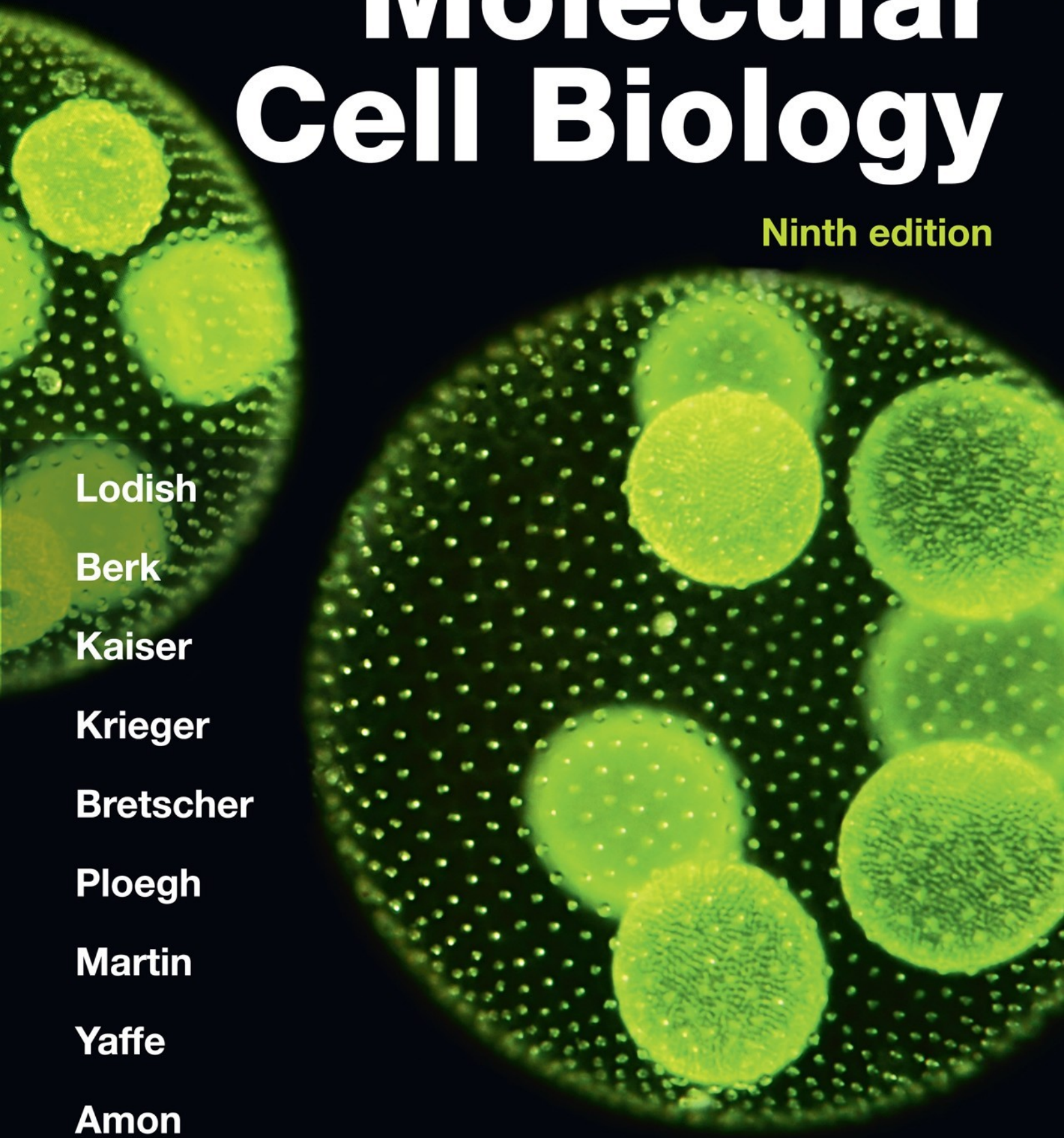
Molecular Cell Biology

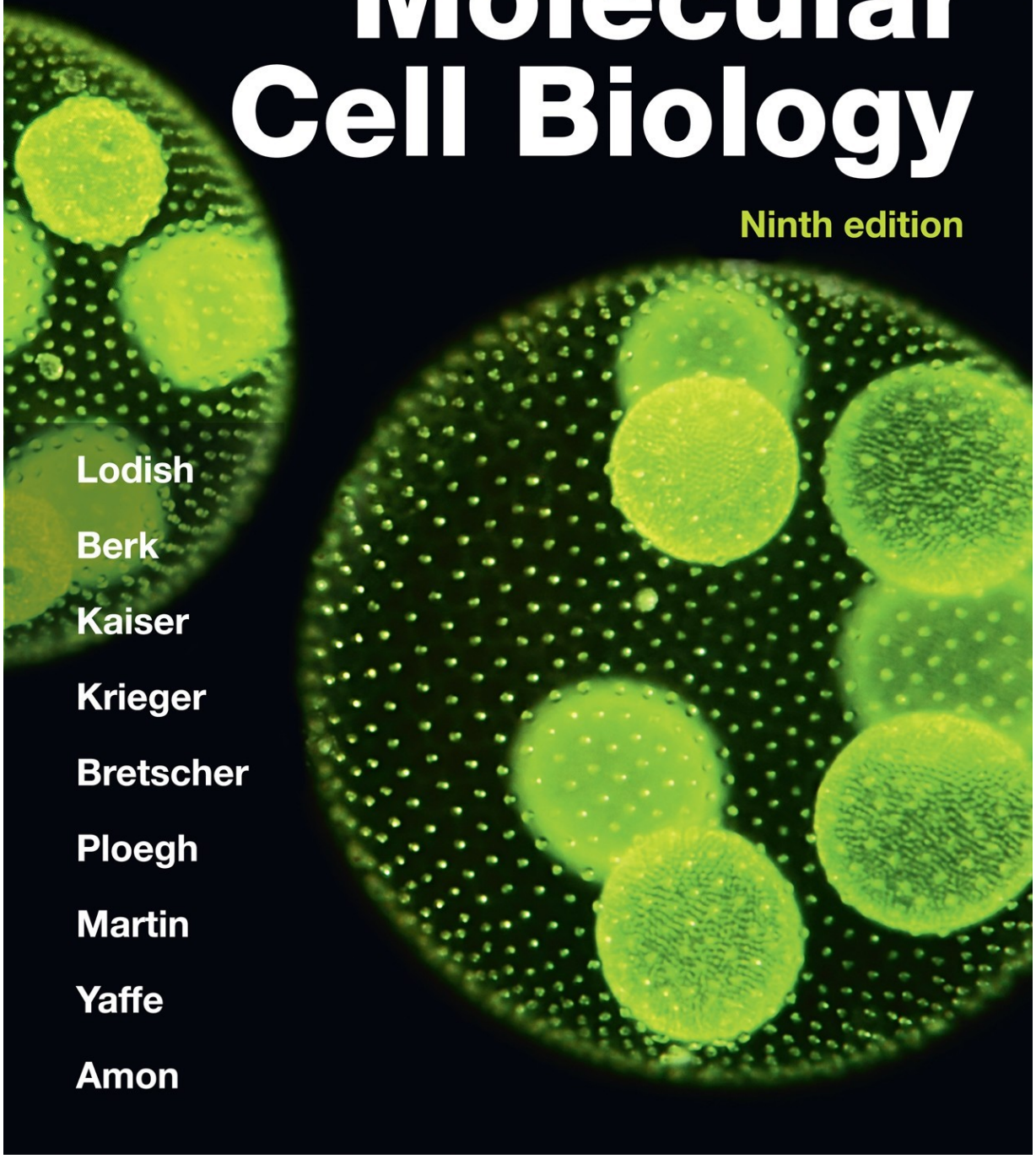
Ninth edition



Lodish
Berk
Kaiser
Krieger
Bretscher
Ploegh
Martin
Yaffe
Amon

Molecular Cell Biology

Ninth edition



Lodish
Berk
Kaiser
Krieger
Bretscher
Ploegh
Martin
Yaffe
Amon

Molecular Cell Biology

Molecular Cell Biology

NINTH EDITION

Harvey Lodish

Arnold Berk

Chris A. Kaiser

Monty Krieger

Anthony Bretscher

Hidde Ploegh

Kelsey C. Martin

Michael B. Yaffe

Angelika Amon

macmillan international
HIGHER EDUCATION

Senior Vice President, STEM: Daryl Fox
Executive Program Director, Life Sciences: Sandy Lindelof
Executive Program Manager: Lisa Lockwood
Executive Marketing Manager: Will Moore
Executive Content Development Manager, STEM: Debbie Hardin
Developmental Editors: Anna Bristow, Maria Lokshin, Susan Moran, Michael Zierler
Executive Project Manager, Content, STEM: Katrina Mangold
Editorial Project Manager: Karen S. Misler
Director of Content, Life Sciences: Jennifer Driscoll Hollis
Senior Media Editor: Cassandra Korsvik
Editorial Assistant: Nathan Livingston
Marketing Assistant: Morgan Psiuk
Director of Content Management Enhancement: Tracey Kuehn
Senior Managing Editor: Lisa Kinne
Senior Content Project Manager: Peter Jacoby
Senior Workflow Project Manager: Paul Rohloff
Director of Design, Content Management: Diana Blume
Design Services Manager: Natasha Wolfe
Cover Design Manager: John Callahan
Cover Designer: Laura De Grasse
Text Designer: Lumina Datamatics, Inc.
Art Manager: Matthew McAdams
Illustrations: Emiko Paul, Adam Steinberg, Troutt Visual Services

Senior Director, Digital Production: Keri deManigold
Senior Media Project Manager: Daniel Comstock
Permissions Manager: Jennifer MacMillan
Photo Editors: Krystyna Borgen, Lumina Datamatics, Inc.
Text Permissions: Michael McCarty
Production Supervisor: Robert Cherry
Composition: Lumina Datamatics, Inc.

ISBN-13: 978-1-319-38360-2 (ePub)

© 2021, 2016, 2013, 2008 by W. H. Freeman and Company

All rights reserved.

1 2 3 4 5 6 25 24 23 22 21 20

Macmillan Learning
One New York Plaza
Suite 4600
New York, NY 10004-1562
www.macmillanlearning.com

TO OUR STUDENTS AND TO OUR TEACHERS,

from whom we continue to learn,

AND TO OUR FAMILIES,

for their support, encouragement, and love

About the Authors

Photo credit: John Soares

HARVEY LODISH is Professor of Biology and Professor of Biological Engineering at the Massachusetts Institute of Technology and a Founding Member of the Whitehead Institute for Biomedical Research. Dr. Lodish is also a member of the National Academy of Sciences and the American Academy of Arts and Sciences and was President (2004) of the American Society for Cell Biology. He is well known for his work on cell-membrane physiology, particularly the biosynthesis of many cell-surface proteins, and on the cloning and functional analysis of several cell-surface receptor proteins, such as the erythropoietin and TGF-β receptors. His laboratory also studies long noncoding RNAs and micro-RNAs that regulate the development and function of hematopoietic cells and adipocytes.

Dr. Lodish teaches undergraduate and graduate courses in cell biology and biotechnology.

Photo credit: Penny Jennings/UCLA Department of Chemistry & Biochemistry

ARNOLD BERK holds the UCLA Presidential Chair in Molecular Cell Biology in the Department of Microbiology, Immunology, and Molecular Genetics and is a member of the Molecular Biology Institute at the University of California, Los Angeles. Dr. Berk is also a fellow of the American Academy of Arts and Sciences. He is one of the discoverers of RNA splicing and of mechanisms for gene control in viruses. His laboratory studies the molecular interactions that regulate transcription initiation in mammalian cells, focusing in particular on adenovirus regulatory proteins. He teaches an advanced undergraduate course in cell biology of the nucleus and a graduate course in the biochemistry of gene expression.

Photo credit: Chris Kaiser

CHRIS A. KAISER is the Amgen Inc. Professor in the Department of Biology at the Massachusetts Institute of Technology. He is also a former Department Head and former Provost at MIT. His laboratory uses genetic and cell biological methods to understand how newly synthesized membrane and secretory proteins are folded and stored in the compartments of the secretory pathway. Dr. Kaiser is recognized as a top educator at MIT, where he has taught genetics to undergraduate and graduate students for many years.

Photo credit: Monty Krieger

MONTY KRIEGER is the Whitehead Professor in the Department of Biology at the Massachusetts Institute of Technology and a Senior Associate Member of the Broad Institute of MIT and Harvard. Dr. Krieger is also a member of the National Academy of Sciences. For his innovative teaching of undergraduate biology and human physiology as well as graduate cell biology courses, he has received numerous awards. His laboratory has made contributions to our understanding of membrane trafficking through the Golgi apparatus and has cloned and characterized receptor proteins important for pathogen recognition and the movement of cholesterol into and out of cells, including the HDL receptor.

Photo credit: Anthony Bretscher

ANTHONY BRETSCHER is Professor of Cell Biology at Cornell University and a member of the Weill Institute for Cell and Molecular Biology. His laboratory is well known for identifying and characterizing new components of the actin cytoskeleton and elucidating their biological functions in relation to cell polarity and membrane traffic. For this work, his laboratory exploits biochemical, genetic, and cell biological approaches in two model systems, vertebrate epithelial cells and the budding yeast. He is a fellow of the American Academy of Arts and Sciences. Dr. Bretscher teaches cell biology to undergraduates at Cornell University.

Photo credit: Hidde Ploegh

HIDDE PLOEGH is a senior investigator in the Program of Cellular and Molecular Medicine at Boston Children's Hospital, where he studies the biochemistry of the immune system. A member of the National Academy of Science and the American Academy of Arts and Sciences, he is one of the world's leading researchers in the molecular understanding of immune system cells and the mechanisms by which viruses evade detection by the immune system. Prior to his current position he was a professor at the Massachusetts Institute of Technology and Harvard Medical School, where he taught immunology and cell biology.

Photo credit: Phuong Pham

KELSEY C. MARTIN is Professor of Biological Chemistry and Psychiatry and Dean of the David Geffen School of Medicine at the University of California, Los Angeles. She is the former Chair of the Biological Chemistry Department. Her laboratory studies the ways in which experience changes connections between neurons in the brain to store long-term memories — a process known as *synaptic plasticity*. She has made important contributions to elucidating the molecular and cell biological mechanisms that underlie this process. Dr. Martin teaches basic principles of neuroscience to undergraduates, graduate students, dental students, and medical students.

Photo credit: Michael B. Yaffe, Koch Institute for Integrative Cancer Research

MICHAEL B. YAFFE is the David H. Koch Professor of Science in the Departments of Biology and Biological Engineering at the Massachusetts Institute of Technology and a Senior Associate Member of the Broad Institute of MIT and Harvard. He is also the Academic Editor-in-Chief of the journal *Science Signaling*. His laboratory studies the cellular response to stress and injury, and is well known for the discovery and characterization of modular protein domains and protein kinases that form molecular signaling circuits. Dr. Yaffe teaches undergraduate biochemistry and graduate cell biology courses, for which he has received numerous awards. He also instructs medical students and residents in the intensive care unit at the Beth Israel Deaconess Medical Center, where he is an attending physician.

Photo credit: Pamela DiFraia/Koch Institute/MIT

ANGELIKA AMON is Professor of Biology at the Massachusetts Institute of Technology, a member of the Koch Institute for Integrative Cancer Research, and Investigator at the Howard Hughes Medical Institute. She is also a member of the National Academy of Sciences. Her laboratory studies the molecular mechanisms that govern chromosome segregation during mitosis and meiosis and the consequences — aneuploidy — when these mechanisms fail during normal cell proliferation and cancer development. Dr. Amon teaches undergraduate and graduate courses in cell biology and genetics.

The authoritative cell biology resource, now available with Achieve

Molecular Cell Biology remains the most authoritative cell biology textbook. The ninth edition is now available in Macmillan's new online learning platform, **Achieve**. The author team, consisting of world-class researchers and professors, has incorporated numerous medically relevant examples to illustrate connections between cell biology and human health and disease. Experimental techniques that drive advances in biomedical sciences are emphasized, introducing students to cutting-edge research that teaches them the skills they need for their careers.

Overview of key features

The ninth edition of ***Molecular Cell Biology*** is now available in **Achieve**. Achieve supports educators and students throughout the full range of instruction, including assets suitable for pre-class preparation, in-class active learning, and post-class study and assessment. The pairing of a powerful new platform with

outstanding biology content provides an unrivaled learning experience.

FEATURES OF ACHIEVE INCLUDE

- A **design guided by learning science research**. Developed with extensive collaboration and testing by students and faculty, including two levels of Institutional Review Board approval for every efficacy study of Achieve.
- A **learning path** of powerful content including pre-class, in-class, and post-class activities and assessments.
- A **detailed gradebook with insights for just-in-time teaching** and reporting on student achievement by learning concept.
- Easy **integration and gradebook synchronization with iClicker** classroom engagement solutions.
- Simple **integration with your campus LMS** and **availability through Inclusive Access** programs.

NEW MEDIA AND ASSESSMENT FEATURES IN ACHIEVE INCLUDE

- Five new **in-class activities**, along with revised **classroom response questions** support an active and inclusive classroom.
- **LearningCurve adaptive quizzing** provides formative assessments to help students prepare before class.
- **Case Studies** and **Analyze the Data Problems** focus on real-world applications.

- **Chapter Concept Quizzes** provide easy to assign and automatically graded assessments.

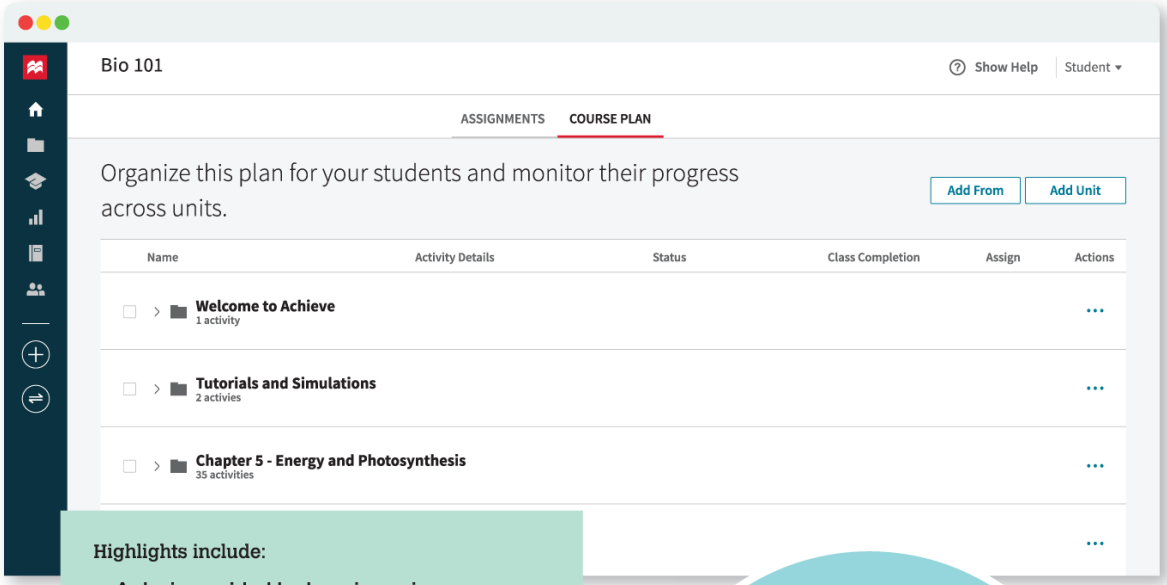
FEATURES OF THE TEXT INCLUDE

- The text and figures **emphasize the experiments** that define our understanding of cell biology, engaging students with the exciting breakthroughs that define the field's history and point to its future.
- **Experimental Figures** illustrate important techniques and experiments and show real data.
- **Medical Connections** are woven throughout the chapters to give students an appreciation for the clinical applications of the basic science they are learning.
- In many chapters, **Plant Biology Connections** spotlight plant biotechnology and unique aspects of cell biology in plants.
- **Key Concepts** for each section provide a learning framework for students as they progress through a chapter.

Achieve

Achieve is the culmination of years of development work put toward creating the most powerful online learning tool for biology students. It houses all of our renowned assessments, multimedia assets, e-books, and instructor resources in a powerful new platform.

Achieve supports educators and students throughout the full range of instruction, including assets suitable for pre-class preparation, in-class active learning, and post-class study and assessment. The pairing of a powerful new platform with outstanding biology content provides an unrivaled learning experience.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Highlights include:

- **A design guided by learning science research.** Co-designed through extensive collaboration and testing by both students and faculty including two levels of Institutional Review Board approval for every study of Achieve.
- **A learning path of powerful content** including pre-class, in-class, and post-class activities and assessments. A detailed gradebook with insights for just-in-time teaching and reporting on student achievement by learning objective.
- **Easy integration and gradebook sync** with iClicker classroom engagement solutions.
- **Simple integration** with your campus LMS and availability through **Inclusive Access** programs.

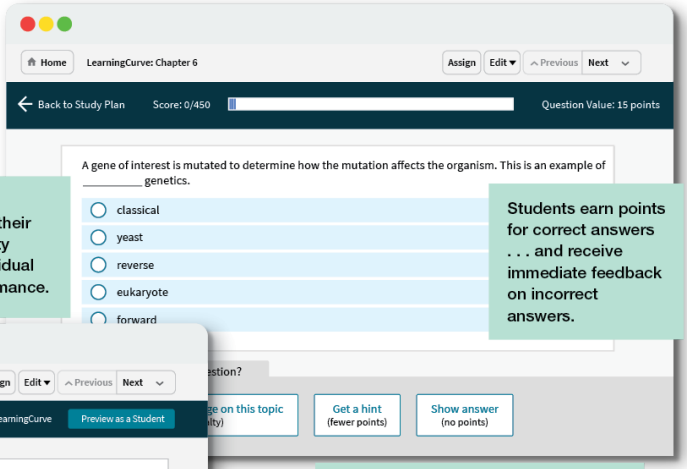
For more information or to sign up for a demonstration of Achieve, contact your local Macmillan representative or visit macmillanlearning.com/achieve.

Full Learning Path

Achieve includes a full learning path of content for pre-class preparation, in-class active learning, and post-class engagement and assessment. It's the perfect solution for engaging all students in an inclusive teaching experience.

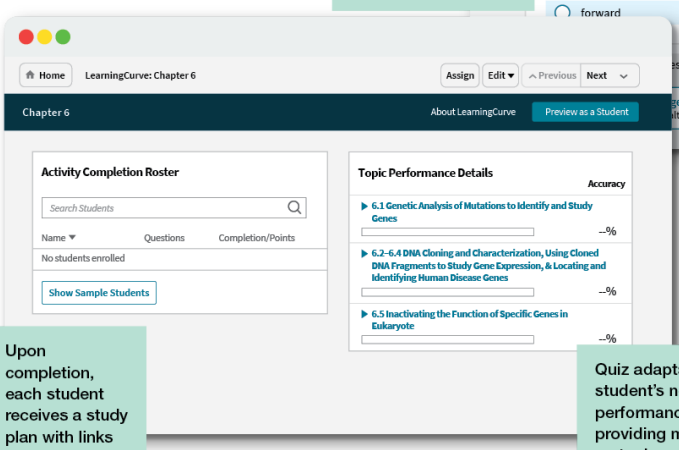
LearningCurve

LearningCurve's game-like quizzing motivates each student to engage with the course content, and reporting tools help teachers get a handle on what their class needs.



Book-specific quizzes adapt their level of difficulty based on individual student performance.

Students earn points for correct answers . . . and receive immediate feedback on incorrect answers.

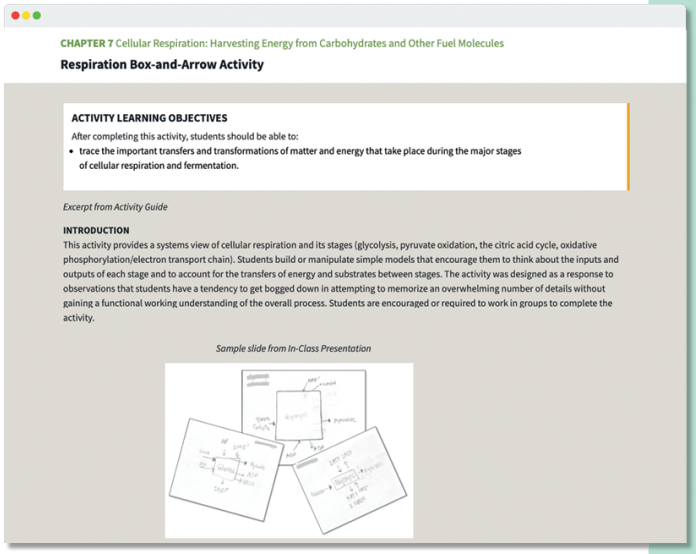


Questions are tagged to sections of the e-book to provide comprehensive coverage and easy-to-find help.

Upon completion, each student receives a study plan with links to additional study tools.

Quiz adapts to each student's needs based on performance, automatically providing more questions on topics where the student is struggling.

Performance by Learning Objectives can also be tracked to give instructors actionable information about topics that need extra emphasis.



In-Class Activities

Active learning exercises provide students with hands-on exploration and help correct common misconceptions.

The activities have been vetted in the classroom and serve as a great off-the-shelf option for incorporating new methods of active learning in your classroom.

The activities include everything you need, including an instructor's guide, presentation decks, and student handouts.

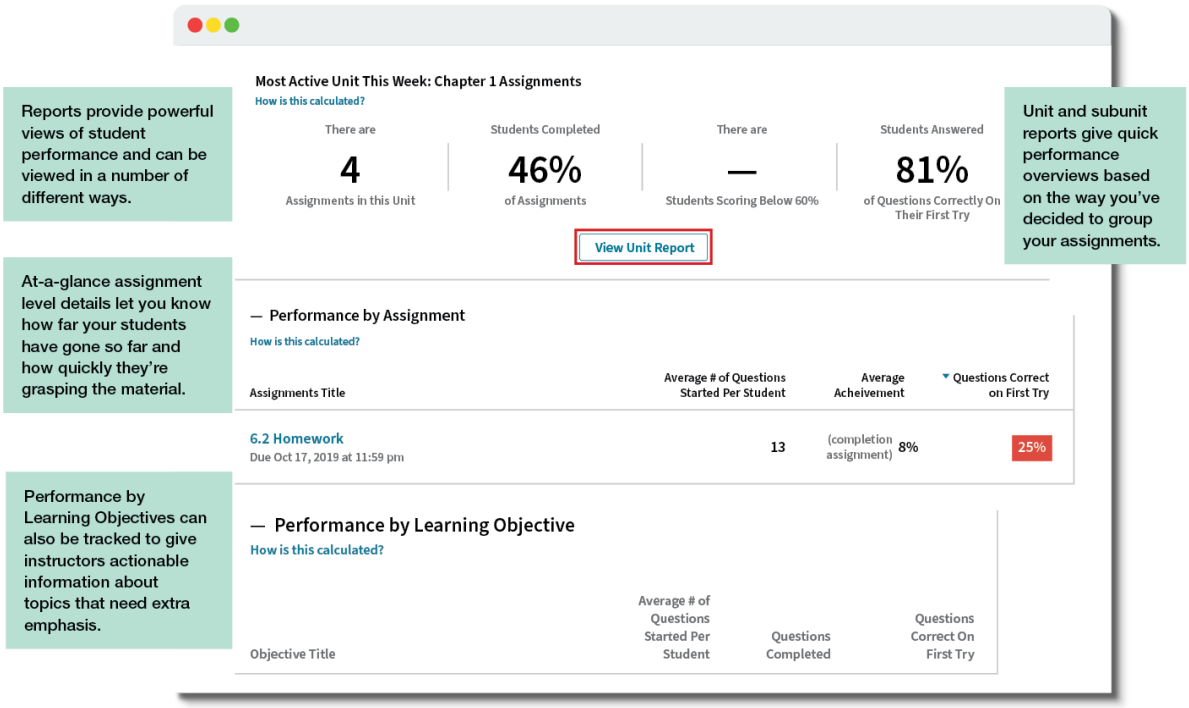
iClicker

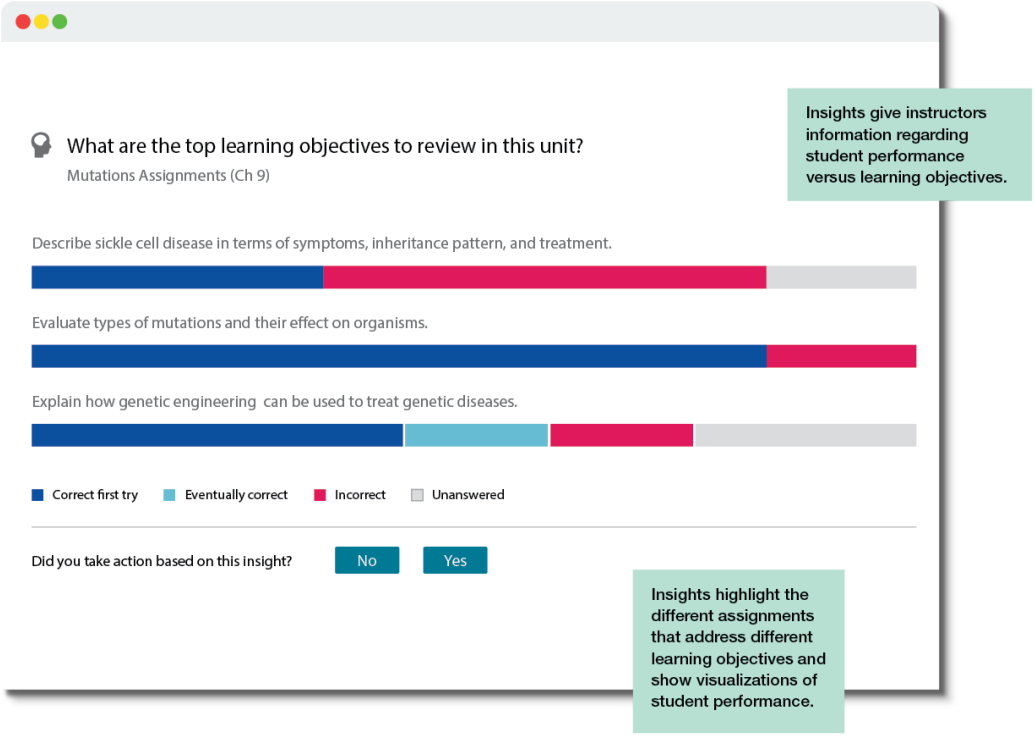
Achieve includes gradebook integration with iClicker, allowing for a seamless pairing of the best classroom response system with the best in-class content.

Powerful analytics, viewable in an elegant dashboard, offer instructors a window into student progress. Achieve gives you

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

the insight to address students' weaknesses and misconceptions before they struggle on a test.





Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

A text grounded in experimentation and research

Molecular Cell Biology **emphasizes the experiments** that define our understanding of cell biology, engaging students with the exciting breakthroughs that define the field's history and point to its future.

There is a good balance between the experimental evidence and the concepts. Many times if a student can understand how the information was gained, it helps them understand the material.

— Jamie Snyder, *California State Polytechnic University, Pomona*

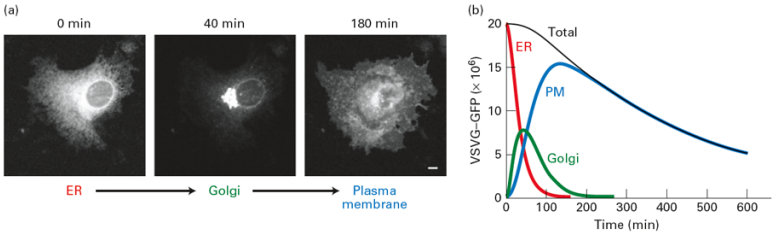
transfer to later compartments can be followed over time. Second, it is necessary to have a way to identify the compartment in which a labeled protein resides. Here we describe two modern experimental procedures for observing the intracellular trafficking of a secretory protein in almost any type of cell.

In both procedures, a gene encoding an abundant membrane glycoprotein (G protein) from vesicular stomatitis virus (VSV) is introduced into cultured mammalian cells either by transfection or simply by infecting the cells with the virus. The transfected cells efficiently produce the VSV G protein, which is then inserted into the membrane of the ER. The use of a mutant gene encoding a temperature-sensitive VSV G protein allows researchers to turn subsequent transport of this protein on and off. At the restrictive temperature of 40 °C, newly made VSV G protein is misfolded and therefore retained within the ER by the quality-control mechanisms discussed in Chapter 13, whereas at the permissive temperature of 32 °C, the protein is correctly folded and transported through the secretory pathway to the cell surface. Importantly, the misfolding of the temperature-sensitive VSV G protein is reversible; thus when cells synthesizing mutant VSV G protein are grown at 40 °C and then shifted to 32 °C, the misfolded mutant VSV G protein that had accumulated in the ER will refold and be transported normally. This clever use of a temperature-sensitive mutation in effect defines a protein cohort whose subsequent transport can be followed.

In two variations of this basic procedure, transport of VSV G protein is monitored by different techniques. Studies using both of these modern trafficking assays came to the same conclusion as Palade's early experiments: in mammalian cells, vesicle-mediated transport of a protein molecule from its site of synthesis on the rough ER to its arrival at the plasma membrane takes from 30 to 60 minutes.

Microscopy of GFP-Labeled VSV G Protein One approach for observing the transport of VSV G protein employs a hybrid gene in which the viral gene is fused to the gene encoding green fluorescent protein (GFP), a naturally occurring fluorescent protein (see Chapter 4). The hybrid gene is transfected into cultured cells by techniques described in Chapter 6. When cells expressing the temperature-sensitive form of the hybrid protein (VSVG-GFP) are grown at the restrictive temperature, VSVG-GFP accumulates in the ER, which appears as a lacy network of membranes when the cells are observed in a fluorescent microscope. When the cells are subsequently shifted to a permissive temperature, the VSVG-GFP can be seen to move first to the membranes of the Golgi complex, which are densely concentrated at the edge of the nucleus, and then to the cell surface (Figure 14-3a). By observing the distribution of VSVG-GFP at different times after shifting cells to the permissive temperature, researchers have determined how long VSVG-GFP resides in each organelle of the secretory pathway (Figure 14-3b).

Detection of Compartment-Specific Oligosaccharide Modifications A second way to follow the transport of secretory proteins takes advantage of modifications to their carbohydrate side chains that occur at different stages of the secretory pathway. To understand this approach, recall that many secretory proteins leaving the ER are carrying one or more copies of the *N*-linked oligosaccharide $Man_8(GlcNAc)_2$, which are synthesized and attached to secretory proteins in the ER (see Figure 13-18). As a protein moves through the Golgi complex, different enzymes localized to the *cis*-, *medial*-, and *trans*-Golgi cisternae catalyze an ordered series of modifications to these core $Man_8(GlcNAc)_2$ chains, as discussed in a later section of this chapter. For instance, glycosidases that reside specifically in the *cis*-Golgi compartment



EXPERIMENTAL FIGURE 14-3 Protein transport through the secretory pathway can be visualized by fluorescence microscopy of cells producing a GFP-tagged membrane protein. Cultured cells were transfected with a hybrid gene encoding the viral membrane glycoprotein VSV G linked to the gene for green fluorescent protein (GFP). A temperature-sensitive mutant version of the viral gene was used so that newly made hybrid protein (VSVG-GFP) was retained in the ER at 40 °C but was released for transport at 32 °C. (a) Fluorescence micrographs of cells just before and at two times after they were shifted to the lower temperature. Movement of VSVG-GFP from the ER to the Golgi and finally to the cell surface occurred within 180 minutes. The scale bar is 5 μm. (b) Plot of the amount of VSVG-GFP in the endoplasmic reticulum (ER), Golgi, and plasma membrane (PM) at different times after the shift to the permissive temperature. The kinetics of transport from one organelle to another can be reconstructed from computer analysis of these data. The decrease in total fluorescence that occurs at later times probably results from slow inactivation of GFP fluorescence. [Jennifer Lippincott-Schwartz and Koret Hirschberg, Metabolism Branch, National Institute of Child Health and Human Development.]

Relevant connections for students

Medical Connections are woven throughout the chapters to give students an appreciation for the clinical applications of the basic science they are learning.

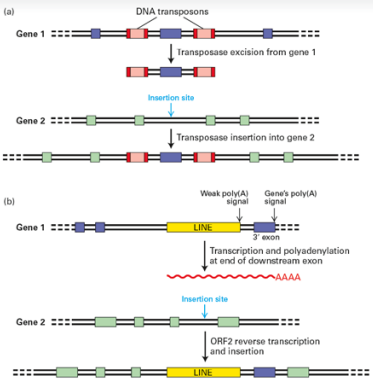


FIGURE 7-19 Exon shuffling by transposition. (a) Transposition of an exon flanked by homologous DNA transposons into an intron on a second gene. As we saw in Figure 7-10, step 1, transposase can recognize and cleave the DNA at the ends of the transposon inverted repeats. In gene 1, if the transposase cleaves at the left end of the transposon on the left and at the right end of the transposon on the right, it can transpose all the intervening DNA, including the exon from gene 1, to a new site in an intron of gene 2. The net result is an insertion of the exon from gene 1 into gene 2. (b) Integration of an exon into another gene via LINE transposition. Some LINEs have weak poly(A) signals. If such a LINE is in the 3′-most intron of gene 1, during transposition its transcription may continue beyond its own poly(A) signals and extend into the 3′ exon, transcribing the cleavage and polyadenylation signals of gene 1 itself. This RNA can then be reverse-transcribed and integrated by the LINE ORF2 protein (see Figure 7-17) into an intron on gene 2, introducing a new 3′ exon (from gene 1) into gene 2.

combined effects of several enhancer elements. Insertion of mobile elements near such transcription-control regions probably contributed to the evolution of new combinations of enhancer sequences. These combinations, in turn, control which specific genes are expressed in particular cell types and the amount of the encoded protein produced in modern organisms, as we discuss in the next chapter.

These considerations suggest that the early view of mobile DNA elements as completely selfish molecular parasites misses the mark. Rather, these elements have contributed profoundly to the evolution of multicellular organisms by promoting (1) the generation of gene families via gene duplication, (2) the creation of new genes via shuffling of preexisting exons, and (3) the formation of more complex regulatory regions that provide multifaceted control of gene expression. Today researchers are attempting to harness transposition mechanisms to insert therapeutic genes into patients as a form of gene therapy.

A process analogous to that shown in Figure 7-19a is largely responsible for the rapid spread of antibiotic resistance among pathogenic bacteria, a major problem in modern medicine. Bacterial genes encoding enzymes that inactivate antibiotics (drug resistance genes) have become flanked by insertion sequences, generating drug resistance transposons. The widespread use of antibiotics in medicine and agriculture has led to positive selection on drug resistance transposons. These transposons are often found incorporated into conjugating plasmids. Conjugating plasmids encode proteins that result in their replication and transfer to other bacterial cells—even cells of other related bacterial species—through a complex macromolecular tube called a pilus. These plasmids, called *R factors* (for drug resistance), can carry multiple drug resistance genes introduced by transposition and selected in environments, such as hospitals, where antibiotic-laced cleaning solutions are used to sterilize surfaces. R factors have led to the rapid spread of resistance to multiple antibiotics among pathogenic bacteria. Coping with the spread of R factors is a major challenge for modern medicine. ■

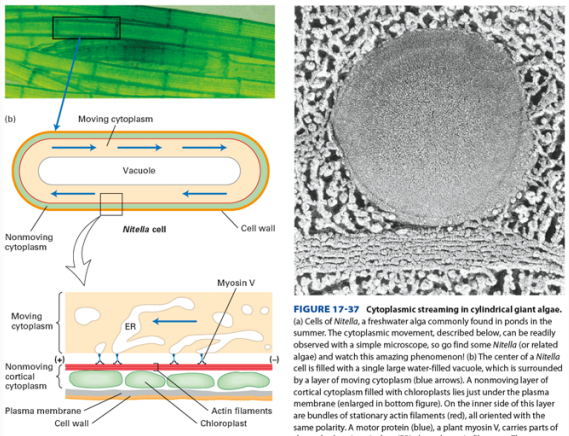


FIGURE 17-37 Cytoplasmic streaming in cylindrical giant algae. (a) Cells of *Nitella*, a freshwater alga commonly found in ponds in the summer. The cytoplasmic movement, described below, can be readily observed with a simple microscope, so go find some *Nitella* (or related algae) and watch this amazing phenomenon! (b) The center of a *Nitella* cell is filled with a single large water-filled vacuole, which is surrounded by a layer of moving cytoplasm (blue arrows). A nonmoving layer of cortical cytoplasm filled with chloroplasts lies just under the plasma membrane (enlarged in bottom figure). On the inner side of this layer are bundles of stationary actin filaments (red), all oriented with the same polarity. A motor protein (blue), a plant myosin V, carries parts of the endoplasmic reticulum (ER) along the actin filaments. The movement of the ER network propels the entire viscous cytoplasm, including organelles that are enmeshed in the ER network. (c) Electron micrograph of the cortical cytoplasm showing a large vesicle connected to an underlying bundle of actin filaments. [Part (c) Courtesy Bechara Kachar.]

so contraction is slower in smooth muscle than in skeletal muscle. Because this regulation involves myosin, it is known as *thick-filament regulation*.

Unlike skeletal muscle, which is stimulated to contract solely by nerve impulses, smooth muscle cells and nonmuscle cells are regulated by many types of external signals. For example, norepinephrine, angiotensin, endothelin, histamine, and other signaling molecules modulate or induce the contraction of smooth muscle or elicit changes in the shape and adhesion of nonmuscle cells by triggering various signal transduction pathways. Some of these pathways lead to an increase in the cytosolic Ca^{2+} concentration, which can stimulate myosin activity by activating MLC kinase (see Figure 17-35). As we will see in Section 17.7, other signaling pathways activate *Rho kinase*, which also activates myosin activity by phosphorylating the regulatory light chain, but in a Ca^{2+}-independent manner.

Perhaps the most dramatic example of movement by myosin V is seen in giant green algae, such as *Nitella* and *Chara*. These algae can be found in ponds during the summer and their movement easily observed using a simple microscope. In their large cells, which can be as much as 2 cm in length, cytosol flows rapidly, at a rate approaching 4.5 mm per minute, in an endless loop around the inner circumference of the cell (Figure 17-37). This *cytoplasmic streaming* is a principal mechanism for distributing cellular metabolites and is also seen in other large cells such as plant cells and amoebae. The algal cells have bundles of actin filaments aligned along the length of the cell, lying just above the stationary chloroplasts located adjacent to the membrane. The bulk cytosol is propelled by myosin V (also known as *myosin XI* in plants) attached to parts of the ER adjacent to the actin filaments. The flow rate of the cytosol in *Nitella* is about 15 times faster than movement produced by any other myosins. ■

Plant Biology Connections spotlight plant biotechnology and unique aspects of cell biology in plants.

"

I love the many great medical examples highlighted explicitly by "medical icons," including Ras in oncogenic transformation, HER2 in breast cancer, "bubble boy," etc. I have a lot of premed students who take my Cell Bio class, and they would love this.

— Adam Bohnert, *Louisiana State University*

Preface

In writing the ninth edition of *Molecular Cell Biology*, we have incorporated many of the spectacular advances made over the past four years in biomedical science, driven in part by new experimental technologies that have revolutionized many fields. Fast techniques for sequencing DNA, allied with efficient methods to generate and study mutations in model organisms and to map disease-causing mutations in humans, have illuminated a basic understanding of the functions of many cellular components, including hundreds of human genes that affect diseases such as diabetes and cancer.

Exploring the most current developments in the field is always a priority in writing a new edition, but it is also important to us to communicate the basics of cell biology clearly by stripping away as much extraneous detail as possible to focus attention on the fundamental concepts of cell biology. To this end, in addition to introducing new discoveries and technologies, we have streamlined and reorganized several chapters to clarify processes and concepts for students.

New Co-Author, Michael B. Yaffe

The ninth edition of *Molecular Cell Biology* introduces a new member to our author team. Dr. Yaffe is the David H. Koch Professor of Science in the Departments of Biology and Biological Engineering at the Massachusetts Institute of Technology and a Senior Associate Member of the Broad Institute of MIT and Harvard. He is also the Academic Editor-in-Chief of the journal *Science Signaling*. His laboratory studies the cellular response to stress and injury and is well known for the discovery and characterization of modular protein domains and protein kinases that form molecular signaling circuits. Dr. Yaffe teaches undergraduate biochemistry and graduate cell biology courses, for which he has received numerous awards. He also instructs medical students and residents in the intensive care unit at the Beth Israel Deaconess Medical Center, where he is an attending physician.

Revised, Cutting-Edge Content

The ninth edition of *Molecular Cell Biology* includes new and improved chapters:

- "Fundamental Molecular Genetic Mechanisms" (Chapter 5) has been reorganized to begin with the structure and replication of DNA, followed by DNA repair and recombination. A new discussion of the information-carrying capacity of DNA has been added, as well as using the genetic code to predict the effect of mutations.

- “Genes, Chromatin, and Chromosomes” (Chapter 7) has been renamed to reflect our much greater understanding of the roles chromatin plays in gene expression. Section 7.4 provides details of the molecular differences between euchromatin and heterochromatin, the processes that occur as regions of chromatin shift from one form to the other, and the protein modifications and signaling pathways that cause these shifts to occur.
- “Biomembrane Structure” (Chapter 10) has been moved later in Part II of the book (previously Chapter 7) to directly precede “Transmembrane Transport of Ions and Small Molecules” and refine the progression of Part II.
- “Cellular Energetics” (Chapter 12) has been revised to approach energetics from the conceptual point of optimizing output. New discussions have been added on intra- and intercellular mitochondrial movements as well as membrane contact sites (MCSs).
- “The Eukaryotic Cell Cycle” (Chapter 19) now opens with an overview of the cell cycle that is presented in a story-like manner, outlining the flow from G_1 phase around the cycle to M phase and back into G_1. Section 19.3 presents our latest understanding about feedback loops and protein modifications that control the cell cycle. Figures have been extensively revised and many new figures have been added, providing overviews of the tiered regulatory processes that control cyclin-dependent kinases and showing clearly how protein complexes perform the many carefully controlled steps of the cell cycle.

- "Responding to the Cellular Environment" (Chapter 21) is a new chapter. It covers many important signaling pathways by which cells respond to changes of nutrients in their environment — glucose, amino acids, cholesterol, and oxygen — as well as to changes in temperature and contact with the extracellular matrix and other cells. Circadian rhythms are also covered.

New Discoveries, New Methodologies

- *Schizosaccharomyces pombe* (cell cycle control) (Ch. 1)
- Neanderthal evolution and genomics (Ch. 1)
- Single-cell RNA sequencing and identification of new types of cells (Chs. 1 and 6)
- Increased focus on regulation of protein activity (Ch. 3)
- Increased coverage of kinases: family tree, kinase structure, and regulation (Ch. 3)
- Recent breakthroughs in protein structure prediction (Ch. 3)
- Biomolecular condensates, including methods to study and identify them (Ch. 3)
- Updated mechanisms and functions of chaperones and proteasomes (Ch. 3)
- iTRAQ and TMT methods for simultaneous, multiplex analysis of samples (Ch. 3)
- Updated resolution of cryoEM (Ch. 3)
- Use of enzyme-catalyzed, proximity-dependent labeling coupled with mass spectrometry for localization of proteins in cells (Ch. 3)
- Generation and use of organoids (Ch. 4)

- Use of optogenetics (Ch. 4)
- New discussion of information-carrying capacity of DNA (Ch. 5)
- New discussion of reading genetic code in predicting the effect of mutations (Ch. 5)
- New discussion of methylation in distinguishing parent and daughter strands (Ch. 5)
- New section on global analysis of double mutant combinations that reveal networks of gene functions (Ch. 6)
- Expanded discussion of cDNA libraries and their limitations (Ch. 6)
- New section on medically important genes identified by alleles that protect against disease (Ch. 6)
- New section on identification of causative mutations in cancer cells (Ch. 6)
- Updated discussion of CRISPR-Cas genome editing (Ch. 6)
- Revised section on chromatin structure (Ch. 7)
- New section on chromatin packing in the nucleus and updated information about SMC protein complexes' role in chromatin compaction (Ch. 7)
- Cryoelectron microscopy models of the preinitiation complex (Ch. 8)
- New section on transcriptional condensates (Ch. 8)
- New section on transcription occurring in bursts (Ch. 8)
- Updated discussion of spliceosome (Ch. 9)
- New section on abnormal RNA splicing and disease (Ch. 9)
- Expanded and updated discussion of V-class proton pumps (Ch. 11)
- Updated discussion of K^+ channels (Ch. 11)

- Expanded discussion on bacterial Na^+ galactose symporter (Ch. 11)
- Expanded discussion of Anion Exchanger 1 (Ch. 11)
- New and updated discussion of mitochondrial structure and dynamics (e.g., fusion and fission) based on state-of-the-art super-resolution microscopy (e.g., grazing incidence-structured illumination microscopy) (Ch. 12)
- Updates on relationship of mitochondria to Parkinson's disease (Ch. 12)
- New discussion of membrane contact sites (Ch. 12)
- Updated discussion of structure of electron-transport chain supercomplexes (Ch. 12)
- Reorganized and updated coverage of photosynthesis (Ch. 12)
- Updated discussion on mechanisms that protect cells from damage by reactive oxygen species, such as singlet oxygen, during photoelectron transport (Ch. 12)
- Families of human G protein–coupled receptors (Ch. 15)
- Gene therapy to correct a form of human blindness (Ch. 15)
- Structure and activation of the insulin receptor (Ch. 16)
- Interleukin 2 (IL-2) receptors and gene therapy to correct severe combined immune deficiency (Ch. 16)
- Molecular structure of gamma secretase (Ch. 16)
- Giant cell-surface signalsomes (Ch. 16)
- New description of microtubule nucleator augmin (Ch. 18)
- New section on spindle positioning during mitosis (Ch. 18)
- New section on regulation of the contractile ring during cytokinesis (Ch. 18)

- New sections on cyclin discovery and cyclin-CDK regulation (Ch. 19)
- Thiol-labeled γ phosphate-based approaches to identify CDK substrates with genetic engineering (Ch. 19)
- Updated discussion of SCF ubiquitin ligases and the anaphase-promoting complex (Ch. 19)
- New sections on how feedback loops and phosphoserine/threonine-binding domains control the cell cycle (Ch. 19)
- Expanded discussion of Aurora and Polo-like kinases (Ch. 19)
- New section on centrosome duplication (Ch. 19)
- Extensive updated coverage of cytokinesis (Ch. 19)
- Updated discussion of DNA damage and cell cycle surveillance (Ch. 19)
- Protocadherin mechanism of adhesion and role in bar-coding neurons (Ch. 20)
- Updated discussion of mechanotransduction (Ch. 20)
- Updated discussion of integrin activation, including new discussion of three conformation models (Ch. 20)
- Mechanism of talin-mediated inside-out and outside-in signaling and mechanotransduction (Ch. 20)
- Updated discussion of O-linked glycoproteins and O-mannose linked oligosaccharides (Ch. 20)
- Updated discussion of claudin-mediated paracellular transport (Ch. 20)
- Updated discussion of plant plasmodesmata structure and composition (Ch. 20)

- New section on biomolecular condensates and adhesion (Ch. 20)
- Extensive updates on plant adhesion molecules and mechanotransduction (Ch. 20)
- Updated and expanded discussion of mTOR pathway and integration of cell growth signals with nutrient and energy levels (Ch. 21)
- New section on HIF-1α and the response of cells to low oxygen levels (Ch. 21)
- New section on the heat-shock response (Ch. 21)
- New section on circadian rhythms (Ch. 21)
- Updated and expanded discussion of mechanotransduction (Ch. 21)
- Updated and expanded discussion of cells in intestinal crypts that dedifferentiate to replenish intestinal stem cells (Ch. 22)
- Updated discussion of Cdc42 and feedback loops that control cell polarity (Ch. 22)
- New section on “eat me” phosphatidylserine signals on apoptotic cells (Ch. 22)
- Updated discussion of necroptosis, an inflammatory form of programmed cell death (Ch. 22)
- Multiple revised figures to clearly demonstrate neurons and neural structures (Ch. 23)
- Expanded discussion of the blood-brain barrier (Ch. 23)
- New discussion of Piezo1 and Piezo2 receptors (Ch. 23)
- New section on natural killer (NK) cells, now called *ILCs* (Ch. 24)

- New discussion of herd immunity and the benefits of childhood vaccinations (Ch. 24)
- New section on immunotherapies in the treatment of cancer (Ch. 25)
- Updated discussion of cancer genome sequencing (Ch. 25)

Medical Connections

Many advances in basic cellular and molecular biology have led to new treatments for cancer and other human diseases. Examples of such medical advances are woven throughout the chapters to give students an appreciation for the clinical applications of the basic science they are learning. Many of these applications hinge on a detailed understanding of multiprotein complexes in cells — complexes that catalyze cell movements; regulate DNA transcription, replication, and repair; coordinate metabolism; and connect cells to other cells and to proteins and carbohydrates in their extracellular environment.

- Stereoisomers of small molecules as drugs — sterically pure molecules have different effects from mixtures (Ch. 2)
- Cholesterol is hydrophobic and must be transported by lipoprotein carriers LDL and HDL (Ch. 2)
- Essential amino acids must be provided in livestock feed (Ch. 2)
- Saturated, unsaturated, and trans fats: their molecular structures and nutritional consequences (Ch. 2)
- Protein misfolding and amyloids in neurodegenerative diseases such as Alzheimer's and Parkinson's (Ch. 3)

- Small molecules that inhibit enzyme activity can be used as drugs (aspirin) or in chemical warfare (sarin gas) (Ch. 3)
- Small-molecule inhibitors of the proteasome are used to treat certain cancers (Ch. 3)
- Disruptions of GTPases, GAPs, GEFs, and GDIs by mutations and pathogens cause a wide variety of diseases (Ch. 3)
- 3-D printing technology may be used to grow replacement organs (Ch. 4)
- Monoclonal antibodies can be used as diagnostic and therapeutic tools in medicine (Ch. 4)
- Mutations in mismatch repair proteins lead to hereditary nonpolyposis colorectal cancer (Ch. 5)
- Nucleotide excision-repair proteins were identified in patients with xeroderma pigmentosum (Ch. 5)
- The high-resolution structures of ribosomes are used in the design of new antibiotics (Ch. 5)
- The sickle-cell allele is an example of one that exhibits both dominant and recessive properties depending on the phenotype being examined (Ch. 6)
- DNA microarrays can be useful as medical diagnostic tools (Ch. 6)
- Recombinant DNA techniques are used to mass-produce therapeutically useful proteins such as insulin and G-CSF (Ch. 6)
- Microsatellite repeats have a tendency to expand and can cause neuromuscular diseases such as Huntington disease and myotonic dystrophy (Ch. 7)

- L1 transposable elements can cause genetic diseases by inserting into new sites in the genome (Ch. 7)
- Exon shuffling can result in bacterial resistance to antibiotics, a growing challenge in hospitals (Ch. 7)
- Telomerase is abnormally activated in most cancers (Ch. 7)
- TFIIH subunits were first identified based on mutations in those subunits that cause defects in DNA repair associated with a stalled RNA polymerase (Ch. 8)
- HIV encodes the Tat protein, which inhibits termination of transcription by RNA polymerase II (Ch. 8)
- Synthetic oligonucleotides are being used in treatment of Duchenne muscular dystrophy (Ch. 9)
- Mutations in splicing enhancers can cause exon skipping, as in spinal muscular atrophy (Ch. 9)
- Myelodysplastic syndromes (MDS) are a group of blood cell diseases characterized by deregulated production of abnormal myelocytes (phagocytic white blood cells). Mutations in individual splicing factor genes that change protein function have been identified in MDS patients (Ch. 9)
- Thalassemia commonly results from mutations in globin-gene splice sites that decrease splicing efficiency but do not prevent association of the pre-mRNA with snRNPs (Ch. 9)
- ABO blood types are determined by the carbohydrates attached to glycoproteins on the surfaces of erythrocytes (Ch. 10)
- Atherosclerosis, marked by accumulation of cholesterol, other lipids, and other biological substances in an artery, is responsible for the majority of deaths due to cardiovascular disease in the United States (Ch. 10)

- Aquaporin 2 levels control the rate of water resorption from urine being formed by the kidney (Ch. 11)
- Certain cystic fibrosis patients are being treated with a small molecule that allows a mutant protein to traffic normally to the cell surface (Ch. 11)
- Inhibitors of the human Na^+/glucose symporter SGLT2 are in development or have been approved for treatment of type II diabetes. Antidepressants and other therapeutic drugs, as well as drugs of abuse, target Na^+-powered symporters because of their role in the reuptake and recycling of neurotransmitters (Ch. 11)
- Drugs that inhibit the Na^+/K^+ ATPase in cardiac muscle cells are used in treating congestive heart failure (Ch. 11)
- Rehydration therapy is a simple, effective means of treating cholera and other diseases caused by intestinal pathogens (Ch. 11)
- Mutations in CIC-7, a chloride ion channel, result in defective bone resorption characteristic of the hereditary bone disease osteopetrosis (Ch. 11)
- The sensitivity of mitochondrial ribosomes to the aminoglycoside class of antibiotics, including chloramphenicol, can cause toxicity in patients (Ch. 12)
- Mutations and large deletions in mtDNA cause certain diseases, such as Leber's hereditary optic neuropathy and Kearns-Sayre syndrome (Ch. 12)
- Mutations in two genes encoding the proteins PINK1 and Parkin are responsible for hereditary early onset Parkinson's disease (Ch. 12)

- Cyanide is toxic because it blocks ATP production in mitochondria (Ch. 12)
- Reactive oxygen species are by-products of electron transport that can damage cells (Ch. 12)
- ATP/ADP antiporter activity was first studied over 2000 years ago through the examination of the effects of poisonous herbs (Ch. 12)
- A hereditary form of emphysema results from misfolding of proteins in the endoplasmic reticulum (Ch. 13)
- Autosomal recessive mutations that cause defective peroxisome assembly can lead to several developmental defects often associated with craniofacial abnormalities, such as those associated with Zellweger syndrome (Ch. 13)
- Certain cases of cystic fibrosis are caused by mutations in the CFTR protein that prevent movement of this chloride channel from the ER to the cell surface (Ch. 14)
- Study of lysosomal storage diseases has revealed key elements of the lysosomal sorting pathway (Ch. 14)
- The hereditary disease familial hypercholesterolemia results from a variety of mutations in the *LDLR* gene (Ch. 14)
- EGF stimulates the proliferation of many types of epithelial cells. In about 25 percent of breast cancers, the tumor cells have elevated levels of one particular type of EGF receptor called *HER2* (Ch. 15)
- The agonist isoproterenol binds more strongly to epinephrine-responsive receptors on bronchial smooth muscle cells than does epinephrine and is used to treat bronchial asthma, chronic bronchitis, and emphysema (Ch. 15)

- Some bacterial toxins (e.g., *Bordetella pertussis*, *Vibrio cholerae*, certain strains of *E. coli*) catalyze a modification of a G protein in intestinal cells, increasing intracellular cAMP, which leads to loss of electrolytes and fluids (Ch. 15)
- In 2017, the U.S. Food and Drug Administration approved the first in vivo human gene therapy in the United States to treat blindness based on a defective RPE65 enzyme (Ch. 15)
- Monoclonal antibodies that bind HER2 and thereby block signaling by EGF are useful in treating tumors that overexpress HER2 (Ch. 16)
- Mutant Ras proteins that bind but cannot hydrolyze GTP, and are therefore locked in an active GTP-bound state, contribute to oncogenic transformation (Ch. 16)
- Potent and selective inhibitors of Raf are being clinically tested in patients with melanomas caused by mutant Raf proteins (Ch. 16)
- The deletion of the *PTEN* gene in multiple types of advanced cancers results in the loss of the PTEN protein, contributing to the uncontrolled growth of cells (Ch. 16)
- Epo and G-CSF are used to boost red blood cells in patients with kidney disease and during certain cancer therapies that affect blood cell formation in the bone marrow (Ch. 16)
- Many cases of severe combined immunodeficiency (SCID) result from a deficiency in the IL-2 receptor gamma chain and can be treated by gene therapy (Ch. 16)
- Many tumors contain inactivating mutations in either TGF-β receptors or Smad proteins and are resistant to growth inhibition by TGF-β (Ch. 16)

- High levels of free β-catenin, caused by aberrant hyperactive Wnt signaling, are associated with the activation of growth-promoting genes in many cancers (Ch. 16)
- Inappropriate activation of Hh signaling associated with primary cilia is the cause of several types of tumors (Ch. 16)
- Increased activity of ADAMs can promote cancer development and heart disease (Ch. 16)
- Hereditary spherocytic anemias can be caused by mutations in spectrin, band 4.1, and ankyrin (Ch. 17)
- Duchenne muscular dystrophy affects the protein dystrophin, resulting in progressive weakening of skeletal muscle (Ch. 17)
- Hypertrophic cardiomyopathies result from various mutations in proteins of the heart contractile machinery (Ch. 17)
- Blood tests that measure the level of cardiac-specific troponins are used to determine the severity of a heart attack (Ch. 17)
- Some drugs (e.g., colchicine) bind tubulin dimers and restrain them from polymerizing into microtubules, whereas others (e.g., taxol) bind microtubules and prevent depolymerization (Ch. 18)
- Defects in LIS1 cause Miller-Dieker lissencephaly in early brain development, leading to abnormalities (Ch. 18)
- Some diseases, such as ADPKD and Bardet-Biedl syndrome, have been traced to defects in primary cilia and intraflagellar transport (Ch. 18)
- Keratin filaments are important to maintaining the structural integrity of epithelial tissues by mechanically reinforcing the connections between cells (Ch. 18)

- Mutations in the human gene for lamin A cause a wide variety of diseases termed *laminopathies* (Ch. 18)
- In cohesinopathies, mutations in cohesion subunits or cohesion loading factors disrupt expression of genes critical for development, resulting in limb and craniofacial abnormalities and intellectual disabilities (Ch. 19)
- Mutations in the DNA repair protein BRCA1 underlie familial breast cancer (Ch. 19)
- Aneuploidy leads to misregulation of genes and can contribute to cancer development. Aneuploid eggs are largely caused by chromosome mis-segregation in meiosis I or nondisjunction, leading to miscarriage or Down syndrome (Ch. 19)
- The protein CDHR3 enables class C rhinoviruses (RV-C) to bind to airway epithelial cells, enter them, and replicate, causing respiratory diseases and exacerbating asthma (Ch. 20)
- The cadherin desmoglein is the predominant target of auto-antibodies in the skin disease pemiphigus vulgaris (Ch. 20)
- Some pathogens, such as hepatitis C virus and the enteric bacterium *Vibrio cholerae*, have evolved to exploit the molecules in tight junctions (Ch. 20)
- Mutations in connexin genes cause a variety of diseases (Ch. 20)
- Defects in the glomerular basement membrane can lead to renal failure (Ch. 20)
- In cells deprived of ascorbate, the pro-α collagen chains are not hydroxylated sufficiently to form the structural support of collagen necessary for healthy blood vessels, tendons, and skin, resulting in scurvy (Ch. 20)

- Mutations affecting type I collagen and its associated proteins cause a variety of diseases, including osteogenesis imperfecta (Ch. 20)
- A variety of diseases, often involving skeletal and cardiovascular abnormalities (e.g., Marfan syndrome), result from mutations in the genes encoding the structural proteins of elastic fibers or the proteins that contribute to their proper assembly (Ch. 20)
- Connections between the extracellular matrix and cytoskeleton are defective in muscular dystrophy (Ch. 20)
- Leukocyte-adhesion deficiency is caused by a genetic defect that results in the leukocytes' inability to fight infection, thereby increasing susceptibility to repeated bacterial infections (Ch. 20)
- Diabetes mellitus is characterized by impaired regulation of blood glucose, which can lead to major complications if left untreated (Ch. 21)
- Genes encoding components of the mTORC1 pathway are mutated in many cancers, and mTor inhibitors combined with other therapies may suppress tumor growth (Ch. 21)
- In kidney tumors with a mutation in the von Hippel-Landau (VHL) gene, the loss of VHL protein activity leads to an increase in the expression of Hif-1α target genes, including VEGF (Ch. 21)
- The importance of circadian rhythms ranges from the use of cells in tissue culture to disorders such as Alzheimer's disease and autism spectrum disorder (Ch. 21)
- Because iPS cells can be derived from somatic cells of patients with difficult-to-understand diseases, they have proved

invaluable in uncovering the molecular and cellular basis of several afflictions, including amyotrophic lateral sclerosis (ALS) (Ch. 22)

- The stem cells in transplanted bone marrow can generate all types of functional blood cells, which makes such transplants useful for patients with certain hereditary blood diseases as well as cancer patients who have received irradiation or chemotherapy (Ch. 22)
- Channelopathies, including some forms of epilepsy, are caused by mutations in genes that encode ion channels (Ch. 23)
- The topical anesthetic lidocaine works by binding to amino acid residues along the voltage-gated Na^+ channel, locking it in the open but occluded state (Ch. 23)
- The cause of multiple sclerosis is not known, but seems to involve either the body's production of auto-antibodies that react with myelin basic protein or the secretion of proteases that destroy myelin proteins (Ch. 23)
- Peripheral myelin is a target of autoimmune disease, mainly involving the formation of antibodies against P_0 (Ch. 23)
- The key role of VAMP in neurotransmitter exocytosis can be seen in the mechanism of action of botulinum toxin (Ch. 23)
- Neurotransmitter transporters are targets of a variety of drugs of abuse (e.g., cocaine) as well as therapeutic drugs commonly used in psychiatry (e.g., Prozac, Zoloft, Paxil) (Ch. 23)
- Nicotinic acetylcholine receptors produced in brain neurons are important in learning and memory; loss of these receptors is observed in schizophrenia, epilepsy, drug addiction, and Alzheimer's disease (Ch. 23)

- People vary significantly in their ability to detect certain odors (Ch. 23)
- Synaptic translation of localized mRNAs is critical to the formation and the experience-dependent plasticity of neural circuits; alterations in this process result in neurodevelopmental and cognitive disorders (Ch. 23)
- People with defects in RAG gene function have no B cells and suffer from severe immunodeficiency (Ch. 24)
- The immunosuppressant drug cyclosporine inhibits calcineurin activity through the formation of a cyclosporine-cyclophilin complex, thus enabling successful allogenic tissue transplantation (Ch. 24)
- Vaccines elicit protective immunity against a variety of pathogens (Ch. 24)
- The immune system may also help in warding off cancer (Ch. 24)
- Increased understanding of the molecular cell biology of tumors is revolutionizing the way cancers are diagnosed and treated (Ch. 25)
- Using the loss of DNA repair in tumors for targeted chemotherapy (Ch. 25)
- Breast cancer as an example of how molecular cell biology techniques have affected both curative and palliative cancer treatment (Ch. 25)

Plant Biology Connections

Developments in agriculture, environmental science, and alternative energy production have demonstrated that the molecular cell biology of plants is increasingly relevant to our lives. Understanding photosynthesis and chloroplasts is just the beginning of plant biology. Throughout the text, we have highlighted plant-specific topics, including aspects of cell structure and function that are unique to plants, plant development, and plant biotechnology applications directed toward solving problems in agriculture and medicine.

- Vascular plants have rigid cell walls and use turgor pressure to stand upright and grow (Ch. 11)
- Transgenic plants have been produced that overexpress the vacuolar Na^+/H^+ antiporter, and can therefore grow successfully in soils containing high salt concentrations (Ch. 11)
- Chloroplast transformation has led to engineered plants that are resistant to infections as well as plants that can be used to make protein drugs (Ch. 12)
- In giant green algae such as *Nitella*, the cytosol flows rapidly due to use of myosin V (Ch. 17)
- Formation of the spindle and cytokinesis have unique features in plants (Ch. 18)
- The Venus flytrap is a well-known example of macroscopic plant mechanosensation (Ch. 20)
- Expression of a group of cell-wall degrading enzymes and their remodeling of cell walls by degrading pectin plays a key role in determining the rate and extent of the softening of fruit (Ch. 20)
- Meristems are niches for stem cells in plants (Ch. 22)

- A negative feedback loop maintains the size of the shoot apical stem-cell population (Ch. 22)
- The root meristem resembles the shoot meristem in structure and function (Ch. 22)

Acknowledgments

This edition would not have been possible without the careful and committed collaboration of our publishing partners at Macmillan Learning. We thank Daryl Fox, Sandy Lindelof, Lisa Lockwood, Will Moore, Debbie Hardin, Katrina Mangold, Karen S. Misler, Nathan Livingston, Tracey Kuehn, Jennifer MacMillan, Krystyna Borgen, Cassandra Korsvik, Lisa Kinne, Peter Jacoby, Diana Blume, Paul Rohloff, Natasha Wolfe, John Callahan, Matt McAdams, Keri deManigold, and Daniel Comstock for their labor and for their willingness to work overtime to produce a book that excels in every way.

In particular, we would like to acknowledge the talent and commitment of our text editors, Michael Zierler, Maria Lokshin, and Anna Bristow. They are remarkable editors. Thank you for all you've done in this edition.

We are also indebted to H. Adam Steinberg for his pedagogical insight and his development of beautiful molecular models and illustrations.

We would like to acknowledge those whose direct contributions to previous editions continue to influence in this edition, especially Ruth Steyn.

Thanks to our own staff: Sally Bittancourt, Diane Bush, Mary Anne Donovan, Carol Eng, James Evans, Julie Knight, George Kokkinogenis, Guicky Waller, Nicki Watson, and Rob Welsh.

Special thanks to our families for inspiring us and for granting us the time it takes to work on such a book and to our mentors and advisers for encouraging us in our studies and teaching us much of what we know: *(Harvey Lodish)* my wife, Pamela; my children and grandchildren Heidi and Eric Steinert and Emma and Andrew Steinert; Martin Lodish, Kristin Schardt, and Sophia, Joshua, and Tobias Lodish; and Stephanie Lodish, Bruce Peabody, and Isaac and Violet Peabody; mentors Norton Zinder and Sydney Brenner; and also David Baltimore and Jim Darnell for collaborating on the first editions of this book; *(Arnold Berk)* my wife, Sally, Jerry Berk, Shirley Berk, Angelina Smith, David Clayton, and Phil Sharp; *(Chris A. Kaiser)* my wife, Kathy O'Neill, my mentors David Botstein and Randy Schekman; *(Monty Krieger)* my wife, Nancy Krieger, parents I. Jay Krieger and Mildred Krieger, children Joshua and Ilana Krieger and Jonathan Krieger and Sofia Colucci, and grandchildren Joaquin, Alejandro, and Ari Krieger; my mentors Robert Stroud, Michael Brown, and Joseph Goldstein; *(Anthony Bretscher)* my wife, Janice, and daughters Heidi and Erika, and advisers A. Dale Kaiser and Klaus Weber; *(Kelsey C. Martin)* my husband, Joel Braslow, children Seth, Ben, Sam, and Maya, father George M. Martin, and mentors Ari Helenius and Eric Kandel; *(Michael B. Yaffe)* my wife Sung-Yun Pai and sons Samkyu and Minkyu, and my mentors Himan

Sternlicht and Lewis Cantley; *(Angelika Amon)* my husband, Johannes Weis, Theresa and Clara Weis, Gerry Fink and Frank Solomon.

Finally, in updating, revising, and rewriting this book, we were given invaluable help by many colleagues. We thank the following people who generously gave of their time and expertise by making contributions to specific chapters in their areas of interest, providing us with detailed information about their courses, or by reading and commenting on one or more chapters.

Jennifer Achiro, *University of California, Los Angeles*
Christa Adam, *Prairie State College*
R. Claudio Aguilar, *Purdue University–Main Campus*
Mohammed AlQuraishi, *Harvard University*
Kyle Anderson, *University of Saskatchewan*
Dragana Antic, *University of California, Berkeley Extension*
David Baker, *University of Washington*
Benjamin Barad, *The Scripps Research Institute*
David P. Bartel, *Whitehead Institute for Biomedical Research*
Uwe Beffert, *Boston University*
Jeffrey Benovic, *Philadelphia University and Thomas Jefferson University*
Michael W. Black, *California Polytech State University, San Luis Obispo*
Jonathan Bogan, *Yale University, Yale School of Medicine*
Adam Bohnert, *Louisiana State University & A&M College*
Jim Bonacum, *University of Illinois Springfield*

Roberto Botelho, *Ryerson University*
Federica Brandizzi, *Michigan State University, East Lansing*
Chloe Bulinski, *Columbia University in the City of New York*
Sharon K. Bullock, *University of North Carolina at Charlotte*
Dean Buonomano, *University of California, Los Angeles*
Samantha Butler, *University of California, Los Angeles*
Kevin P. Campbell, *University of Iowa*
Steve Carr, *Broad Institute*
David C. Chan, *California Institute of Technology*
Chaoping Chen, *Fort Lewis College*
James Darnell, *Rockefeller University*
Zhicheng Dou, *Clemson University*
Arri Eisen, *Emory University*
Ray Enke, *University of Delaware*
Irene M. Evans, *Rochester Institute of Technology*
Daniel Finley, *Harvard Medical School*
Christine M. Fleet, *Emory & Henry College*
James C. Fleet, *Purdue University–Main Campus*
Kenneth Frauwirth, *University of Maryland College Park*
Margaret Fuller, *Stanford University*
Mary Gehring, *Whitehead Institute and Massachusetts Institute of Technology*
Michael A. Glotzer, *University of Chicago*
Daniel A. Gold, *Saint Edward's University*
Alfred Goldberg, *Harvard Medical School*
Ben Goult, *University of Kent*

Berthold Gottgens, *Wellcome Trust; MRC Stem Cell Institute; Cambridge University*

Eric Guisbert, *Florida Institute of Technology–Melbourne*

Ronald G. Haller, *University of Texas Southwestern Medical Center*

Wendy Hanna-Rose, *Pennsylvania State University–Main Campus*

Craig Hart, *Louisiana State University & A&M College*

Ian C. Haydon, *University of Washington*

Andreas Herrlich, *Washington University School of Medicine*

Barry H. Honig, *Columbia University*

Richard O. Hynes, *Massachusetts Institute of Technology*

Mack Ivey, *University of Arkansas–Fayetteville*

Rudolf Jaenisch, *Whitehead Institute and Massachusetts Institute of Technology*

Michael G. Jonz, *University of Ottawa*

Torah Kachur, *University of Alberta, Edmonton*

Fedor Karginov, *University of California, Riverside*

Naohiro Kato, *Louisiana State University & A&M College*

Randy Kaufman, *Sanford Burnham Prebys Medical Discovery Institute*

Amy Keating, *Massachusetts Institute of Technology*

Eirini Kefalogianni, *Washington University School of Medicine*

Thomas Keller, *Florida State University–Tallahassee*

Greg Kelly, *University of Western Ontario*

Baljit Khakh, *University of California, Los Angeles*

Lou W. Kim, *Florida International University*

Maria Krasilnikova, *Pennsylvania State University–Main Campus*

Arnold Kriegstein, *University of California, San Francisco*

Iwo Kucinsky, *Wellcome Trust–MRC Stem Cell Institute–Cambridge University*

Salil Lachke, *University of Delaware*

Pranav Lalgudi, *Massachusetts Institute of Technology*

Vivienne K. Y. Lam, *University of British Columbia*

Eric Lander, *Broad Institute and Massachusetts Institute of Technology*

Daniel Leahy, *University of Texas at Austin*

Pulin Li, *Whitehead Institute and Massachusetts Institute of Technology*

Song-Tao Liu, *University of Toledo*

Andrew Lombardo, *Cornell University*

Richard L. Londraville, *University of Akron*

Sebastian Lourido, *Whitehead Institute and Massachusetts Institute of Technology*

Quentin Machingo, *Manhattan College*

Anita Mandal, *Edwards Waters College*

Emilie Marcus, *University of California, Los Angeles*

Heidi McBride, *McGill University*

Mill Miller, *Wright State University–Main Campus*

Vamsi Mootha, *Massachusetts General Hospital and Harvard Medical School*

Aristidis Moustakas, *Uppsala University*

Alexander Murashov, *East Carolina University*

Brion W. Murray, *Turning Point Therapeutics*

Hao Nguyen, *California State University, Sacramento*

Benjamin Novitch, *University of California, Los Angeles*

Oliver R. Oakley, *Eastern Kentucky University*

Yuki Oka, *California Institute of Technology*

Jamie Parker, *Fordham University–Bronx*
Joel Parker, *State University of New York, Plattsburgh*
Samantha Parks, *Georgia State University*
Ardem Patapoutian, *The Scripps Research Institute*
Mark Peifer, *University of North Carolina*
Richard Posner, *Northern Arizona University*
Marianne Poxleitner, *Gonzaga University*
Peter Reddien, *Whitehead Institute and Massachusetts Institute of Technology*
David Reich, *Harvard Medical School*
Kacper B. Rogala, *Whitehead Institute*
Rajat Rohatgi, *Stanford University*
Jatin Roper, *Duke University Medical School*
Joel Rovnak, *Colorado State University–Fort Collins*
Richard Roy, *McGill University*
Alapakkam Sampath, *UCLA School of Medicine*
Philip J. Santangelo, *Georgia Institute of Technology*
Gabriela S. Schlau-Cohen, *Massachusetts Institute of Technology*
Felix Schweizer, *University of California, Los Angeles*
Hao Shen, *University of Washington*
Peter Sicinski, *Harvard Medical School*
Jamie Siders, *Ohio Northern University*
Jeffrey D. Singer, *Portland State University*
Jonathan Snow, *Barnard College*
Jamie C. Snyder, *California Polytech State University–Pomona*
Michelle Snyder, *Towson University*
Merav Socolovsky, *University of Massachusetts Medical School*

Bao-Liang Song, *Wuhan University*

Kay Song, *Georgia State University*

Susan Spencer, *Saint Louis University*

Timothy Springer, *Boston Children's Hospital and Harvard Medical School*

Christine Suetterlin, *University of California, Irvine*

Sergei I. Sukharev, *University of Maryland*

Aurelian Tartar, *Nova Southeastern University*

Alison Taylor, *Harvard University*

Robert G. Van Buskirk, *State University of New York, Binghamton*

Matthew G. Vander Heiden, *Massachusetts Institute of Technology*

John P. Vessey, *University of Guelph*

Crystal M. Weyman, *Cleveland State University*

Forest White, *Massachusetts Institute of Technology*

Amy M. Wiles, *Mercer University–Macon*

Roderick Wilson, *University of Connecticut–Stamford*

Hao Wu, *Boston Children's Hospital and Harvard Medical School*

Hao Xu, *University of Southern Mississippi*

Yukiko Yamashita, *University of Michigan*

Omer Yilmaz, *Massachusetts Institute of Technology*

Junying Yuan, *Harvard Medical School*

Riasat Zaman, *Cornell University*

Brief Contents

Part I Chemical and Molecular Foundations

1 Evolution: Molecules, Genes, Cells, and Organisms

2 Chemical Foundations

3 Protein Structure and Function

4 Culturing and Visualizing Cells

Part II Biomembranes, Genes, and Gene Regulation

5 Fundamental Molecular Genetic Mechanisms

6 Molecular Genetic Techniques

7 Genes, Chromatin, and Chromosomes

8 Transcriptional Control of Gene Expression

9 Post-Transcriptional Gene Control

10 Biomembrane Structure

Part III Cellular Organization and Function

11 Transmembrane Transport of Ions and Small Molecules

12 Cellular Energetics

13 Moving Proteins into Membranes and Organelles

14 Vesicular Traffic, Secretion, and Endocytosis

15 Receptors, Hormones, and Cell Signaling

16 Growth Factor and Cytokine Signaling Pathways That Control Gene Expression

17 Cell Organization and Movement I: Microfilaments

18 Cell Organization and Movement II: Microtubules and Intermediate Filaments

19 The Eukaryotic Cell Cycle

Part IV Cell Growth and Differentiation

20 Integrating Cells into Tissues

21 Responding to the Cellular Environment

22 Stem Cells, Cell Asymmetry, and Regulated Cell Death

23 Cells of the Nervous System

24 Immunology

25 Cancer

Contents

Preface

Part I Chemical and Molecular Foundations

1 Evolution: Molecules, Genes, Cells, and Organisms

1.1 The Molecules of Life

- Proteins Give Cells Structure and Perform Most Cellular Tasks
- Nucleic Acids Carry Coded Information for Making Proteins at the Right Time and Place
- Phospholipids Are the Conserved Building Blocks of All Cellular Membranes
- Quality Control of All Cellular Macromolecules Is Essential for Life

1.2 Prokaryotic Cell Structure and Function

- Prokaryotes Comprise Two Kingdoms: Archaea and Eubacteria
- Many Bacteria Including *Escherichia coli* Are Widely Used in Biological Research

1.3 Eukaryotic Cell Structure and Function

- The Cytoskeleton Has Many Important Functions
- The Nucleus Contains the DNA Genome, Apparatuses for Synthesis of DNA and RNA, and a Fibrous Matrix

The Endoplasmic Reticulum Is the Site of Synthesis of Most Membrane and Secreted Proteins as Well as Many Lipids

The Golgi Complex Sorts Secreted Proteins and Many Membrane Proteins to Their Final Destinations in the Cell

Endosomes Bring Proteins and Particles from the Outside into Cells

Lysosomes Are Cellular Recycling Centers

Plant Vacuoles Store Water, Ions, and Small-Molecule Nutrients Such as Sugars and Amino Acids

Peroxisomes and Plant Glyoxisomes Metabolize Fatty Acids and Other Small Molecules Without Producing ATP from ADP and P_i

Mitochondria Are the Principal Sites of ATP Production in Aerobic Cells

Chloroplasts Contain Internal Compartments in Which Photosynthesis Takes Place

Many Organelle-Like Structures Are Unbounded by a Membrane

All Eukaryotic Cells Use a Similar Cycle to Regulate Their Division

1.4 Unicellular Eukaryotic Organisms Widely Used in Cell Biology Research

Yeasts Are Used to Study Fundamental Aspects of Eukaryotic Cell Structure and Function

Mutations in Yeast Led to the Identification of Key Cell Cycle Proteins

Studies in the Alga *Chlamydomonas reinhardtii* Led to the Development of a Powerful Technique to Study Brain Function

The Parasite That Causes Malaria Has Novel Organelles That Allow It to Undergo a Remarkable Life Cycle

1.5 Metazoan Structure, Function, Evolution, and Differentiation

Multicellularity Requires Cell-Cell and Cell-Matrix Adhesions

Epithelia Originated Early in Evolution

Cells Are Organized into Tissues and Tissues into Organs

Genomics Has Revealed Important Aspects of Metazoan Evolution and Cell Function

Development Uses a Conserved Set of Master Transcription Factors and Involves Epigenetic Modifications to DNA and Its Associated Histone Proteins

1.6 Metazoan Organisms Widely Used in Cell Biology Research

Drosophila melanogaster and *Caenorhabditis elegans* Are Used to Identify Genes That Regulate Animal Development

Planaria Are Used to Study Stem Cells and Tissue Regeneration

Studies on Fish, Mice, and Other Vertebrate Organisms Inform the Study of Human Development and Disease

Human Genetic Diseases Elucidate Important Aspects of Cell Function

Unbiased Single Cell Sequencing Experiments Identify Altogether New Cell Types

The Following Chapters Present Many Experimental Techniques and Much Experimental Data That Explains How We Know What We Know About Cell Structure and Function

2 Chemical Foundations

2.1 Covalent Bonds and Noncovalent Interactions

The Electronic Structure of an Atom Determines the Number and Geometry of the Covalent Bonds It Can Make

All Covalent Bonds Are Not Equal: Electrons May Be Shared Equally or Unequally in Covalent Bonds

Covalent Bonds Are Much Stronger and More Stable Than Noncovalent Interactions

Ionic Bonds Are Noncovalent Interactions Formed by the Electrostatic Attractions Between Oppositely Charged Ions

Hydrogen Bonds Are Noncovalent Interactions That Determine the Properties of Water and the Water Solubility of Uncharged Molecules

Van der Waals Interactions Are Weak Attractive Interactions Caused by Transient Dipoles

The Hydrophobic Effect Causes Nonpolar Molecules to Adhere to One Another

Molecular Complementarity Due to Noncovalent Interactions Leads to a Lock-and-Key Fit Between Biomolecules

2.2 Chemical Building Blocks of Cells

Amino Acids Differing Only in Their Side Chains Compose Proteins

Five Different Nucleotides Are Used to Build Nucleic Acids

Monosaccharides Covalently Assemble into Linear and Branched Polysaccharides

Phospholipids Associate Noncovalently to Form the Basic Bilayer Structure of Biomembranes

2.3 Chemical Reactions and Chemical Equilibrium

A Chemical Reaction Is in Equilibrium When the Rates of the Forward and Reverse Reactions Are Equal

The Equilibrium Constant Reflects the Extent of a Chemical Reaction

Chemical Reactions in Cells Are at Steady State

Dissociation Constants of Binding Reactions Reflect the Affinity of Interacting Molecules

Biological Fluids Have Characteristic pH Values

Hydrogen Ions Are Released by Acids and Taken Up by Bases

Buffers Maintain the pH of Intracellular and Extracellular Fluids

2.4 Biochemical Energetics

Several Forms of Energy Are Important in Biological Systems

Cells Can Transform One Type of Energy into Another

The Change in Free Energy Determines If a Chemical Reaction Will Occur Spontaneously

The $\Delta G^{\circ\prime}$ of a Reaction Can Be Calculated from Its K_{eq}

The Rate of a Reaction Depends on the Activation Energy Necessary to Energize the Reactants into a Transition State

Life Depends on the Coupling of Energetically Unfavorable Chemical Reactions with Energetically Favorable Ones

Hydrolysis of ATP Releases Substantial Free Energy and Drives Many Cellular Processes

ATP Is Generated During Photosynthesis and Respiration

NAD^+ and FAD Couple Many Biological Oxidation and Reduction Reactions

3 Protein Structure and Function

3.1 Hierarchical Structure of Proteins

The Primary Structure of a Protein Is Its Linear Arrangement of Amino Acids

Secondary Structures Are the Core Elements of Protein Architecture

Structural Motifs Are Regular Combinations of Secondary Structures

Tertiary Structure Is the Overall Folding of a Polypeptide Chain

Different Ways of Depicting the Conformation of Proteins Convey Different Types of Information

Domains Are Modules of Tertiary Structure

Comparing Protein Sequences and Structures Provides Insight into Protein Function and Evolution

There Are Four Broad Structural Categories of Proteins

Multiple Polypeptides Assemble into Quaternary Structures, Supramolecular Complexes, and Biomolecular Condensates

3.2 Protein Folding

Planar Peptide Bonds Limit the Shapes into Which Proteins Can Fold

Protein Folding Is Promoted by Proline Isomerases

The Amino Acid Sequence of a Protein Determines How It Will Fold

Folding of Proteins In Vivo Is Promoted by Chaperones

Abnormally Folded Proteins Can Form Amyloids That Are Implicated in Diseases

3.3 Protein Binding and Enzyme Catalysis

Specific Binding of Ligands Underlies the Functions of Most Proteins

Enzymes Are Highly Efficient and Specific Catalysts

An Enzyme's Active Site Binds Substrates and Carries Out Catalysis

Serine Proteases Demonstrate How an Enzyme's Active Site Works

Enzymes in a Common Pathway Are Often Physically Associated with One Another

3.4 Regulating Protein Function

Regulated Synthesis and Degradation of Proteins Is a Fundamental Property of Cells

The Proteasome Is a Molecular Machine Used to Degrade Proteins

Ubiquitin Marks Cytosolic Proteins for Degradation in Proteasomes

Noncovalent Binding Permits Allosteric, or Cooperative, Regulation of Proteins

Noncovalent Binding of Calcium and GTP Are Widely Used as Allosteric Switches to Control Protein Activity

Covalent Modification of Proteins Can Regulate their Activities

Phosphorylation and Dephosphorylation Covalently Regulate Protein Activity

The Structure and Function of Protein Kinase A Is Typical of Many Kinases

Protein Kinase Activity Is Often Regulated by Phosphorylation of the Kinase

Ubiquitinylation and Deubiquitinylation Covalently Regulate Protein Activity

Proteolytic Cleavage Irreversibly Activates or Inactivates Some Proteins

Higher Order Regulation Includes Control of Protein Location

3.5 Purifying, Detecting, and Characterizing Proteins

Centrifugation Can Separate Particles and Molecules That Differ in Mass or Density

Electrophoresis Separates Molecules on the Basis of Their Charge-to-Mass Ratio

Liquid Chromatography Resolves Proteins by Mass, Charge, or Affinity

Highly Specific Enzyme and Antibody Assays Can Detect Individual Proteins

Radioisotopes Are Indispensable Tools for Detecting Biological Molecules

Mass Spectrometry Can Determine the Mass and Sequence of Proteins

Protein Primary Structure Can Be Determined by Chemical Methods and from Gene Sequences

Protein Conformation Is Determined by Sophisticated Physical Methods

3.6 Proteomics

Proteomics Is the Study of All or a Large Subset of Proteins in a Biological System

Advanced Techniques in Mass Spectrometry Are Critical to Proteomic Analysis

4 Culturing and Visualizing Cells

4.1 Growing and Studying Cells in Culture

Culture of Animal Cells Requires Nutrient-Rich Media and Special Solid Surfaces

Primary Cell Cultures and Cell Strains Have a Finite Life Span

Transformed Cells Can Grow Indefinitely in Culture

Flow Cytometry Separates Different Cell Types

Growth of Cells in Two-Dimensional and Three-Dimensional Culture Mimics the In Vivo Environment

Stem Cells Can Differentiate in Culture to Make Organoids

Hybridomas Produce Abundant Monoclonal Antibodies

A Wide Variety of Cell Biological Processes Can Be Studied with Cultured Cells

Drugs Are Commonly Used in Cell Biological Research

4.2 Light Microscopy: Exploring Cell Structure and Visualizing Proteins Within Cells

The Resolution of the Conventional Light Microscope Is About 0.2 μm

Phase-Contrast and Differential-Interference-Contrast Microscopy Visualize Unstained Live Cells

Imaging Subcellular Details Often Requires That Specimens Be Fixed, Sectioned, and Stained

Fluorescence Microscopy Can Localize and Quantify Specific Molecules in Live Cells

Intracellular Ion Concentrations Can Be Determined with Ion-Sensitive Fluorescent Dyes

Immunofluorescence Microscopy Can Detect Specific Proteins in Fixed Cells

Tagging with Fluorescent Proteins Allows the Visualization of Specific Proteins in Live Cells

Deconvolution and Confocal Microscopy Enhance Visualization of Three-Dimensional Fluorescent Objects

Two-Photon Excitation Microscopy Allows Imaging Deep into Tissue Samples

TIRF Microscopy Provides Exceptional Imaging in One Focal Plane

FRAP Reveals the Dynamics of Cellular Components

FRET Measures Distance Between Fluorochromes

Optogenetics Allows Light to Regulate Events in a Spatial and Temporal Manner

Point Source Fluorescent Objects Can Be Located at Nanometer Resolution

Super-Resolution Microscopy Can Localize Proteins to Nanometer Accuracy

Light-Sheet Microscopy Can Rapidly Image Cells in Living Tissue

4.3 Electron Microscopy: High-Resolution Imaging

Single Molecules or Structures Can Be Imaged Using a Negative Stain or Metal Shadowing

Cells and Tissues Are Cut into Thin Sections for Viewing by Electron Microscopy

Immunoelectron Microscopy Localizes Proteins at the Ultrastructural Level

Cryoelectron Microscopy Allows Visualization of Specimens Without Fixation or Staining

Scanning Electron Microscopy of Metal-Coated Specimens Reveals Surface Features

4.4 Isolation of Cell Organelles

Disruption of Cells Releases Their Organelles and Other Contents

Centrifugation Can Separate Many Types of Organelles

Organelle-Specific Antibodies Are Useful in Preparing Highly Purified Organelles

Proteomics Reveals the Protein Composition of Organelles

Part II Biomembranes, Genes, and Gene Regulation

5 Fundamental Molecular Genetic Mechanisms

5.1 The Double-Helical Structure of DNA

Native DNA Is a Double Helix of Complementary Antiparallel Strands

DNA Can Undergo Reversible Strand Separation

DNA Molecules Can Acquire Torsional Stress

5.2 DNA Replication

DNA Polymerases Require a Template and a Primer to Replicate DNA

Duplex DNA Is Unwound, and Daughter Strands Are Formed at the DNA Replication Fork

A DNA Replication Fork Advances by Cooperation of Multiple Proteins

DNA Replication Occurs Bidirectionally from Each Origin

5.3 DNA Repair and Recombination

Chemical and Radiation Damage to DNA Can Lead to Mutations

High-Fidelity DNA Excision-Repair Systems Recognize and Repair Damage

Base Excision Repairs T-G Mismatches and Damaged Bases

Mismatch Excision Repairs Other Mismatches and Small Insertions and Deletions

Nucleotide Excision Repairs Chemical Adducts That Distort Normal DNA Shape

Two Systems Use Recombination to Repair Double-Strand Breaks in DNA

Homologous Recombination Can Repair DNA Damage and Generate Genetic Diversity

5.4 Transcription of Protein-Coding Genes and Formation of mRNA

A Template DNA Strand Is Transcribed into a Complementary RNA Strand by RNA Polymerase

Eukaryotic Precursor mRNAs Are Processed to Form Functional mRNAs

Alternative RNA Splicing Increases the Number of Proteins That Can Be Expressed from a Single Eukaryotic Gene

5.5 The Decoding of mRNA by tRNAs

Messenger RNA Carries Information from DNA in a Three-Letter Genetic Code

The Folded Structure of tRNA Promotes Its Decoding Functions

Nonstandard Base Pairing Often Occurs Between Codons and Anticodons

Amino Acids Are Linked to Their Cognate tRNAs with Great Accuracy

5.6 Stepwise Synthesis of Proteins on Ribosomes

Ribosomes Are Protein-Synthesizing Machines

Methionyl-$tRNA_i^{Met}$ Recognizes the AUG Start Codon

Eukaryotic Translation Initiation Usually Occurs at the First AUG Downstream from the 5′ End of an mRNA

During Chain Elongation Each Incoming Aminoacyl-tRNA Moves Through Three Ribosomal Sites

Translation Is Terminated by Release Factors When a Stop Codon Is Reached

Polysomes and Rapid Ribosome Recycling Increase the Efficiency of Translation

GTPase-Superfamily Proteins Function in Several Quality-Control Steps of Translation

Nonsense Mutations Can Be Bypassed by Suppressing tRNA Mutations

5.7 Viruses: Parasites of the Cellular Genetic System

Most Viral Host Ranges Are Narrow

Viral Capsids Are Regular Arrays of One or a Few Types of Protein

Lytic Viral Growth Cycles Lead to Death of Host Cells

Viral DNA Is Integrated into the Host-Cell Genome in Some Nonlytic Viral Growth Cycles

6 Molecular Genetic Techniques

6.1 Using Genetic Analysis of Mutations to Identify and Study Genes

Recessive and Dominant Mutant Alleles Generally Have Opposite Effects on Gene Function

Segregation of Mutations in Breeding Experiments Reveals Whether They Are Dominant or Recessive

Conditional Mutations Can Be Used to Study Essential Genes in Yeast

Recessive Lethal Mutations in Diploids Can Be Identified by Inbreeding and Maintained in Heterozygotes

Complementation Tests Determine Whether Different Recessive Mutations Are in the Same Gene

Double Mutants Are Useful in Assessing the Order in Which Proteins Function

Genetic Suppression and Synthetic Lethality Can Reveal Interacting or Redundant Proteins

Global Analysis of Double Mutant Combinations Can Reveal Networks of Gene Functions

6.2 DNA Cloning and Characterization

Restriction Enzymes and DNA Ligases Allow Insertion of DNA Fragments into Cloning Vectors

Isolated DNA Fragments Can Be Cloned into *E. coli* Plasmid Vectors

Yeast Genomic Libraries Can Be Constructed with Shuttle Vectors and Screened by Functional Complementation

cDNA Libraries Represent the Sequences of Protein-Coding Genes

The Polymerase Chain Reaction Amplifies a Specific DNA Sequence from a Complex Mixture

Cloned DNA Molecules Can Be Sequenced Rapidly by Methods Based on PCR

6.3 Using Sequence Information to Identify Genes and Deduce Their Function

Most Genes Can Be Readily Identified Within Genomic DNA Sequences

Bioinformatic Principles Can Be Used to Deduce the Likely Functional Consequences of Mutations

The Function and Evolutionary Origins of Genes and Proteins Can Be Deduced from Their Sequence

Comparison of Related Sequences from Different Species Can Give Clues to Evolutionary Relationships Among Proteins

Biological Complexity of an Organism Is Not Directly Related to the Number of Protein-Coding Genes in the Genome

6.4 Locating and Identifying Genes That Specify Human Traits

Monogenic Diseases Show One of Three Patterns of Inheritance

DNA Polymorphisms Are Used as Markers for Linkage Mapping of Human Mutations

Human Linkage Studies Can Map Disease Genes with a Resolution of About 1 Mbp

Further Analysis Is Needed to Locate a Disease Gene in Cloned DNA

Most Inherited Diseases Result from Multiple Genetic Defects

Identifying Component Genetic Risk Factors of Complex Traits

Medically Important Genes Can Be Identified by Alleles That Protect Against Disease

Identification of Causative Mutations in Cancer Cells

6.5 Using Cloned DNA Fragments to Study Gene Expression

In Situ Hybridization Techniques Permit Detection of Specific mRNAs

DNA Microarrays Can Be Used to Evaluate the Expression of Many Genes at Once

Cluster Analysis of Multiple Expression Experiments Identifies Co-Regulated Genes

Sequencing of cDNAs Allows Analysis of Gene Expression in Individual Cells

E. coli Expression Systems Can Produce Large Quantities of Proteins from Cloned Genes

Plasmid Expression Vectors Can Be Designed for Use in Animal Cells

6.6 Altering the Function of Specific Genes by Design

Normal Yeast Genes Can Be Replaced with Mutant Alleles by Homologous Recombination

Engineered CRISPR Systems Allow Precise Genome Editing

Somatic Cell Recombination Can Inactivate Genes in Specific Tissues of Mice

RNA Interference Causes Gene Inactivation by Destroying the Corresponding mRNA

7 Genes, Chromatin, and Chromosomes

7.1 Eukaryotic Gene Structure and Organization

Most Genes of Multicellular Eukaryotes Contain Introns and Produce mRNAs Encoding Single Proteins

Simple and Complex Transcription Units Are Found in Eukaryotic Genomes

Protein-Coding Genes May Be Solitary or Belong to a Gene Family

Heavily Used Gene Products Are Encoded by Multiple Copies of Genes

Nonprotein-Coding Genes Encode Functional RNAs

7.2 Chromosomal Organization of Genes and Noncoding DNA

Genomes of Many Organisms Contain a Large Fraction of Noncoding DNA

Most Simple-Sequence DNAs Are Concentrated in Specific Chromosomal Locations

DNA Fingerprinting Depends on Differences in Length of Simple-Sequence DNAs

Unclassified Intergenic DNA Occupies a Significant Portion of the Genome

7.3 Transposable (Mobile) DNA Elements

Movement of Mobile Elements Involves a DNA or an RNA Intermediate

Most Mobile Elements in Bacteria Are DNA Transposons Known as Insertion Sequences

Eukaryotic DNA Transposons Move Using a Cut-and-Paste Process

LTR Retrotransposons Behave Like Intracellular Retroviruses

Non-LTR Retrotransposons Transpose by a Distinct Mechanism

Other Retroposed RNAs Are Found in Genomic DNA

Mobile DNA Elements Have Significantly Influenced Evolution

7.4 Structural Organization of Eukaryotic Chromatin and Chromosomes

Chromatin Is Made of Nucleosomes

Chromatin Structure Is Conserved Among Eukaryotes

Chromatin Is a Disordered Chain of Nucleosomes Packed Together at Different Concentration Densities in the Nucleus

Modifications of Histone Tails Control Chromatin Condensation and Function

Additional Nonhistone Proteins Regulate Transcription and Replication

7.5 Morphology and Functional Elements of Eukaryotic Chromosomes

Chromosome Number, Size, and Shape at Metaphase Are Species-Specific

During Metaphase, Chromosomes Can Be Distinguished by Banding Patterns and Chromosome Painting

Chromosome Painting and DNA Sequencing Reveal the Evolution of Chromosomes

Interphase Polytene Chromosomes Arise by DNA Amplification

Three Functional Elements Are Required for Replication and Stable Inheritance of Chromosomes

Centromere Sequences Vary Greatly in Length and Complexity

Addition of Telomeric Sequences by Telomerase Prevents Shortening of Chromosomes

8 Transcriptional Control of Gene Expression

8.1 Overview of Eukaryotic Transcription

Regulatory Elements in Eukaryotic DNA Are Found Both Close to and Many Kilobases Away from Transcription Start Sites

Three Eukaryotic Nuclear RNA Polymerases Catalyze Formation of Different RNAs

The Clamp Domain Enables RNA Polymerase II to Transcribe Long Stretches of DNA

The Largest Subunit in RNA Polymerase II Has an Essential Carboxy-Terminal Repeat

8.2 RNA Polymerase II Promoters and General Transcription Factors

RNA Polymerase II Initiates Transcription at DNA Sequences Corresponding to the 5′ Cap of mRNAs

The TATA Box, Initiators, and CpG Islands Function as Promoters in Eukaryotic DNA

General Transcription Factors Position RNA Polymerase II at Transcription Start Sites and Assist in Initiation

Elongation Factors Regulate the Initial Stages of Transcription in the Promoter-Proximal Region

8.3 Regulatory Sequences for Protein-Coding Genes and the Proteins Through Which They Function

Promoter-Proximal Elements Help Regulate Eukaryotic Genes

Distant Enhancers Often Stimulate Transcription by RNA Polymerase II

Most Eukaryotic Genes Are Regulated by Multiple Transcription-Control Elements

DNase I Footprinting and EMSA Detect Protein-DNA Interactions

Activators Are Composed of Distinct Functional Domains

Repressors Are the Functional Converse of Activators

DNA-Binding Domains Can Be Classified into Numerous Structural Types

Structurally Diverse Activation and Repression Domains Regulate Transcription

Transcription Factor Interactions Increase Gene-Control Options

Multiprotein Complexes Form on Enhancers

8.4 Molecular Mechanisms of Transcription Repression and Activation

Formation of Heterochromatin Silences Gene Expression at Telomeres, near Centromeres, and in Other Regions

Repressors Can Direct Histone Deacetylation at Specific Genes

Activators Can Direct Histone Acetylation at Specific Genes

Chromatin-Remodeling Complexes Help Activate or Repress Transcription

Pioneer Transcription Factors Initiate the Process of Gene Activation During Cellular Differentiation

The Mediator Complex Forms a Molecular Bridge Between Activation Domains and Pol II

Transcriptional Condensates Greatly Increase the Rate of Transcription Initiation

Transcription Occurs in Bursts

8.5 Regulation of Transcription-Factor Activity

DNase I Hypersensitive Sites Reflect the Developmental History of Cellular Differentiation

Nuclear Receptors Are Regulated by Lipid-Soluble Hormones

All Nuclear Receptors Share a Common Domain Structure

Nuclear-Receptor Response Elements Contain Inverted or Direct Repeats

Hormone Binding to a Nuclear Receptor Regulates Its Activity as a Transcription Factor

Metazoans Regulate the RNA Polymerase II Transition from Initiation to Elongation

Termination of Transcription Is Also Regulated

8.6 Epigenetic Regulation of Transcription

DNA Methylation Regulates Transcription

Methylation of Specific Histone Lysines Is Linked to Epigenetic Mechanisms of Gene Repression

Epigenetic Control by Polycomb and Trithorax Complexes

Long Noncoding RNAs Direct Epigenetic Repression in Metazoans

8.7 Other Eukaryotic Transcription Systems

Transcription Initiation by Pol I and Pol III Is Analogous to That by Pol II

9 Post-Transcriptional Gene Control

9.1 Processing of Eukaryotic Pre-mRNA

The 5′ Cap Is Added to Nascent RNAs Shortly After Transcription Initiation

Chain Elongation by RNA Polymerase II Is Coupled to the Presence of RNA Processing Factors

A Diverse Set of Proteins with Conserved RNA-Binding Domains Associate with Pre-mRNAs

Splicing Occurs at Short, Conserved Sequences in Pre-mRNAs via Two Transesterification Reactions

During Splicing, snRNAs Base-Pair with Pre-mRNA to Select Splice Sites and Guide the Transesterification Reactions

Spliceosomes Catalyze Pre-mRNA Splicing

3′ Cleavage and Polyadenylation of Pre-mRNAs Are Tightly Coupled

9.2 Regulation of Pre-mRNA Processing

Additional Nuclear Proteins Contribute to Splice-Site Selection in the Long Pre-mRNAs of Humans and Other Vertebrates

Expression and Function of Related K^+-Channel Protein Isoforms in Vertebrate Inner Ear Hair Cells

Regulation of RNA Splicing Through Splicing Enhancers and Silencers Controls *Drosophila* Sexual Differentiation

Splicing Repressors and Activators Control Splicing at Alternative Sites

Expression of *Dscam* Isoforms in *Drosophila* Retinal Neurons

Abnormal RNA Splicing and Disease

Self-Splicing Group II Introns Provide Clues to the Evolution of snRNAs

Nuclear Exonucleases and the Exosome Degrade RNA That Is Processed out of Pre-mRNAs

RNA Processing Solves the Problem of Pervasive Transcription of the Genome in Mammalian Cells

RNA Editing Alters the Sequences of Some Pre-mRNAs

9.3 Transport of mRNA Across the Nuclear Envelope

SR Proteins Mediate Nuclear Export of mRNA

Pre-mRNAs Associated with Spliceosomes Are Not Exported from the Nucleus

HIV Rev Protein Regulates the Transport of Unspliced Viral mRNAs

9.4 Cytoplasmic Mechanisms of Post-Transcriptional Control

The Concentration of an mRNA in the Cytoplasm Is Determined by Its Rate of Synthesis and Its Rate of Degradation

Degradation of mRNAs in the Cytoplasm Occurs by Several Mechanisms

MicroRNAs Repress Translation and Induce Degradation of Specific mRNAs

RNA Interference Induces Degradation of Precisely Complementary mRNAs

Cytoplasmic Polyadenylation Promotes Translation of Some mRNAs

Protein Synthesis Can Be Globally Regulated

Sequence-Specific RNA-Binding Proteins Control Specific mRNA Translation

Surveillance Mechanisms Prevent Translation of Improperly Processed mRNAs

Localization of mRNAs Permits Production of Proteins at Specific Regions Within the Cytoplasm

9.5 Processing of rRNA and tRNA

Pre-rRNA Genes Are Similar in All Eukaryotes and Function as Nucleolar Organizers

Small Nucleolar RNAs Assist in Processing Pre-rRNAs

Self-Splicing Group I Introns Were the First Examples of Catalytic RNA

Pre-tRNAs Undergo Extensive Modification in the Nucleus

9.6 Nuclear Bodies Are Functionally Specialized Nuclear Domains

Cajal Bodies

Nuclear Speckles

Nuclear Paraspeckles

Promyelocytic Leukemia (PML) Nuclear Bodies

Nucleolar Functions in Addition to Ribosomal Subunit Synthesis

10 Biomembrane Structure

10.1 The Lipid Bilayer: Composition and Structural Organization

Phospholipids Spontaneously Form Bilayers

Phospholipid Bilayers Form a Sealed Compartment Surrounding an Internal Aqueous Space

Biomembranes Contain Three Principal Classes of Lipids

Most Lipids and Many Proteins Are Laterally Mobile in Biomembranes

Lipid Composition Influences the Physical Properties of Membranes

Lipid Composition Is Different in the Exoplasmic and Cytosolic Leaflets

Cholesterol and Sphingolipids Cluster with Specific Proteins in Membrane Microdomains

Cells Store Excess Lipids in Lipid Droplets

10.2 Membrane Proteins: Structure and Basic Functions

Proteins Interact with Membranes in Three Different Ways

Most Transmembrane Proteins Have Membrane-Spanning α Helices

Multiple β Strands in Porins Form Membrane-Spanning "Barrels"

Covalently Attached Lipids Anchor Some Proteins to Membranes

All Transmembrane Proteins and Glycolipids Are Asymmetrically Oriented in the Bilayer

Lipid-Binding Motifs Help Target Peripheral Proteins to the Membrane

Proteins Can Be Removed from Membranes by Detergents or High-Salt Solutions

10.3 Phospholipids, Sphingolipids, and Cholesterol: Synthesis and Intracellular Movement

Fatty Acids Are Assembled from Two-Carbon Building Blocks by Several Important Enzymes

Small Cytosolic Proteins Facilitate Movement of Fatty Acids

Fatty Acids Are Incorporated into Phospholipids Primarily on the ER Membrane

Flippases Move Phospholipids from One Membrane Leaflet to the Opposite Leaflet

Cholesterol Is Synthesized by Enzymes in the Cytosol and ER Membrane

Cholesterol and Phospholipids Are Transported Between Organelles by Several Mechanisms

Part III Cellular Organization and Function

11 Transmembrane Transport of Ions and Small Molecules

11.1 Overview of Transmembrane Transport

Only Gases and Small Uncharged Molecules Cross Membranes by Simple Diffusion

Three Main Classes of Membrane Proteins Transport Molecules and Ions Across Cellular Membranes

11.2 Facilitated Transport of Glucose and Water

Uniport Transport Is Faster and More Specific than Simple Diffusion

The Low K_m of the GLUT1 Uniporter Enables It to Transport Glucose into Most Mammalian Cells

The Human Genome Encodes a Family of Sugar-Transporting GLUT Proteins

Transport Proteins Can Be Studied Using Artificial Membranes and Recombinant Cells

Osmotic Pressure Causes Water to Move Across Membranes

Aquaporins Increase the Water Permeability of Cellular Membranes

11.3 ATP-Powered Pumps and the Intracellular Ionic Environment

There Are Four Main Classes of ATP-Powered Pumps

ATP-Powered Ion Pumps Generate and Maintain Ionic Gradients Across Cellular Membranes

Muscle Relaxation Depends on Ca^{2+} ATPases That Pump Ca^{2+} from the Cytosol into the Sarcoplasmic Reticulum

The Mechanism of Action of the Ca^{2+} Pump Is Known in Detail

The Na^+/K^+ ATPase Maintains the Intracellular Na^+ and K^+ Concentrations in Animal Cells

V-Class H^+ ATPases Maintain the Acidity of Lysosomes and Vacuoles

ABC Proteins Export a Wide Variety of Drugs and Toxins from the Cell

Certain ABC Proteins "Flip" Phospholipids and Other Lipid-Soluble Substrates from One Membrane Leaflet to the Other

The ABC Cystic Fibrosis Transmembrane Regulator Is a Chloride Channel, Not a Pump

11.4 Nongated Ion Channels and the Resting Membrane Potential

Selective Movement of Ions Creates a Transmembrane Electric Gradient

The Resting Membrane Potential in Animal Cells Depends Largely on the Outward Flow of K^+ Ions Through Open K^+ Channels

Ion Channels Are Selective for Certain Ions by Virtue of a Molecular Selectivity Filter

Patch Clamps Permit Measurement of Ion Movements Through Single Channels

Novel Ion Channels Can Be Characterized by a Combination of Oocyte Expression and Patch Clamping

11.5 Cotransport by Symporters and Antiporters

Na^+ Entry into Mammalian Cells Is Thermodynamically Favored

Na^+-Linked Symporters Enable Animal Cells to Import Sugars Including Glucose and Galactose as Well as Amino Acids Against High Concentration Gradients

A Bacterial Na^+/Amino Acid Symporter Reveals How Symport Works

A Na^+-Linked Ca^{2+} Antiporter Regulates the Strength of Cardiac Muscle Contraction

Several Cotransporters Regulate Cytosolic pH

An Anion Antiporter Is Essential for Transport of CO_2 by Erythrocytes

Numerous Transport Proteins Enable Plant Vacuoles to Accumulate Metabolites and Ions

11.6 Transcellular Transport

Multiple Transport Proteins Are Needed to Move Glucose and Amino Acids Across Epithelia

Simple Rehydration Therapy Depends on the Osmotic Gradient Created by Absorption of Glucose and Na^+

Parietal Cells Acidify the Stomach Contents While Maintaining a Neutral Cytosolic pH

Bone Resorption Requires the Coordinated Function of a V-Class Proton Pump and a Specific Chloride Channel

12 Cellular Energetics

12.1 Chemiosmosis, Electron Transport, the Proton-Motive Force, and ATP Synthesis

12.2 First Step of Harvesting Energy from Glucose: Glycolysis

- During Glycolysis (Stage I), Cytosolic Enzymes Convert Glucose to Pyruvate
- The Rate of Glycolysis Is Adjusted to Meet the Cell's Need for ATP
- Glucose Is Fermented When Oxygen Is Scarce

12.3 The Structure of Mitochondria

- Mitochondria Are Abundant, Multifunctional Organelles
- Mitochondria Have Two Structurally and Functionally Distinct Membranes
- Mitochondria Contain DNA and Evolved from a Single Endosymbiotic Event Involving Alphaproteobacterium
- The Size, Structure, and Coding Capacity of mtDNA Vary Considerably Among Organisms
- Mitochondrial DNA Is Located in the Matrix and Transferred During Mitosis to Daughter Cells by Cytoplasmic Inheritance
- Products of Mitochondrial Genes Are Not Exported
- Mitochondrial Genetic Codes Can Differ from the Standard Nuclear Code
- Mutations in Mitochondrial DNA Cause Several Genetic Diseases in Humans

12.4 The Dynamics of Mitochondria and Mitochondrial-ER Membrane Contact Sites

Mitochondria Are Dynamic Organelles

Mitochondrial Function and Dynamics Can Depend on Direct Contacts with Other Organelles

12.5 The Citric Acid Cycle and Fatty Acid Oxidation

In the First Part of Stage II, Pyruvate Is Converted to Acetyl CoA and High-Energy Electrons

In the Second Part of Stage II, the Citric Acid Cycle Oxidizes the Acetyl Group in Acetyl CoA to CO_2 and Generates High-Energy Electrons

Transporters in the Inner Mitochondrial Membrane Help Maintain Appropriate Cytosolic and Matrix Concentrations of NAD^+ and NADH

Mitochondrial Oxidation of Fatty Acids Generates ATP

Peroxisomal Oxidation of Fatty Acids Generates No ATP

12.6 The Electron-Transport Chain and Generation of the Proton-Motive Force

Oxidation of NADH and $FADH_2$ Releases a Substantial Amount of Energy

Electron Transport in Mitochondria Is Coupled to Proton Pumping

Electrons Release Energy as They Flow Downhill Through a Series of Electron Carriers

Four Large Multiprotein Complexes (I–IV) Couple Electron Transport to Proton Pumping Across the Inner Mitochondrial Membrane

The Reduction Potentials of Electron Carriers in the Electron-Transport Chain Favor Electron Flow from NADH to O_2

The Multiprotein Complexes of the Electron-Transport Chain Assemble into Supercomplexes

Reactive Oxygen Species Are By-Products of Electron Transport

Experiments Using Purified Electron-Transport Chain Complexes Established the Stoichiometry of Proton Pumping

The Proton-Motive Force in Mitochondria Is Due Largely to a Voltage Gradient Across the Inner Membrane

12.7 Harnessing the Proton-Motive Force to Synthesize ATP

The Mechanism of ATP Synthesis Is Shared Among Bacteria, Mitochondria, and Chloroplasts

ATP Synthase Comprises F_0 and F_1 Multiprotein Complexes

Rotation of the F_1 γ Subunit, Driven by Proton Movement Through F_0, Powers ATP Synthesis

Multiple Protons Must Pass Through ATP Synthase to Synthesize One ATP

F_0 c Ring Rotation Is Driven by Protons Flowing Through Transmembrane Channels

ATP-ADP Exchange and Phosphate Transport Across the Inner Mitochondrial Membrane Are Required to Supply ADP and Phosphate for ATP Synthesis

The Rate of Mitochondrial Oxidation Normally Depends on ADP Levels

Mitochondria in Brown Fat Use the Proton-Motive Force to Generate Heat

12.8 Chloroplasts and Photosynthesis

Thylakoid Membranes in Chloroplasts Are the Sites of Photosynthesis in Plants

Chloroplasts Contain Large DNAs Often Encoding More Than a Hundred Proteins

Light Absorption by Photosystems in Chloroplasts Provides the Energy That Drives the Synthesis of NADPH and ATP and the Generation of O_2 from H_2O

Three of the Four Stages in Photosynthesis Occur on the Thylakoid Membrane and Only During Illumination

Stages 1 and 2 of Photosynthesis Convert Sunlight into High Energy Electrons That Generate a Proton-Motive Force and NADPH

Core Antenna Complexes and Light-Harvesting Complexes Increase the Efficiency of Photosynthesis

Multiple Mechanisms Protect Cells Against Damage from Reactive Oxygen Species During Photoelectron Transport

12.9 Use of Light Energy to Generate Molecular Oxygen, NADPH, and ATP in Stages 1–3 of Photosynthesis

The First Three Stages of Photosynthesis

Relative Activities of Photosystems I and II Are Regulated

12.10 ATP and NADPH Drive Carbon Fixation in the Calvin Cycle and Carbohydrate Synthesis in Stage 4 of Photosynthesis

Rubisco Fixes CO_2 in the Chloroplast Stroma

Photorespiration Competes with Carbon Fixation and Is Reduced in C_4 Plants

13 Moving Proteins into Membranes and Organelles

13.1 Targeting Proteins to and Across the ER Membrane

Pulse-Chase Experiments with Purified ER Membranes Demonstrated That Secreted Proteins Cross the ER Membrane

A Hydrophobic N-Terminal Signal Sequence Targets Nascent Secretory Proteins to the ER

Cotranslational Translocation Is Initiated by Two GTP-Hydrolyzing Proteins

Passage of Growing Polypeptides Through the Translocon Is Driven by Translation

ATP Hydrolysis Powers Post-Translational Translocation of Some Secretory Proteins in Yeast

13.2 Insertion of Membrane Proteins into the ER

Several Topological Classes of Integral Membrane Proteins Are Synthesized on the ER

Internal Stop-Transfer Anchor and Signal-Anchor Sequences Determine Topology of Single-Pass Proteins

Type IV (Multipass) Proteins

A Phospholipid Anchor Tethers Some Cell-Surface Proteins to the Membrane

The Topology of a Membrane Protein Can Often Be Deduced from Its Sequence

13.3 Protein Modifications, Folding, and Quality Control in the ER

A Preformed *N*-Linked Oligosaccharide Is Added to Many Proteins in the Rough ER

Oligosaccharide Side Chains May Promote Folding and Stability of Glycoproteins

Disulfide Bonds Are Formed and Rearranged by Proteins in the ER Lumen

Chaperones and Other ER Proteins Facilitate Folding and Assembly of Proteins

Improperly Folded Proteins in the ER Induce Expression of Protein-Folding Catalysts

Unassembled or Misfolded Proteins in the ER Are Often Transported to the Cytosol for Degradation

13.4 Targeting of Proteins to Mitochondria and Chloroplasts

Amphipathic N-Terminal Targeting Sequences Direct Proteins to the Mitochondrial Matrix

Mitochondrial Protein Import Requires Outer-Membrane Receptors and Translocons in Both Membranes

Studies with Chimeric Proteins Demonstrate Important Features of Mitochondrial Protein Import

Three Energy Inputs Are Needed to Import Proteins into Mitochondria

Multiple Signals and Pathways Target Proteins to Submitochondrial Compartments

Import of Chloroplast Stromal Proteins Is Similar to Import of Mitochondrial Matrix Proteins

Proteins Are Targeted to Thylakoids by Mechanisms Related to Bacterial Protein Translocation

13.5 Targeting of Peroxisomal Proteins

A Cytosolic Receptor Targets Proteins with an SKL Sequence at the C-Terminus to the Peroxisomal Matrix

Peroxisomal Membrane and Matrix Proteins Are Incorporated by Different Pathways

13.6 Transport into and out of the Nucleus

Large and Small Molecules Enter and Leave the Nucleus via Nuclear Pore Complexes

Nuclear Transport Receptors Escort Proteins Containing Nuclear-Localization Signals into the Nucleus

A Second Type of Nuclear Transport Receptor Escorts Proteins Containing Nuclear-Export Signals out of the Nucleus

Most mRNAs Are Exported from the Nucleus by a Ran-Independent Mechanism

14 Vesicular Traffic, Secretion, and Endocytosis

14.1 Techniques for Studying the Secretory Pathway

Transport of a Protein Through the Secretory Pathway Can Be Assayed in Live Cells

Yeast Mutants Define Major Stages and Components of Vesicular Transport

Cell-Free Transport Assays Allow Dissection of Individual Steps in Vesicular Transport

14.2 Molecular Mechanisms of Vesicle Budding and Fusion

Assembly of a Protein Coat Drives Vesicle Formation and Selection of Cargo Molecules

A Conserved Set of GTPase Switch Proteins Controls the Assembly of Different Vesicle Coats

Targeting Sequences on Cargo Proteins Make Specific Molecular Contacts with Coat Proteins

Rab GTPases Control Docking of Vesicles on Target Membranes

Paired Sets of SNARE Proteins Mediate Fusion of Vesicles with Target Membranes

Dissociation of SNARE Complexes After Membrane Fusion Is Driven by ATP Hydrolysis

14.3 Early Stages of the Secretory Pathway

COPII Vesicles Mediate Transport from the ER to the Golgi

COPI Vesicles Mediate Retrograde Transport Within the Golgi and from the Golgi to the ER

Anterograde Transport Through the Golgi Occurs by Cisternal Maturation

14.4 Later Stages of the Secretory Pathway

Vesicles Coated with Clathrin and Adapter Proteins Mediate Transport from the *trans*-Golgi

Dynamin Is Required for Pinching Off of Clathrin-Coated Vesicles

Mannose 6-Phosphate Residues Target Resident Enzymes to Lysosomes

Study of Lysosomal Storage Diseases Revealed Key Components of the Lysosomal Sorting Pathway

Protein Aggregation in the *trans*-Golgi May Function in Sorting Proteins to Regulated Secretory Vesicles

Some Proteins Undergo Proteolytic Processing After Leaving the *trans*-Golgi

Distinct Pathways Sort Membrane Proteins to the Apical or Basolateral Region of Polarized Cells

14.5 Receptor-Mediated Endocytosis

Cells Take Up Lipids from the Blood in the Form of Large, Well-Defined Lipoprotein Complexes

Receptors for Macromolecular Ligands Contain Sorting Signals That Target Them for Endocytosis

The Acidic pH of Late Endosomes Causes Most Receptor-Ligand Complexes to Dissociate

Receptor-Mediated Endocytosis Can Down-Regulate Signaling Receptors

14.6 Directing Membrane Proteins and Cytosolic Materials to the Lysosome for Degradation

Multivesicular Endosomes Segregate Membrane Proteins Destined for the Lysosomal Membrane from Proteins Destined for Lysosomal Degradation

Retroviruses Bud from the Plasma Membrane by a Process Similar to Formation of Multivesicular Endosomes

The Autophagic Pathway Delivers Cytosolic Proteins or Entire Organelles to Lysosomes

15 Receptors, Hormones, and Cell Signaling

15.1 Signal Transduction Pathways: From Extracellular Signal to Cellular Response

Signaling Molecules Can Act Locally or at a Distance

Signal Transduction Pathways Can Produce Rapid, Short-Term or Slow, Long-Term Changes in Cells, or Both

Receptors Are Allosteric Proteins That Activate Signal Transduction Pathways

Receptors Can Be in the Cytosol, Nucleus, or on the Cell Surface Membrane

Most Receptors Bind Only a Single Type of Ligand or a Group of Closely Related Ligands

Most Receptors Bind Their Ligands with High Affinity

Second Messengers Are Used in Most Signal Transduction Pathways

Protein Kinases and Phosphatases Participate in Signal Transduction Pathways by Covalently Modifying and Thus Activating or Inhibiting a Wide Variety of Proteins That Control Cellular States

GTP-Binding Proteins Are Frequently Used in Signal Transduction Pathways as On/Off Switches

Signal Amplification and Feedback Repression Characterize Most Signal Transduction Pathways

15.2 Studying Cell-Surface Receptors and Signal Transduction Proteins

Binding Assays Are Used to Detect Receptors and Determine Their Affinity and Specificity for Ligands

Near-Maximal Cellular Response to a Signaling Molecule Usually Does Not Require Activation of All Receptors

Sensitivity of a Cell to External Signals Is Determined by the Number of Cell-Surface Receptors and Their Affinity for Ligand

Chemical Analogs of Signaling Molecules Are Used to Study Receptors and Are Widely Used as Drugs

Receptors Can Be Purified by Affinity Chromatography Techniques

Immunoprecipitation Assays and Affinity Techniques Can Be Used to Study the Activity of Protein Kinases

Immunoprecipitation of Kinases

Western Blotting with a Monoclonal Antibody Specific for a Phosphorylated Amino Acid in a Protein

GTP-Binding Signal Transduction Proteins Can Be Isolated and Their Activities Measured Using Pull-Down Assays

Concentrations of Free Ca^{2+} in the Mitochondrial Matrix, ER, and Cytosol Can Be Measured with Targeted Fluorescent Proteins

15.3 Structure and Mechanism of G Protein–Coupled Receptors

All G Protein–Coupled Receptors Share the Same Basic Structure

Ligand-Activated G Protein–Coupled Receptors Catalyze Exchange of GTP for GDP on the α Subunit of a Heterotrimeric G Protein

Different G Proteins Are Activated by Different GPCRs and in Turn Regulate Different Effector Proteins

Analysis of GPCRs Has Identified Important Human Hormones

15.4 Regulating Metabolism of Many Cells: G Protein–Coupled Receptors That Activate or Inhibit Adenylyl Cyclase

Adenylyl Cyclase Is Stimulated and Inhibited by Different Receptor-Ligand Complexes

cAMP Activates Protein Kinase A by Releasing Its Inhibitory Subunits

Glycogen Catabolism Is Stimulated by Hormone-Induced Activation of PKA

Signal Amplification Occurs in the cAMP-PKA Glycogen Degradation Pathway

cAMP-Mediated Activation of PKA Produces Diverse Responses in Different Cell Types

CREB Links cAMP and PKA to Activation of Gene Transcription

Anchoring Proteins Localize Effects of cAMP to Specific Regions of the Cell

Multiple Feedback Mechanisms Suppress Signaling by the GPCR/cAMP/PKA Pathway

15.5 Regulating Protein Secretion and Muscle Contraction: Ca^{2+} Ions as Second Messengers in Multiple Signal Transduction Pathways

Products of the Hydrolysis of the Membrane Lipid Phosphatidylinositol 4,5-Bisphosphate by Phospholipase C Elevate Cytosolic Ca^{2+} Levels

Ca^{2+} Release from the ER Triggered by IP_3

IP_3-triggered Ca^{2+} Transport from the ER to the Mitochondrial Matrix

The Store-Operated Ca^{2+} Channel in the Plasma Membrane

Feedback Loops in ER and Cytosol Cycling of Ca^{2+} Trigger Oscillations in the Cytosolic Ca^{2+} Concentration

DAG Activates Protein Kinase C

Integration of Ca^{2+} and cAMP Second Messengers Regulates Glycogenolysis

15.6 Vision: How the Eye Senses Light

Light Activates Rhodopsin in Rod Cells of the Eye

Activation of Rhodopsin by Light Leads to Closing of cGMP-Gated Cation Channels

Signal Amplification Makes the Rhodopsin Signal Transduction Pathway Exquisitely Sensitive

Rapid Termination of the Rhodopsin Signal Transduction Pathway Is Essential for the Temporal Resolution of Vision

Termination of the Signal from Light-Activated Rhodopsin (R*) by Rhodopsin Phosphorylation and Binding of Arrestin

Termination of the Signal from Activated $G_{\alpha t}$·GTP by GTP Hydrolysis

Rod Cells Adapt to Varying Levels of Ambient Light by Intracellular Trafficking of Arrestin and Transducin

16 Growth Factor and Cytokine Signaling Pathways That Control Gene Expression

16.1 Growth Factors and Their Receptor Tyrosine Kinases

Binding of Ligand to the Extracellular Domain of an RTK Leads to Dimerization and Activation of Its Intrinsic Cytosolic Tyrosine Kinase

Homo- and Hetero-Oligomers of Epidermal Growth Factor Receptors Bind Members of the Epidermal Growth Factor Family

EGF Receptor Homodimers

EGF Receptor Heterodimers with HER2

Ligand Binding to the EGF Receptor and Receptor Dimerization Results in the Formation of an Active Asymmetric Kinase Domain Dimer

Signal Transduction After Activation of RTKs: Phosphotyrosine Residues on the Receptor Are Binding Surfaces for Multiple Proteins with SH2 Domains

Receptor-Mediated Endocytosis and Lysosomal Degradation Squelch Signaling from RTKs

16.2 The Ras/MAP Kinase Signal Transduction Pathway

Ras, a GTPase Switch Protein, Operates Downstream of Most RTKs and Cytokine Receptors

Receptor Tyrosine Kinases Are Linked to Ras by Adapter Proteins

Binding of Sos to Inactive Ras Causes a Conformational Change That Triggers an Exchange of GTP for GDP

Signals Pass from Activated Ras to a Cascade of Protein Kinases Ending with MAP Kinase

MAP Kinase Regulates the Activity of Many Transcription Factors Controlling Early Response Genes

Multiple Feedback Mechanisms Restrain MAP Kinase Activation

Scaffold Proteins Isolate Distinct MAP Kinase Pathways in the Same Cell from One Another

16.3 Phosphoinositide Signal Transduction Pathways

Phospholipase C_γ Is Activated by Many RTKs and Cytokine Receptors

Recruitment of PI-3 Kinase to Activated Receptors Leads to Accumulation of Phosphatidylinositol-3-Phosphates in the Plasma Membrane and to Activation of Several Downstream Kinases

Activated Protein Kinase B Induces Many Cellular Responses

The PI-3 Kinase Pathway Is Negatively Regulated by PTEN Phosphatase

16.4 Cytokines, Cytokine Receptors, and the JAK/STAT Signaling Pathway

Cytokines Regulate the Development and Function of Many Cell Types

Binding of a Cytokine to Its Receptor Activates One or More Tightly Bound JAK Protein Tyrosine Kinases

JAK Kinases Phosphorylate and Activate STAT Transcription Factors

Multiple Mechanisms Suppress Signaling from Cytokine Receptors

Phosphotyrosine Phosphatases

SOCS Proteins

16.5 The TGF-β Family of Growth Factors, Their Receptor Serine Kinases, and the Smad Transcription Factors They Activate

TGF-β Proteins Are Stored in an Inactive Form in the Extracellular Matrix

Three Separate TGF-β Receptor Proteins Participate in Binding TGF-β and Activating Signal Transduction

Activated RI TGF-β Receptors Phosphorylate Smad Transcription Factors

The R-Smad/co-Smad Complex Activates Expression of Different Genes in Different Cell Types

Negative Feedback Loops Restrain TGF-β/Smad Signaling

16.6 Signal Transduction Pathways That Utilize Regulated, Site-Specific Protein Cleavage: Notch/Delta and EGF Precursors

On Binding Delta, the Notch Receptor Is Cleaved, Releasing a Component Transcription Factor

Metalloproteases Catalyze Cleavage of Many Signaling Proteins from the Cell Surface

16.7 Signal Transduction Pathways That Utilize Proteasomal Degradation of Signaling Components: Wnt, Hedgehog, and the Many Hormones That Activate NF-κB

Wnt Signaling Prevents Destruction of a Transcription Factor by a Cytosolic Protein Complex

Concentration Gradients of Wnt Protein Are Essential for Many Steps in Development

Hedgehog Signaling Relieves Repression of Target Gene Expression

Processing of Hh Precursor Protein

The Hh Receptors Patched and Smoothened and the Downstream Signaling Pathway Were Initially Elucidated by

Genetic Studies of *Drosophila* Development

Feedback Regulation of Hh Signaling

Hedgehog Signaling in Vertebrates Requires Primary Cilia

Degradation of an Inhibitor Protein Activates the NF-κB Transcription Factor

Enormous Signalsomes with Polyubiquitin Chain Scaffolds Link Many Cell Surface Receptors to Downstream Proteins in the NF-κB Pathway

17 Cell Organization and Movement I: Microfilaments

17.1 Microfilaments and Actin Structures

Actin Is Ancient, Abundant, and Highly Conserved

G-Actin Monomers Assemble into Long, Helical F-Actin Polymers

F-Actin Has Structural and Functional Polarity

17.2 Dynamics of Actin Filaments

Actin Polymerization In Vitro Proceeds in Three Steps

Actin Filaments Grow Faster at (+) Ends than at (−) Ends

Actin Filament Treadmilling Is Accelerated by Profilin and Cofilin

Thymosin-β_4 Provides a Reservoir of Actin for Polymerization

Capping Proteins Block Assembly and Disassembly at Actin Filament Ends

17.3 Mechanisms of Actin Filament Assembly

Formins Assemble Unbranched Filaments

The Arp2/3 Complex Nucleates Branched Filament Assembly

Intracellular Movements Can Be Powered by Actin Polymerization

Microfilaments Function in Endocytosis

Toxins That Perturb the Pool of Actin Monomers Are Useful for Studying Actin Dynamics

17.4 Organization of Actin-Based Cellular Structures

Cross-Linking Proteins Organize Actin Filaments into Bundles or Networks

Adapter Proteins Link Actin Filaments to Membranes

17.5 Myosins: Actin-Based Motor Proteins

Myosins Have Head, Neck, and Tail Domains with Distinct Functions

Myosins Make Up a Large Family of Mechanochemical Motor Proteins

Conformational Changes in the Myosin Head Couple ATP Hydrolysis to Movement

Myosin Heads Take Discrete Steps Along Actin Filaments

17.6 Myosin-Powered Movements

Myosin Thick Filaments and Actin Thin Filaments in Skeletal Muscle Slide Past Each Other During Contraction

Skeletal Muscle Is Structured by Stabilizing and Scaffolding Proteins

Contraction of Skeletal Muscle Is Regulated by Ca^{2+} and Actin-Binding Proteins

Actin and Myosin II Form Contractile Bundles in Nonmuscle Cells

Myosin-Dependent Mechanisms Regulate Contraction in Smooth Muscle and Nonmuscle Cells

Myosin V Carries Vesicles Along Actin Filaments

17.7 Cell Migration: Mechanism, Signaling, and Chemotaxis

Cell Migration Coordinates Force Generation with Cell Adhesion and Membrane Recycling

The Small GTP-Binding Proteins Cdc42, Rac, and Rho Control Actin Organization

Cell Migration Involves the Coordinate Regulation of Cdc42, Rac, and Rho

Migrating Cells Are Steered by Chemotactic Molecules

18 Cell Organization and Movement II: Microtubules and Intermediate Filaments

18.1 Microtubule Structure and Organization

Microtubule Walls Are Polarized Structures Built from αβ-Tubulin Dimers

Microtubules Are Assembled from MTOCs to Generate Diverse Configurations

18.2 Microtubule Dynamics

Individual Microtubules Exhibit Dynamic Instability

Localized Assembly and *Search and Capture* Help Organize Microtubules

Drugs Affecting Tubulin Polymerization Are Useful Experimentally and in Treatment of Diseases

18.3 Regulation of Microtubule Structure and Dynamics

Microtubules Are Stabilized by Side-Binding Proteins

+TIPs Regulate the Properties and Functions of the Microtubule (+) End

Other End-Binding Proteins Also Promote Microtubule Disassembly

Severing Proteins Also Regulate Microtubule Dynamics

18.4 Kinesins and Dyneins: Microtubule-Based Motor Proteins

Organelles in Axons Are Transported Along Microtubules in Both Directions

Kinesin-1 Powers Anterograde Transport of Vesicles Down Axons Toward the (+) Ends of Microtubules

The Kinesins Form a Large Protein Superfamily with Diverse Functions

Kinesin-1 Is a Processive Motor

Dynein Motors Transport Organelles Toward the (–) Ends of Microtubules

Kinesins and Dyneins Cooperate in the Transport of Organelles Throughout the Cell

Tubulin Modifications Distinguish Different Classes of Microtubules and Their Accessibility to Motors

18.5 Cilia and Flagella: Microtubule-Based Surface Structures

Eukaryotic Cilia and Flagella Contain Long Doublet Microtubules Bridged by Dynein Motors

Ciliary and Flagellar Beating Are Produced by Controlled Sliding of Outer Doublet Microtubules

Intraflagellar Transport Moves Material Up and Down Cilia and Flagella

Primary Cilia Are Sensory Organelles on Interphase Cells

Defects in Primary Cilia Underlie Many Diseases

18.6 Mitosis

Centrosomes Duplicate Early in the Cell Cycle in Preparation for Mitosis

Mitosis Can Be Divided into Five Stages

The Mitotic Spindle Contains Three Classes of Microtubules

Microtubule Dynamics Increase Dramatically in Mitosis

Chromosomes Are Captured and Oriented During Prometaphase

Duplicated Chromosomes Are Aligned by Motors and Microtubule Dynamics

The Chromosomal Passenger Complex Regulates Microtubule Attachment at Kinetochores

Anaphase A Moves Chromosomes to Poles by Microtubule Shortening

Anaphase B Separates Poles by the Combined Action of Kinesins and Dynein

The Spindle Is Centered and Oriented by a Dynein/Dynactin-Dependent Pathway

Cytokinesis Splits the Duplicated Cell in Two

Plant Cells Reorganize Their Microtubules and Build a New Cell Wall in Mitosis

18.7 Intermediate Filaments

Intermediate Filaments Are Assembled from Subunit Dimers

Intermediate Filaments Are Dynamic

Cytoplasmic Intermediate Filament Proteins Are Expressed in a Tissue-Specific Manner

Lamins Line the Inner Nuclear Envelope to Provide Organization and Rigidity to the Nucleus

Lamins Are Reversibly Disassembled by Phosphorylation During Mitosis

18.8 Coordination and Cooperation Between Cytoskeletal Elements

Intermediate Filament–Associated Proteins Contribute to Cellular Organization

Microfilaments and Microtubules Cooperate to Transport Melanosomes

Cdc42 Coordinates Microtubules and Microfilaments During Cell Migration

Advancement of Neural Growth Cones Is Coordinated by Microfilaments and Microtubules

19 The Eukaryotic Cell Cycle

19.1 Overview of the Cell Cycle

G_1 Controls Entry into S Phase

G_2 Phase Readies the Cell for Mitosis and Cell Division

Mitosis and Cytokinesis Occur During M-Phase

19.2 Model Organisms and Methods of Studying the Cell Cycle

Budding and Fission Yeasts Are Powerful Systems for Genetic Analysis of the Cell Cycle

Frog Oocytes and Early Embryos Facilitate Biochemical Characterization of the Cell Cycle Machinery

The Study of Tissue Culture Cells Uncovers Cell Cycle Regulation in Mammals

Researchers Use Multiple Tools to Study the Cell Cycle

19.3 Cell Cycle Progression and Control: Feedback Loops and Post-Translational Modification

Cyclin-Dependent Kinases Are Small Protein Kinases That Require a Regulatory Cyclin Subunit for Their Activity

Cyclins Determine the Activity of CDKs

CDKs Are Regulated by Activating and Inhibitory Phosphorylation

CDK Inhibitors Provide Additional Control of Cyclin-CDK Activity

Cyclin Levels Are Regulated by Transcriptional Activation and Ubiquitin-Mediated Protein Degradation

Phosphoserine/Threonine-Binding Domains Build Feedback Loops That Coordinate CDK Activation and Cell Cycle Progression

Mass Spectrometry Studies and Genetically Engineered CDKs Led to the Discovery of New CDK Substrates and Functions

19.4 The Transition from G_1 into S Phase and DNA Replication

The G_1/S Transition in Budding Yeast Is Controlled by Cyclin-CDK Complexes

The G_1–S Phase Transition in Metazoans Involves Cyclin-CDK Control of the E2F Transcription Factor Through Its Regulator Rb

Extracellular Signals Govern Cell Cycle Entry

Degradation of an S Phase CDK Inhibitor Triggers DNA Replication

Replication at Each Origin Is Initiated Once and Only Once During the Cell Cycle

Duplicated DNA Strands Become Linked During Replication

19.5 The G_2/M Transition and the Irreversible Engine of Mitosis

Precipitous Activation of Mitotic CDKs by Positive Feedback Loops Initiates Mitosis

Mitotic CDKs Promote Nuclear Envelope Breakdown

Centrosomes Duplicate During S phase and Separate During Mitosis

Mitotic CDKs, Polo-like Kinases, and Aurora Kinases Drive Assembly of a Mitotic Spindle That Attaches to the Kinetochores of Condensed Chromosomes

Chromosome Condensation Facilitates Chromosome Segregation

19.6 The Mitotic Spindle, Chromosome Segregation, and Exit from Mitosis

Separase-Mediated Cleavage of Cohesins Initiates Chromosome Segregation

APC/C Activates Separase Through Securin Ubiquitinylation

Mitotic CDK Inactivation and Protein Dephosphorylation Triggers Exit from Mitosis

Cytokinesis Creates Two Daughter Cells

19.7 Surveillance Mechanisms in Cell Cycle Regulation

The DNA Damage Response System Halts Cell Cycle Progression and Recruits DNA Repair Machinery When DNA Is Compromised

The Spindle Assembly Checkpoint Pathway Prevents Chromosome Segregation Until Chromosomes Are Accurately Attached to the Mitotic Spindle

19.8 Meiosis: A Special Type of Cell Division

Extracellular and Intracellular Cues Regulate Germ Cell Formation

Several Features Distinguish Meiosis from Mitosis

Recombination and a Meiosis-Specific Cohesin Subunit Are Necessary for the Specialized Chromosome Segregation in Meiosis I

Co-orienting Sister Kinetochores Is Critical for Meiosis I Chromosome Segregation

Part IV Cell Growth and Differentiation

20 Integrating Cells into Tissues

20.1 Cell-Cell and Cell–Extracellular Matrix Adhesion: An Overview

Cell-Adhesion Molecules Bind to One Another and to Intracellular Proteins

The Extracellular Matrix Participates in Adhesion, Signaling, and Other Functions

The Evolution of Multifaceted Adhesion Molecules Enabled the Evolution of Diverse Animal Tissues

Cell-Adhesion Molecules Mediate Mechanotransduction

20.2 Cell-Cell and Cell–Extracellular Matrix Junctions and Their Adhesion Molecules

Epithelial Cells Have Distinct Apical, Lateral, and Basal Surfaces

Three Types of Junctions Mediate Many Cell-Cell and Cell-Matrix Interactions

Cadherins Mediate Cell-Cell Adhesions in Adherens Junctions and Desmosomes

Integrins Mediate Cell-Matrix Adhesions, Including Those in Epithelial-Cell Hemidesmosomes

Tight Junctions Seal Off Body Cavities and Restrict Diffusion of Membrane Components

Gap Junctions Composed of Connexins Allow Small Molecules to Pass Directly Between the Cytosols of Adjacent Cells

Tunneling Nanotubes Can Mediate Metabolic Coupling and Transfer Organelles Between Animal Cells

20.3 The Extracellular Matrix I: The Basal Lamina

The Basal Lamina Provides a Foundation for Assembly of Cells into Tissues

Laminin, a Multi-Adhesive Matrix Protein, Helps Cross-Link Components of the Basal Lamina

Sheet-Forming Type IV Collagen Is a Major Structural Component of the Basal Lamina

Perlecan, a Proteoglycan, Cross-Links Components of the Basal Lamina and Cell-Surface Receptors

20.4 The Extracellular Matrix II: Connective Tissue

Fibrillar Collagens Are the Major Fibrous Proteins in the ECM of Connective Tissues

Fibrillar Collagen Is Secreted and Assembled into Fibrils Outside the Cell

Type I and II Collagens Associate with Nonfibrillar Collagens to Form Diverse Structures

Proteoglycans and Their Constituent GAGs Play Diverse Roles in the ECM

Hyaluronan Resists Compression, Facilitates Cell Migration, and Gives Cartilage Its Gel-Like Properties

Fibronectins Connect Cells and ECM, Influencing Cell Shape, Differentiation, and Movement

Elastic Fibers Permit Many Tissues to Undergo Repeated Stretching and Recoiling

Metalloproteases Remodel and Degrade the Extracellular Matrix

20.5 Adhesive Interactions in Motile and Nonmotile Cells

Integrins Mediate Adhesion and Relay Signals Between Cells and Their Three-Dimensional Environment

Regulation of Integrin-Mediated Adhesion and Signaling Controls Cell Function and Movement

Connections Between the ECM and Cytoskeleton Are Defective in Muscular Dystrophy

IgCAMs Mediate Cell-Cell Adhesion in Neural and Other Tissues

Leukocyte Movement into Tissues Is Orchestrated by a Precisely Timed Sequence of Adhesive Interactions

20.6 Plant Tissues

The Plant Cell Wall, a Plant's ECM, Is a Laminate of Cellulose Fibrils in a Matrix of Polysaccharides and Glycoproteins

Loosening of the Cell Wall Permits Plant Cell Growth

Plasmodesmata Directly Connect the Cytosols of Adjacent Cells

The Molecules That Plants Depend on for Adhesion and Mechanotransduction Differ from Those in Animals

21 Responding to the Cellular Environment

21.1 Regulating Blood Glucose Level

Insulin and Glucagon Work Together to Maintain a Stable Blood Glucose Level

A Rise in Blood Glucose Triggers Insulin Secretion from the β Islet Cells

In Fat and Muscle Cells, Insulin Triggers Fusion of Intracellular Vesicles Containing the GLUT4 Glucose Transporter with the Plasma Membrane, Thus Increasing the Rate of Glucose Uptake

In the Liver, Insulin Inhibits Glucose Synthesis, Accelerates the Rate of Glycolysis, and Enhances Storage of Glucose as Glycogen

21.2 Integrating Cell Growth Signals with Nutrient and Energy Levels

The Active mTORC1 Complex Activates Many Anabolic Signal Transduction Pathways

mTORC1 Kinase Activation Requires Amino Acids, a High ATP:AMP Ratio, and Activation of Signal Transduction Pathways Downstream of Growth-Factor Receptors

21.3 Responding to Changes in the Levels of Cholesterol and Unsaturated Fatty Acids

Fatty Acid and Cholesterol Biosynthesis as Well as Cholesterol Import Are Regulated at the Level of Gene Transcription

The Endoplasmic Reticulum SCAP Protein Senses the Level of Cellular Cholesterol

Regulated Intramembrane Proteolysis of SREBP in the Golgi Releases a bHLH Transcription Factor That Acts to Maintain Appropriate Phospholipid and Cholesterol Levels

21.4 Responding to Low Oxygen

Induction of the Erythropoietin Gene at Low Oxygen Levels

Oxygen Sensing and Regulated Hif-1α Expression Is a Property of All Nucleated Mammalian Cells

Hif-1α Function and Stability Are Blocked at Ambient Oxygen Levels

A Conserved Family of Oxygen-Sensitive Transcription Factors Found in Plants and Animals Is Regulated by Post-Translational Addition of an Arginine Residue

21.5 Responding to Elevated Temperatures

The Heat-Shock Response Is Induced by Unfolded Polypeptide Chains

The Heat-Shock Response Is Regulated Primarily by Related Transcription Factors in All Eukaryotes Called *Heat-Shock Factors*, Including HSF1 in Humans

21.6 Sensing Night and Day: Circadian Rhythms

The Circadian Clock in Most Organisms Relies on a Negative Feedback Loop

The Circadian Clock in Bacteria: A Different Solution

Suprachiasmatic Nucleus: The Master Clock in Mammals

21.7 Sensing and Responding to the Physical Environment

The Hippo Kinase Cascade Pathway in *Drosophila* and Mammals

Regulation of the Hippo Kinase Cascade by Cell Interactions with the Extracellular Matrix and by Tension on Actin Filaments

The Hippo Pathway and Early Embryogenesis

22 Stem Cells, Cell Asymmetry, and Regulated Cell Death

22.1 Early Mammalian Development, Embryonic Stem Cells, and Induced Pluripotent Stem Cells

Fertilization Unifies the Genome

Cleavage of the Mammalian Embryo Leads to the First Differentiation Events

Pluripotent Cells of the Inner Cell Mass Are the Source of ES Cells

Multiple Factors Control the Pluripotency of ES Cells

Animal Cloning Shows That Epigenetic Changes During Differentiation Can Be Reversed

Somatic Cells Can Generate iPS Cells

Patient-Specific iPS Cells Can Be Used to Develop Potential Treatments for Many Diseases

ES and iPS Cells Can Generate Functional Differentiated Human Cells

22.2 Stem Cells and Niches in Multicellular Organisms

Adult Planarians Contain Pluripotent Stem Cells

Multipotent Somatic Stem Cells Give Rise to Both Stem Cells and Differentiating Cells

Stem Cells for Different Tissues Occupy Sustaining Niches

Germ-Line Stem Cells in Many Organisms Produce Sperm or Oocytes

Intestinal Stem Cells Continuously Generate All the Cells of the Intestinal Epithelium

Wnt and R-Spondins Are Essential for Function of the $Lgr5^+$ Intestinal Stem Cells

Hematopoietic Stem Cells Form All Blood Cells and All Cells of the Immune System

Characterizing Hematopoietic Stem Cells by Transplantation

Niches for Hematopoietic Stem Cells and Many Hematopoietic Progenitor Cells

Regulating the Production of Differentiated Hematopoietic Cells

Meristems Are Niches for Stem Cells in Plants

A Negative Feedback Loop Maintains the Size of the Shoot Apical Stem-Cell Population

The Root Meristem Resembles the Shoot Meristem in Structure and Function

22.3 Mechanisms of Cell Polarity and Asymmetric Cell Division

The Intrinsic Polarity Program Depends on a Positive Feedback Loop Involving Cdc42

Cell Polarization Before Cell Division Follows a Common Hierarchy of Steps

Polarized Membrane Traffic Allows Yeast to Grow Asymmetrically During Mating

The Par Proteins Direct Cell Asymmetry in the Nematode Embryo

The Par Proteins and Other Polarity Complexes Are Involved in Epithelial-Cell Polarity

The Planar Cell Polarity Pathway Orients Cells Within an Epithelium

The Par Proteins Are Involved in Asymmetric Division of Stem Cells

22.4 Cell Death and Its Regulation

Most Programmed Cell Death Occurs Through Apoptosis

Evolutionarily Conserved Proteins Participate in the Apoptotic Pathway

Caspases Amplify the Initial Apoptotic Signal and Destroy Key Cellular Proteins

Phosphatidylserine: an “Eat Me” Signal on the Surface of Apoptotic Cells

Neurotrophins Promote Survival of Neurons

Mitochondria Play a Central Role in Regulation of Apoptosis in Vertebrate Cells

The Pro-Apoptotic Proteins Bax and Bak Form Pores and Holes in the Outer Mitochondrial Membrane

Release of SMAC/DIABLO Proteins from Mitochondria Also Promotes Caspase Activation

Trophic Factors Induce Inactivation of Bad, a Pro-apoptotic BH3-Only Protein

Apoptosis in Vertebrates Is Induced by BH3-Only Pro-apoptotic Proteins That Are Activated by Environmental Stresses

Apoptosis and Necroptosis Can Be Triggered by Tumor Necrosis Factor, Fas Ligand, and Related Death Proteins

23 Cells of the Nervous System

23.1 Neurons and Glia: Building Blocks of the Nervous System

Information Flows Through Neurons from Dendrites to Axons

Information Moves Along Axons as Pulses of Ion Flow Called *Action Potentials*

Information Flows Between Neurons via Synapses

The Nervous System Uses Signaling Circuits Composed of Multiple Types of Neurons

Glial Cells Form Myelin Sheaths and Support Neurons

Neural Stem Cells Form Nerve and Glial Cells in the Central Nervous System

23.2 Voltage-Gated Ion Channels and the Propagation of Action Potentials

The Magnitude of the Action Potential Is Close to E_{Na} and Is Caused by Na^+ Influx Through Open Na^+ Channels

Sequential Opening and Closing of Voltage-Gated Na^+ and K^+ Channels Generate Action Potentials

Action Potentials Are Propagated Unidirectionally Without Diminution

All Voltage-Gated Ion Channels Have Similar Structures

Voltage-Sensing S4 α Helices Move in Response to Membrane Depolarization

Movement of the Channel-Inactivating Segment into the Open Pore Blocks Ion Flow

Myelination Increases the Velocity of Impulse Conduction

Action Potentials "Jump" from Node to Node in Myelinated Axons

Two Types of Glia Produce Myelin Sheaths

Light-Activated Ion Channels and Optogenetics

23.3 Communication at Synapses

Formation of Synapses Requires Assembly of Presynaptic and Postsynaptic Structures

Neurotransmitters Are Transported into Synaptic Vesicles by H^+-Linked Antiport Proteins

Three Pools of Synaptic Vesicles Loaded with Neurotransmitter Are Present in the Presynaptic Terminal

Influx of Ca^{2+} Triggers Release of Neurotransmitters

A Calcium-Binding Protein Regulates Fusion of Synaptic Vesicles with the Plasma Membrane

Fly Mutants Lacking Dynamin Cannot Recycle Synaptic Vesicles

Signaling at Synapses Is Terminated by Degradation or Reuptake of Neurotransmitters

Opening of Acetylcholine-Gated Cation Channels Leads to Muscle Contraction

All Five Subunits in the Nicotinic Acetylcholine Receptor Contribute to the Ion Channel

Nerve Cells Integrate Many Inputs to Make an All-or-None Decision to Generate an Action Potential

Gap Junctions Allow Direct Communication Between Neurons and Between Glia

23.4 Sensing the Environment: Touch, Pain, Taste, and Smell

Mechanoreceptors Are Gated Cation Channels

Pain Receptors Are Also Gated Cation Channels

Five Primary Tastes Are Sensed by Subsets of Cells in Each Taste Bud

The Largest Group of G Protein–Coupled Receptors Detect Odors

Each Olfactory Receptor Neuron Expresses a Single Type of Odorant Receptor

23.5 Forming and Storing Memories

Memories Are Formed by Changing the Number or Strength of Synapses Between Neurons

The Hippocampus Is Required for Memory Formation

Multiple Molecular Mechanisms Contribute to Synaptic Plasticity

Formation of Long-Term Memories Requires Gene Expression

24 Immunology

24.1 Overview of Host Defenses

Pathogens Enter the Body Through Different Routes and Replicate at Different Sites

Cells of the Innate and Adaptive Immune Systems Circulate Throughout the Body and Take Up Residence in Tissues and Lymph Nodes

Mechanical and Chemical Boundaries Form a First Layer of Defense Against Pathogens

Innate Immunity Provides a Second Line of Defense

Inflammation Is a Complex Response to Injury That Encompasses Both Innate and Adaptive Immunity and Helps Destroy Pathogens

Adaptive Immunity, the Third Line of Defense, Exhibits Specificity

24.2 Immunoglobulins: Structure and Function

Immunoglobulins Have a Conserved Structure Consisting of Heavy and Light Chains

Multiple Immunoglobulin Isotypes Exist, Each with Different Functions

Each Naive B Cell Produces a Unique Immunoglobulin

Immunoglobulin Domains Have a Characteristic Fold Composed of β Two Sheets Stabilized by a Disulfide Bond

An Immunoglobulin's Constant Region Determines Its Functional Properties

24.3 Generation of Antibody Diversity and B-Cell Development

A Functional Light-Chain Gene Requires Assembly of V and J Gene Segments

Rearrangement of the Heavy-Chain Locus Involves V, D, and J Gene Segments

Somatic Hypermutation Allows the Generation and Selection of Antibodies with Improved Affinities

B-Cell Development Requires Input from a Pre-B-Cell Receptor

During an Adaptive Response, B Cells Switch from Making Membrane-Bound Ig to Making Secreted Ig

B Cells Can Switch the Isotype of Immunoglobulin They Make

24.4 The MHC and Antigen Presentation

The MHC Determines the Ability of Two Unrelated Individuals of the Same Species to Accept or Reject Grafts

The Killing Activity of Cytotoxic T Cells Is Antigen Specific and MHC Restricted

T Cells with Different Functional Properties Are Guided by Two Distinct Classes of MHC Molecules

MHC Molecules Are Highly Polymorphic, Bind Peptide Antigens, and Interact with the T-Cell Receptor

In Antigen Presentation, Protein Fragments Are Complexed with MHC Products and Posted to the Cell Surface

The Class I MHC Pathway Presents Cytosolic Antigens

The Class II MHC Pathway Presents Antigens Delivered to the Endocytic Pathway

24.5 T Cells, T-Cell Receptors, and T-Cell Development

The Structure of the T-Cell Receptor Resembles the F(ab) Portion of an Immunoglobulin

TCR Genes Are Rearranged in a Manner Similar to Immunoglobulin Genes

Many of the Variable Residues of TCRs Are Encoded in the Junctions Between V, D, and J Gene Segments

Signaling via Antigen-Specific Receptors Triggers Proliferation and Differentiation of T and B Cells

T Cells Capable of Recognizing MHC Molecules Develop Through a Process of Positive and Negative Selection

T Cells Commit to the CD4 or CD8 Lineage in the Thymus

T Cells Require Two Types of Signals for Full Activation

Cytotoxic T Cells Carry the CD8 Co-Receptor and Are Specialized for Killing

T Cells Secrete an Array of Cytokines That Provide Signals to Other Immune-System Cells

Helper T Cells Are Divided into Distinct Subsets Based on Their Cytokine Production and Expression of Surface Markers

Innate Lymphoid Cells Regulate Inflammation and the Overall Immune Response

Leukocytes Move in Response to Chemotactic Cues Provided by Chemokines

24.6 Collaboration of Immune-System Cells in the Adaptive Response

Toll-Like Receptors Perceive a Variety of Pathogen-Derived Macromolecular Patterns

Engagement of Toll-Like Receptors Leads to Activation of Antigen-Presenting Cells

Production of High-Affinity Antibodies Requires Collaboration Between B and T cells

Vaccines Elicit Protective Immunity Against a Variety of Pathogens

The Immune System Defends Against Cancer

25 Cancer

25.1 How Tumor Cells Differ from Normal Cells

The Genetic Makeup of Most Cancer Cells Is Dramatically Altered

Uncontrolled Proliferation Is a Universal Trait of Cancer

Cellular Housekeeping Functions Are Fundamentally Altered in Cancer Cells

Cancer Cells Exhibit Altered Cell-Cell Interactions to Form Heterogeneous Organs

Tumor Growth Requires Formation of New Blood Vessels

Invasion and Metastasis Are Late Stages of Tumorigenesis

25.2 Genetic and Genomic Basis of Cancer

Carcinogens Induce Cancer by Damaging DNA

Some Carcinogens Have Been Linked to Specific Cancers

Familial Syndromes that Cause Loss of DNA Repair Can Lead to Cancer

Somatic Mutations in the DNA Damage Response Pathway Are Oncogenic

Cancer Genome Sequencing Reveals an Enormous Diversity of Somatic Mutations

Oncogenes Were Discovered by Their Association with Tumor Viruses

Single Oncogenic Drivers Can Be Activated by Chromosome Rearrangements

Inherited Predisposition for Cancer Enabled Identification of Some Oncogenic Drivers

Oncogenic Driver Mutations Have Been Identified in Many Genes

Oncogenic Driver Mutations Can Be Identified by Comparing Cancer Genomes

Oncogenic Drivers Can Be Gain-of-Function or Loss-of-Function Mutations

Tumor Suppressor Genes and Oncogenes Often Operate in the Same Pathway

MicroRNAs Can Promote and Inhibit Tumorigenesis

Epigenetic Changes Can Contribute to Tumorigenesis

25.3 Dysregulation of Cell Growth and Developmental Pathways Initiates Tumorigenesis

Receptor Mutations Can Cause Proliferation in the Absence of External Growth Factors

Many Oncogenic Mutations Constitutively Activate Signal-Transducing Proteins

Growth Control Pathways Ultimately Regulate Initiation of the Cell Cycle

Inappropriate Production of Nuclear Transcription Factors Can Induce Transformation

Aberrations in Signaling Pathways That Control Development Are Associated with Many Cancers

Experimental Reconstruction of the Multi-Hit Model for Cancer

The Succession of Oncogenic Mutations Can Be Traced in Colon Cancers

Cancer Development Can Be Studied in Animal Models

Molecular Cell Biology Is Changing How Cancer Is Diagnosed and Treated

25.4 Evasion of Programmed Cell Death and Immune Surveillance Processes

Oncogenic Driver Mutations Enable Cancer Cells to Evade Apoptosis

p53 Can Activate Either the DNA Damage Checkpoint or Apoptosis in Response to DNA Damage

The Immune System Is a Second Line of Defense Against Cancer Formation

The Tumor Microenvironment and Immunoediting Limit the Ability of the Immune System to Detect and Kill Established Tumors

Activation of the Immune System Presents a Major Opening for Cancer Therapy

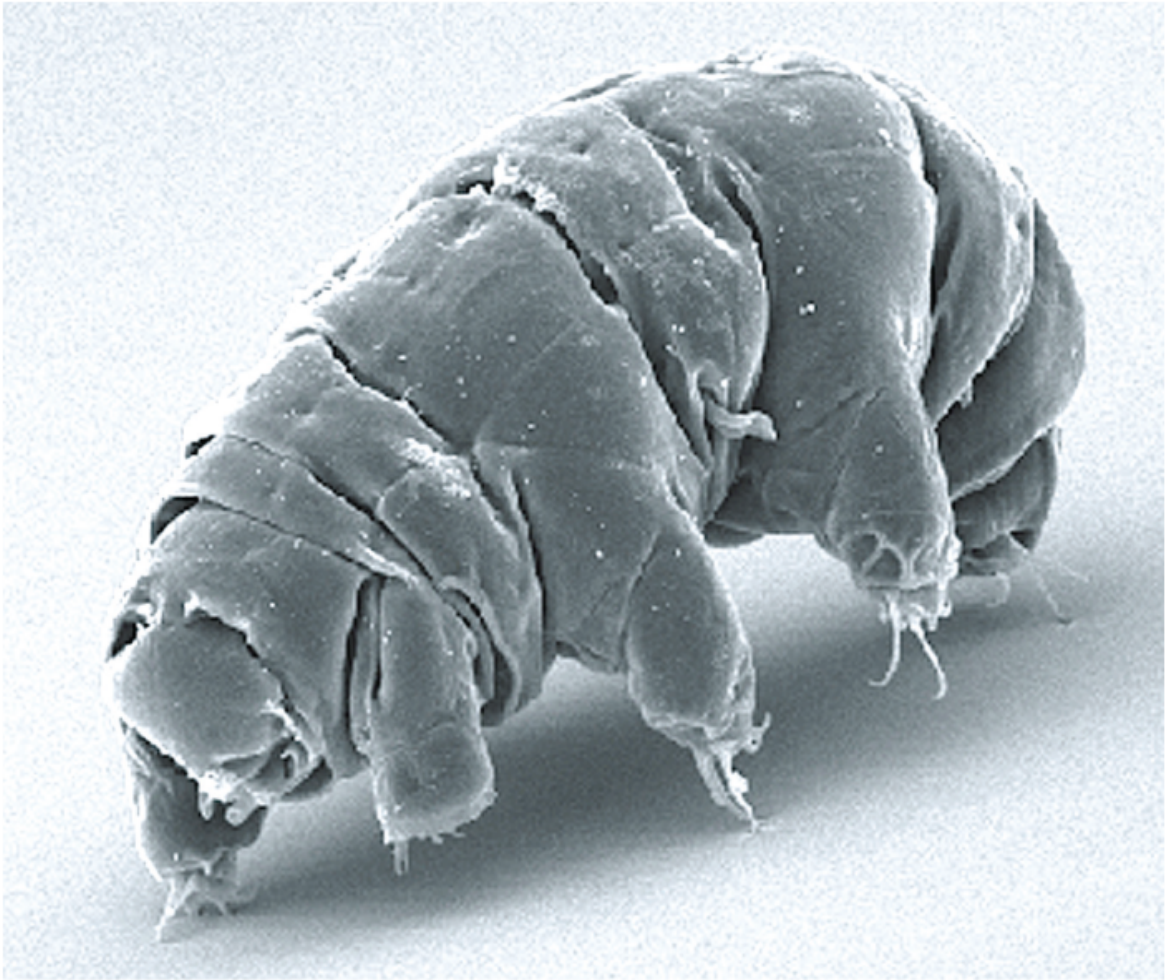
Index

Glossary (online and e-book only)

References (online and e-book only)

CHAPTER 1

Evolution: Molecules, Genes, Cells, and Organisms



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Tardigrades are tiny 0.05-mm-long invertebrates that have evolved to survive exposures to extreme conditions that kill other life forms; they are the only metazoan organisms known to survive exposure to the deadly combination of low pressure and intense radiation in outer space.

[Schokraie E., Warnken U., Hotz-Wagenblatt A., Grohme M. A., Hengherr S., et al. 2012. Comparative proteome analysis of *Milnesium tardigradum* in early embryonic state *versus* adults in active and anhydrobiotic state. *PLoS ONE* **7**(9): e45682. https://doi.org./10.1371/journal.pone.0045682]

OUTLINE

1.1 The Molecules of Life
1.2 Prokaryotic Cell Structure and Function
1.3 Eukaryotic Cell Structure and Function
1.4 Unicellular Eukaryotic Organisms Widely Used in Cell Biology Research
1.5 Metazoan Structure, Function, Evolution, and Differentiation
1.6 Metazoan Organisms Widely Used in Cell Biology Research

> *Nothing in biology makes sense except in the light of evolution.*
>
> — Theodosius Dobzhansky, 1973, essay in *American Biology Teacher* **35**:125–129

Biology is a science fundamentally different from physics or chemistry, which deal with unchanging properties of matter that can be described by mathematical equations. Biological systems, of course, follow the rules of chemistry and physics, but biology is a historical science, as the forms and structures of the living world today are the results of billions of years of evolution. Through evolution, all organisms both living and extinct are related in an

evolutionary family tree extending from primitive single-celled organisms that lived in the distant past to the diverse plants, animals, and microorganisms of the present era (Figure 1-1, Table 1-1). The great insight of Charles Darwin was the principle of natural selection: organisms vary randomly and compete within their environment for resources. Only those that survive and reproduce are able to pass down their genetic traits (Figure 1-2).

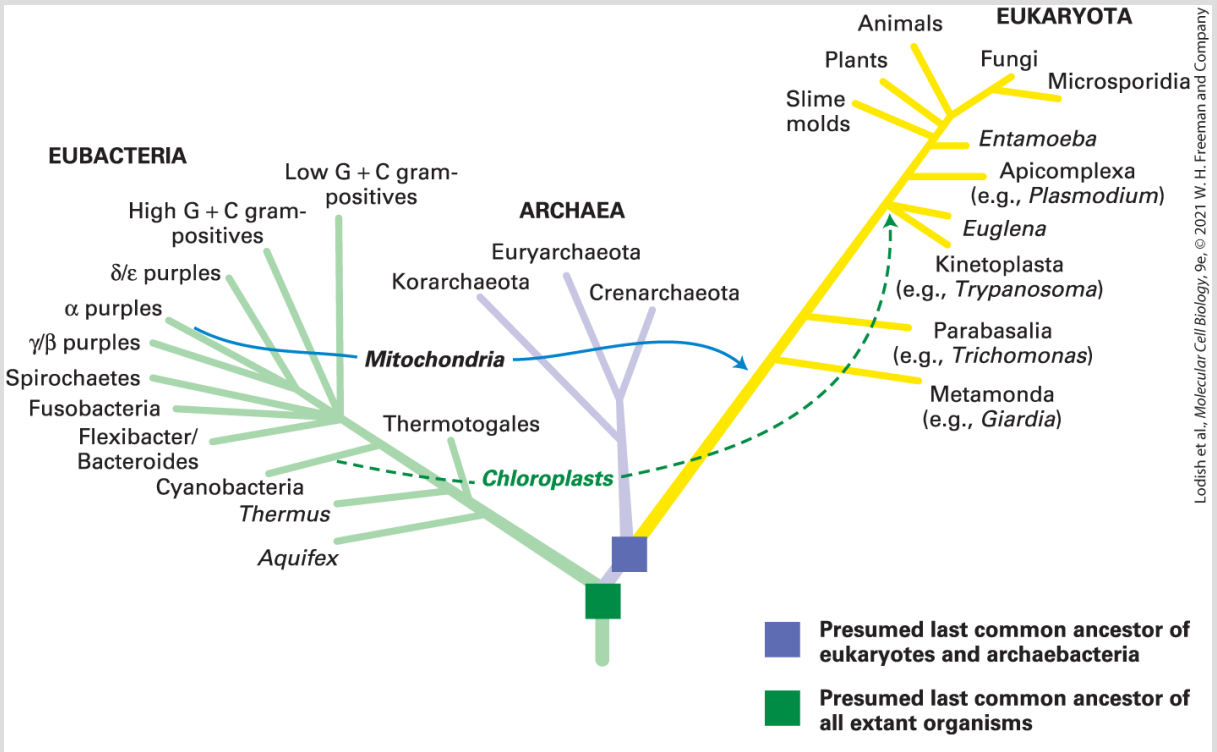


FIGURE 1-1 All living organisms descended from a common ancestral cell. All organisms, from simple bacteria to complex mammals, probably evolved from a common single-celled ancestor. This evolutionary tree depicts the relationships among the three major lineages of organisms. The structure of the tree was initially ascertained from morphological criteria: creatures that look alike were put close together. More recently, the sequences of DNA and proteins found in organisms have provided more information-rich criteria for assigning relationships. The greater the similarities in these macromolecular sequences, the more closely related organisms are thought to be. The trees based on morphological comparisons and the fossil record generally agree well with those based on molecular data. This

evolutionary tree has branches. Organisms like bacteria exchange pieces of their DNA genome with one another, and mitochondria and chloroplasts are thought to derive from primitive bacteria that took up residence inside the organisms that gave rise to present eukaryotes.

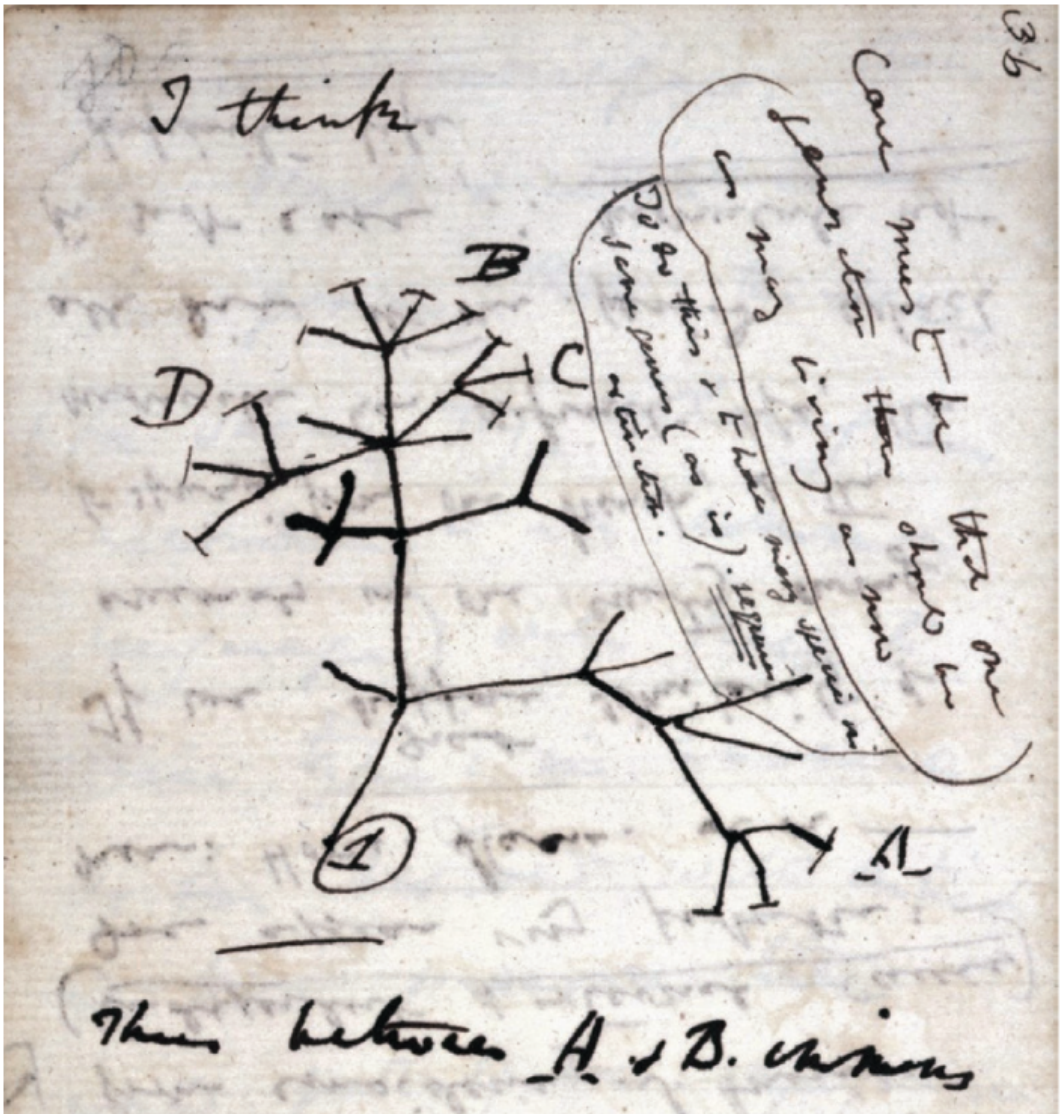
[Data from J. R. Brown, 2005, "Universal Tree of Life," in *Encyclopedia of Life Sciences*, Wiley InterScience (online).]

TABLE 1-1 • Timeline for Evolution of Life on Earth, as Determined from the Fossil Record and from Recent DNA Sequence Analyses

4600 million years ago	The planet Earth forms from material revolving around the young sun.
$\sim$3500 million years ago	Cells resembling prokaryotes appear. These first organisms are chemoautotrophs: they use carbon dioxide as a carbon source and oxidize inorganic materials to extract energy.
3000 million years ago	Photosynthesizing cyanobacteria evolve; they use water as a reducing agent, thereby producing oxygen as a waste product.
1850 million years ago	Unicellular eukaryotic cells appear.
1200 million years ago	Simple multicellular organisms evolve, mostly consisting of cell colonies of limited complexity.
580–500 million years ago	Most modern phyla of animals begin to appear in the fossil record during the Cambrian explosion.

535 million years ago	Major diversification of living things in the oceans: chordates, arthropods (e.g., trilobites, crustaceans), echinoderms, mollusks, brachiopods, foraminifers, radiolarians, etc.
485 million years ago	First vertebrates with true bones (jawless fishes) evolve.
434 million years ago	First primitive plants arise on land.
225 million years ago	Earliest dinosaurs (prosauropods) and teleost fishes appear.
220 million years ago	Gymnosperm forests dominate the land; herbivores grow to huge sizes.
215 million years ago	First mammals evolve.
65.5 million years ago	The Cretaceous-Tertiary extinction event eradicates about half of all animal species, including all of the dinosaurs.
6.5 million years ago	First hominids evolve.
~ 2 million years ago	First members of the genus *Homo* appear in the fossil record.

~450 thousand years ago	Neanderthals appear.
~300 thousand years ago	Anatomically modern humans appear in Africa.
~40 thousand years ago	Extinction of Neanderthals.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

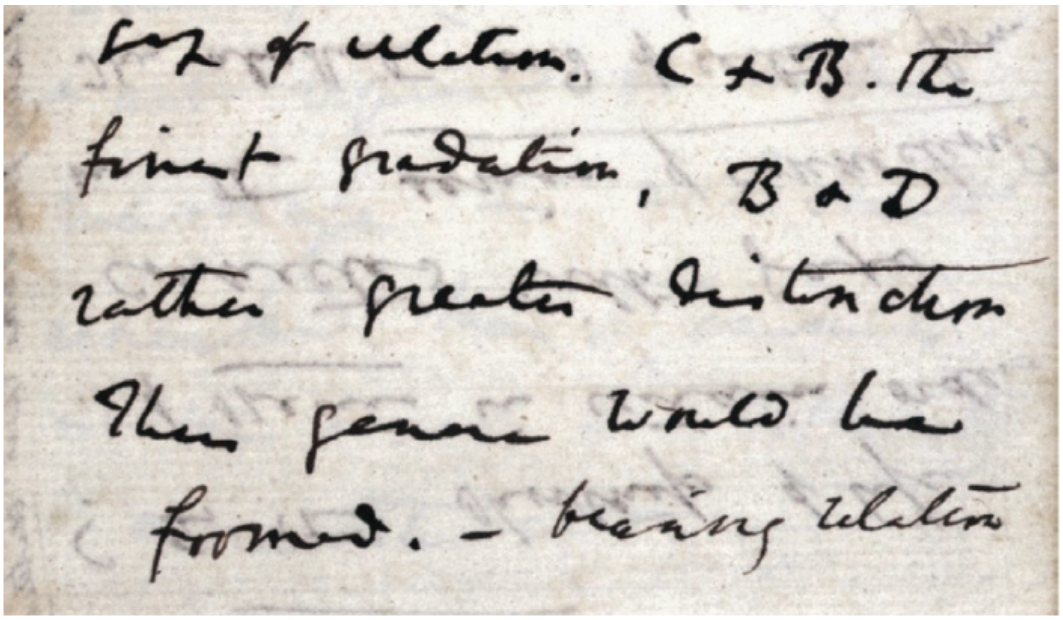


FIGURE 1-2 The first evolutionary tree in Charles Darwin's notebook. One year after his epic voyage on HMS *Beagle*, Darwin had already begun formulating his concept of natural selection, which would be published in his *On the Origin of Species* (1859). Shown here is page 36 of his 1837 "B" notebook on *Transmutation of Species;* he wrote "I think" above his first evolutionary tree. In *On the Origin of Species* he describes in greater detail the evolutionary history of organisms both living and extinct and includes a more elaborate hypothetical evolutionary tree.

[Page 36 of Charles Darwin's 1837 "B" notebook on Transmutation of Species.]

At first glance, the biological universe does appear amazingly diverse — from tiny ferns to tall fir trees, from single-celled bacteria and protozoans visible only under a microscope to multicellular animals of all kinds. The discovery that all living organisms are made up of cells was one of the most remarkable discoveries in biology. Indeed, cells come in an astonishing variety of sizes and shapes (Figure 1-3). Some move rapidly and have fast-changing structures, as we can see in movies of amoebae and rotifers. Others

are largely stationary and structurally stable. Oxygen kills some cells but is an absolute requirement for others. Most cells in multicellular organisms are intimately involved with other cells. Although some unicellular organisms live in isolation (Figure 1-3a), others form colonies or live in close association with other types of organisms (Figure 1-3b, d), such as the bacteria that help plants extract nitrogen from the air or the bacteria that live in our intestines and help us digest food.

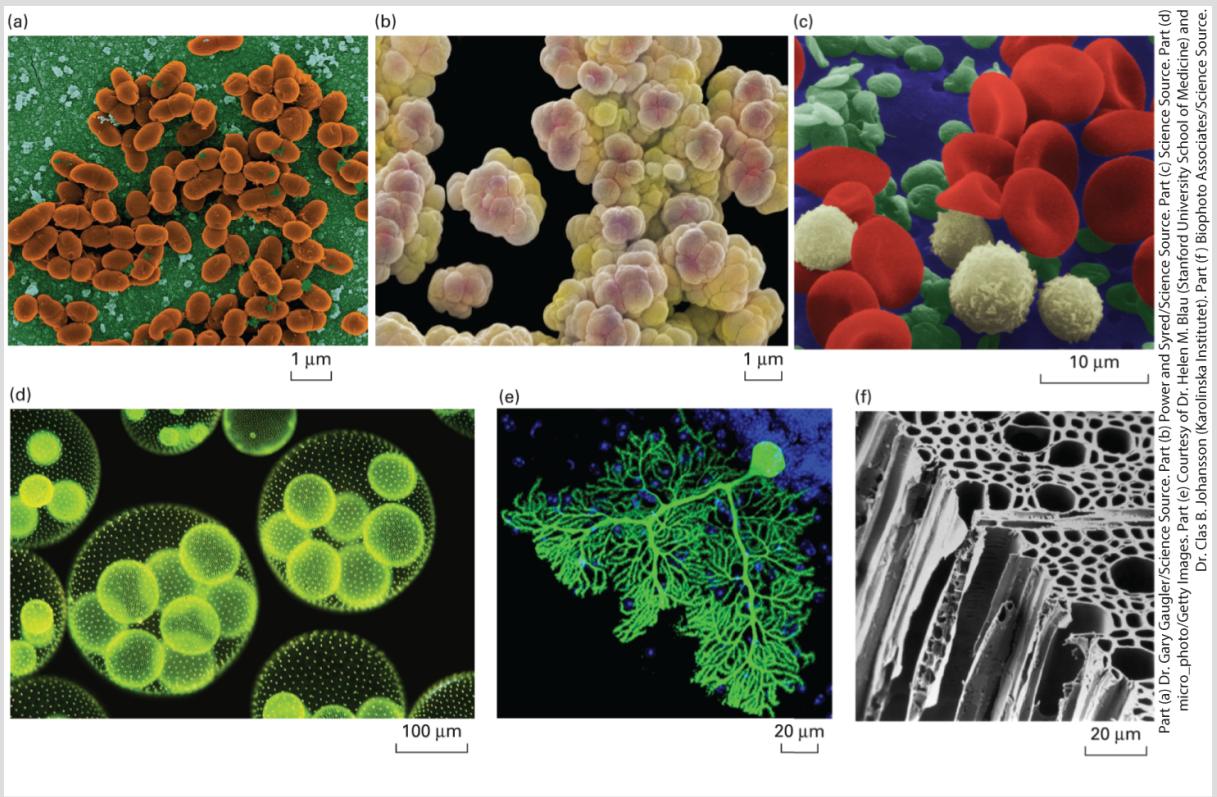


FIGURE 1-3 Cells come in an astounding assortment of shapes and sizes. Some of the morphological variety of cells is illustrated in these photographs. In addition to morphology, cells differ in their ability to move, internal organization (prokaryotic versus eukaryotic cells), and metabolic activities. (a) Eubacteria: *Lactococcus lactis*, which are used to produce cheese such as Roquefort, Brie, and Camembert. Note the dividing cells. (b) A mass of archaeans (*Methanosarcina*) that produce their energy by converting carbon dioxide and hydrogen gas to methane. Some species that live in the rumens of cattle give rise to

>150 liters of methane gas each day. (c) Human blood cells, shown in false color. The red cells are oxygen-bearing erythrocytes, the white cells (leukocytes) are part of the immune system and fight infection, and the green false-colored cell fragments are platelets that plug wounds and contain substances to initiate blood clotting. (d) A colonial single-celled green alga, *Volvox aureus*. The surfaces of the large spheres contain thousands of individual cells, visible as blue or green dots. The yellow masses inside are daughter colonies, each made up of many cells. (e) A single Purkinje neuron of the cerebellum, which can form more than a hundred thousand connections with other cells through its branched network of dendrites. The cell was made visible by introduction of a gene-encoding green fluorescent protein; the cell body with its nucleus is the bulb at the upper right. (f) Plant cells in vascular plants are fixed firmly in place, supported by a rigid cellulose skeleton. Spaces between the cells are joined into tubes for transport of water and food.

Yet the bewildering array of outward biological forms overlies a powerful uniformity: thanks to our common ancestry, all biological systems are composed of cells containing the same types of chemical molecules and employing similar principles of organization and function at the cellular level. Although the basic kinds of biological molecules have been conserved during the billions of years of evolution, the patterns in which they are assembled to form functioning cells and organisms have undergone considerable change.

We now know that **genes**, which chemically are composed of **deoxyribonucleic acid (DNA)**, ultimately define biological structure and maintain the integration of cellular function. Many genes are copied into **messenger RNA (mRNA)** that in turn directs the synthesis of specific **proteins**, the primary molecules that make up cell structures and carry out virtually all cellular activities.

Alterations in the structure and organization of genes, or **mutations**, provide the random variation that can alter biological structure and function. While the vast majority of random mutations have no observable effect on a gene's or protein's function, many are deleterious, and only a few confer an evolutionary advantage on the organism. In all organisms, mutations in DNA are constantly occurring, allowing over time the small alterations in cellular structures and functions that may prove to be advantageous. Entirely new cellular structures are rarely created; more often, existing cellular structures undergo changes that better adapt the organism to new circumstances. Slight changes in a protein can cause important changes in its function or abolish its function entirely.

For instance, in a particular organism, one gene may randomly become duplicated, after which one copy of the gene and its encoded protein retain their original function while, over time, the second copy of the gene mutates such that its protein takes on a slightly different or even a totally new function. During the evolution of some organisms, the entire genome became duplicated, allowing the second copies of many genes to undergo mutations and acquire new functions. The cellular organization of organisms plays a fundamental role in this process because it allows these changes to come about by small alterations in previously evolved cells, giving them new abilities. The result is that closely related organisms have very similar genes and proteins as well as similar cellular and tissue organizations.

Multicellular organisms, including the human body, consist of such closely interrelated elements that no single element can be fully appreciated in isolation from the others. Organisms contain organs, organs are composed of tissues, tissues consist of cells, and cells are formed from molecules (Figure 1-4). The unity of living systems is coordinated by many levels of interrelationship: molecules carry messages from organ to organ and cell to cell, and tissues are delineated and integrated with other tissues by molecules secreted by cells. Generally all the levels into which we fragment biological systems interconnect.

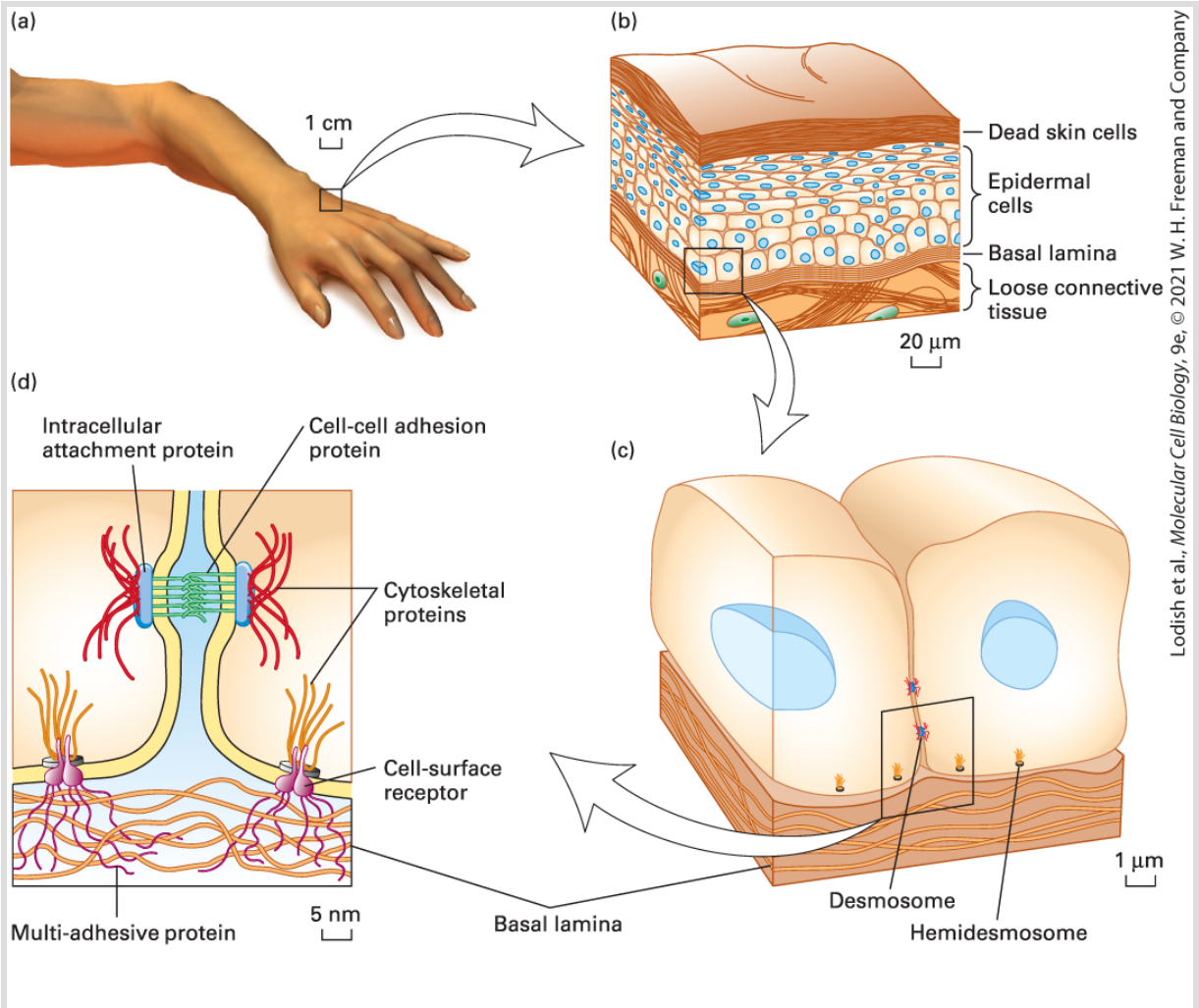


FIGURE 1-4 Living systems such as the human body consist of closely interrelated elements. (a) The surface of the hand is covered by a living organ, skin, that is composed of several layers of tissue. (b) An outer covering of hard, dead skin cells protects the body from injury, infection, and dehydration. This layer is constantly renewed by living epidermal cells, which also give rise to hair and fur in animals. Deeper layers of muscle and connective tissue give skin its tone and firmness. (c) Tissues are formed through subcellular adhesion structures (desmosomes and hemidesmosomes) that join cells to one another and to an underlying layer of supporting fibers. (d) At the heart of cell-cell adhesion are its structural components: phospholipid molecules that make up the cell-surface membrane, and large protein molecules. Protein molecules that traverse the cell membrane often form strong bonds with internal and external fibers made of multiple proteins.

To learn about biological systems, however, we must examine one small portion of a living system at a time. The biology of cells is a logical starting point because an organism can be viewed as consisting of interacting cells, which are the closest thing to autonomous biological units that exist. The last common ancestor of all life on Earth was a single cell (see Figure 1-1), and at the cellular level all life is remarkably similar. All cells use the same molecular building blocks; similar methods for the storage, maintenance, and expression of genetic information; and similar processes of energy metabolism, molecular transport, signaling, development, and structure.

In this chapter, we introduce the common features of cells. We begin with a brief discussion of the principal small molecules and macromolecules found in biological systems. Next we discuss the fundamental aspects of cell structure and function that are

conserved in present-day organisms, focusing first on **prokaryotic organisms** — single-celled organisms without a nucleus — and their uses in studying the basic molecules of life. Then we discuss the structure and function of **eukaryotic cells** — cells with a defined nucleus — focusing on their many organelles. This discussion is followed by a section describing the use of unicellular eukaryotic organisms in investigations of basic aspects of molecular and cellular biology, focusing on yeasts and the ability to generate and analyze mutations in genes that are essential for life. We also discuss the parasite that causes malaria, focusing on the several dramatic transformations this cell undergoes within its human and mosquito hosts.

We then turn to the structure and function of metazoans — multicellular animals — focusing on cell–cell interactions that allow tissues and organs to form. We now have the complete sequences of the genomes of many metazoans, as well as that of several thousand prokaryotic organisms and single-celled eukaryotes; these sequences have provided considerable insight into the evolution of genes and organisms. We discuss how this sequence data has led to the development of many new data-driven and computer-based techniques — the fields of *genomics* and *bioinformatics* — that have greatly informed our knowledge about biology. Analysis of the complete sequences of thousands of human DNAs has led, for example, to identification of mutations in many genes that cause human diseases. Analysis of ancient DNA samples has led to an

understanding of the relationships between modern humans and extinct Neanderthals and the finding that many humans today have up to 2 percent of their DNA genome derived from Neanderthals as a result of matings thousands of years ago (Figure 1-5). Such DNA studies have also elucidated the migratory pathways taken by ancient humans in and from Africa to their present locations. Analysis of messenger RNAs present in individual human cells has led to the discovery of totally new types of cells in the human body. Continuing advances in genomics and bioinformatics, along with developments in artificial intelligence and machine learning over the coming years will certainly illuminate many other aspects of human biology and disease.

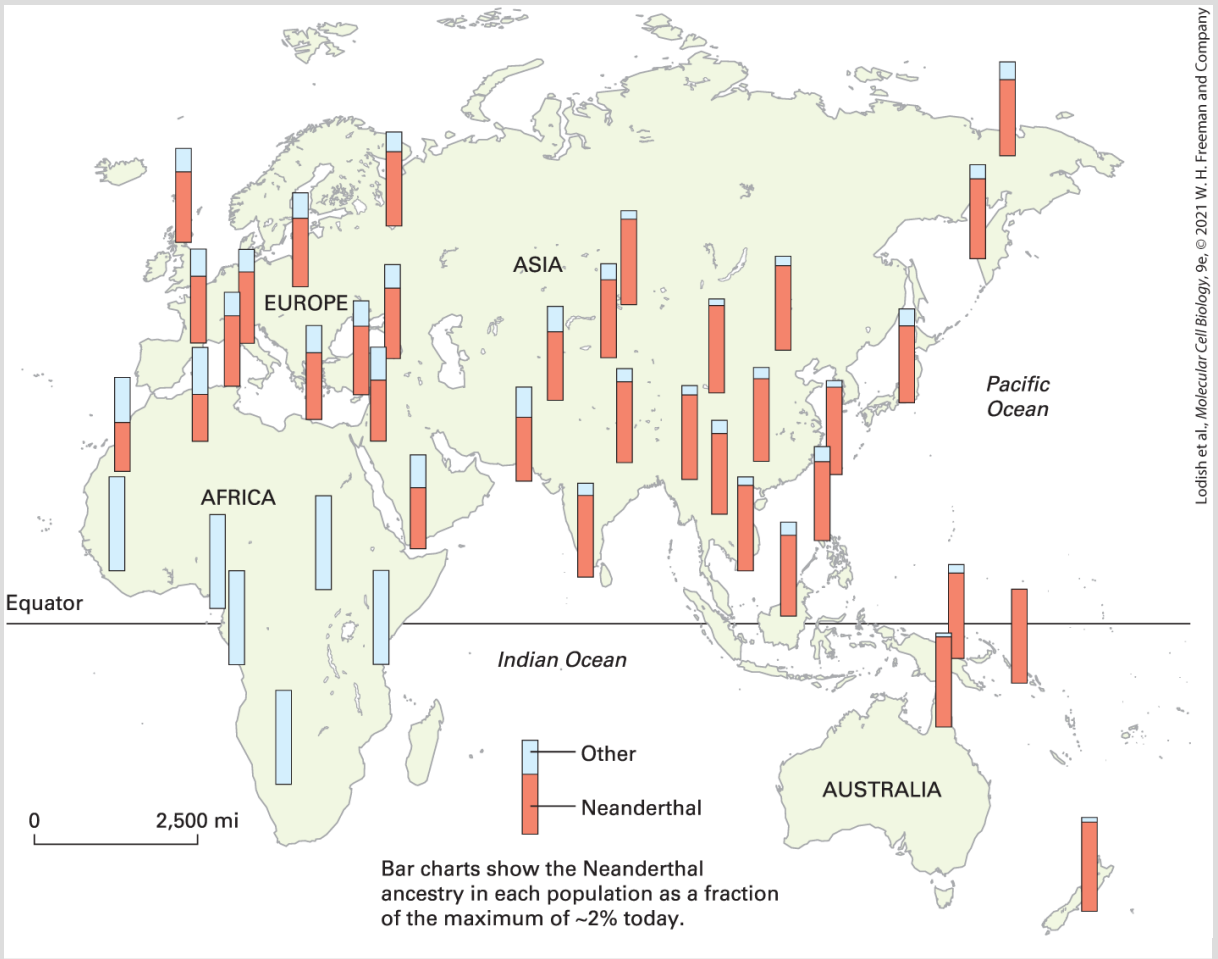


FIGURE 1-5 Approximate proportions of Neanderthal DNA found in present day humans. Humans bred with Neanderthals in the distant past, most likely in present-day Europe, and up to 2 percent of the DNA in present-day humans has a sequence derived from Neanderthal DNA. Plotted in this graph is the amount of Neanderthal DNA in humans in various locations, relative to the maximum observed of 2 percent. The entire Neanderthal genome was first sequenced from several $\sim 40{,}000$-year-old bones found in a cave in Croatia; clearly Neanderthals mated with forerunners of modern humans in Eurasia, and we see different segments of Neanderthal DNA remaining in different modern non-African people. Neanderthal DNA is nearly absent from indigenous people living in South Africa, presumably because the great majority of their ancestors never encountered Neanderthals.

[Data from: "Map" by Oliver Uberti, copyright © 2018 by David Reich & Eugenie Reich; from *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past* by David Reich, Pantheon Books.]

In the final section of this chapter we introduce several widely studied metazoans — the fruit fly *Drosophila melanogaster*, the roundworm *Caenorhabditis elegans*, planaria, and the zebrafish *Brachydanio rerio* — that have been especially useful in uncovering the conserved roles of many genes in metazoan development and function. We show how the accumulating amount of genetic data is and will continue to be analyzed to refine the evolutionary relationships among organisms as well as our understanding of human development, function, and disease. Indeed, biologists use evolution as a research tool: if a gene and its protein have been conserved in all metazoans but are not found in unicellular organisms, the protein probably has an important function in all metazoans and thus can be studied in whatever metazoan organism is most suitable for the investigation. Because the structure and function of many types of metazoan cells are also conserved, we now understand the structure and function of many cell types in considerable detail, including muscle and liver cells and the sheets of epithelial cells that line the intestine and form our skin. But many other cells — especially the multiple types that form our nervous and immune systems — still retain mysteries we need to solve; much important cell biological experimentation is needed on these and other cell systems and organs that form our bodies.

1.1 The Molecules of Life

While large polymers are the focus of molecular and cellular biology, small molecules are the stage on which all cellular processes are set. Water, inorganic ions, and a wide array of relatively small organic molecules (Figure 1-6) account for 75 to 80 percent of living matter by weight, and water accounts for about 75 percent of a cell's volume. These small molecules serve as substrates for many of the chemical reactions that take place inside the cell, including the reactions of energy metabolism and cell signaling. Cells acquire these small molecules in different ways. Ions, water, and many small organic molecules are imported into the cell (see Chapter 11); other small molecules are synthesized within the cell, often by a series of chemical reactions (see Chapter 12).

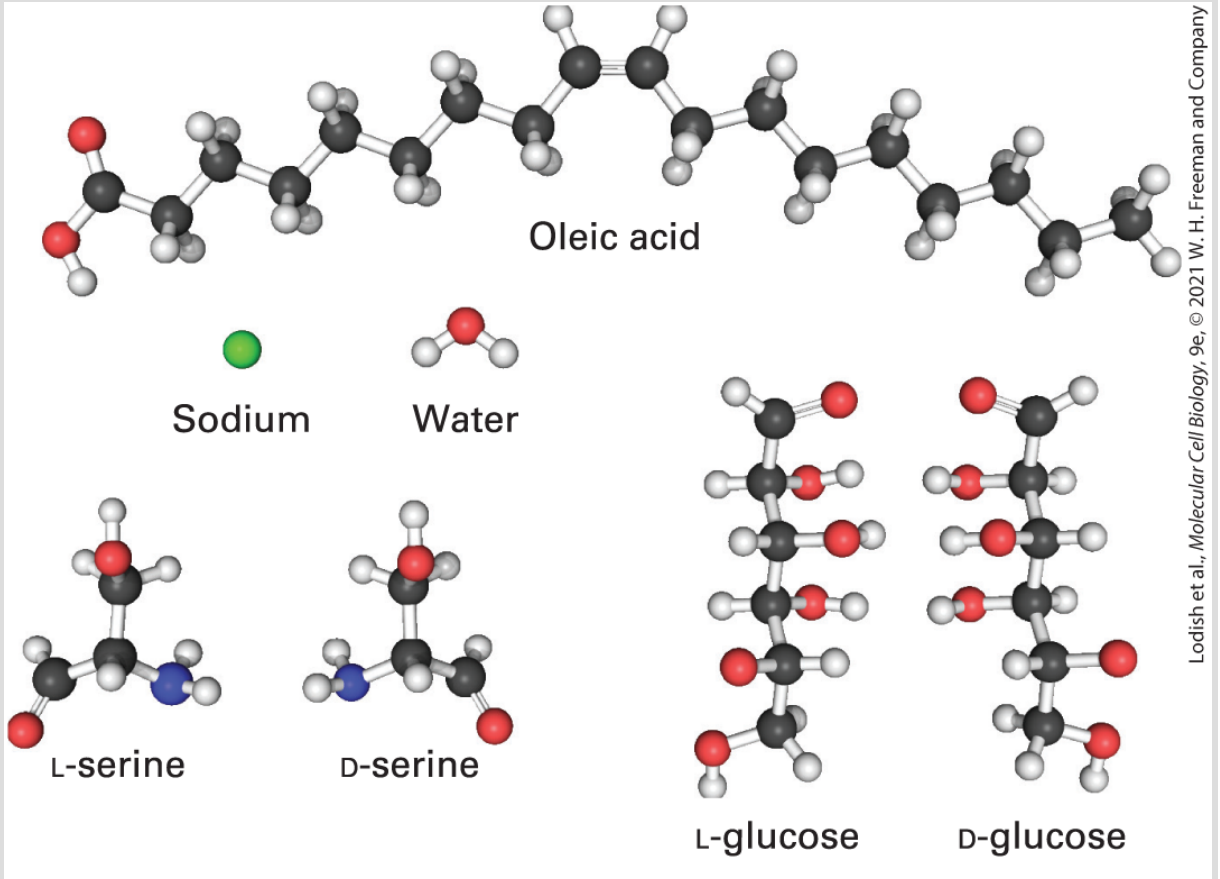


FIGURE 1-6 Some of the many small molecules found in cells. Only the L-forms of amino acids, such as serine depicted here, are found in cells and incorporated into proteins, not their D-mirror images. Only the D-form of glucose, not its L-mirror image, is found in nature and can be metabolized by cells to carbon dioxide and water.

Even in the structures of many small molecules, such as sugars, vitamins, and amino acids, we see the footprint of evolution. For example, all amino acids except glycine have an asymmetric carbon atom, yet only the L-stereoisomer, never the D-stereoisomer, is incorporated into proteins. Similarly, only the D-stereoisomer of glucose is found in organisms, never the mirror-image L-stereoisomer (see Figure 1-6). At an early stage of biological evolution, our common cellular ancestor evolved the ability to

catalyze chemical reactions that act on one stereoisomer instead of the other. How these selections happened is unknown, but now these choices are locked in place.

The small molecules in cells have a variety of functions, For example, the universally conserved small molecule **adenosine triphosphate (ATP)** stores readily available chemical energy in two of its chemical bonds (Figure 1-7). When one of these energy-rich bonds in ATP is broken, forming **ADP (adenosine diphosphate)** and inorganic phosphate, HPO_4^{2-}, often abbreviated P_i, the released energy can be harnessed to power energy-requiring processes such as muscle contraction or protein biosynthesis. To obtain energy for making ATP, many types of cells break down food molecules. For instance, when sugar is degraded to carbon dioxide and water, the energy stored in the sugar molecule's chemical bonds is released, and much of it can be captured in the energy-rich bonds in ATP (see Chapter 12). Bacterial, plant, and animal cells can all make ATP by this process. In addition, plants and a few other organisms can harvest energy from sunlight to form ATP in **photosynthesis**. Some prokaryotes that live in extreme environments produce ATP by catalyzing reactions with reduced compounds like hydrogen sulfide (H_2S).

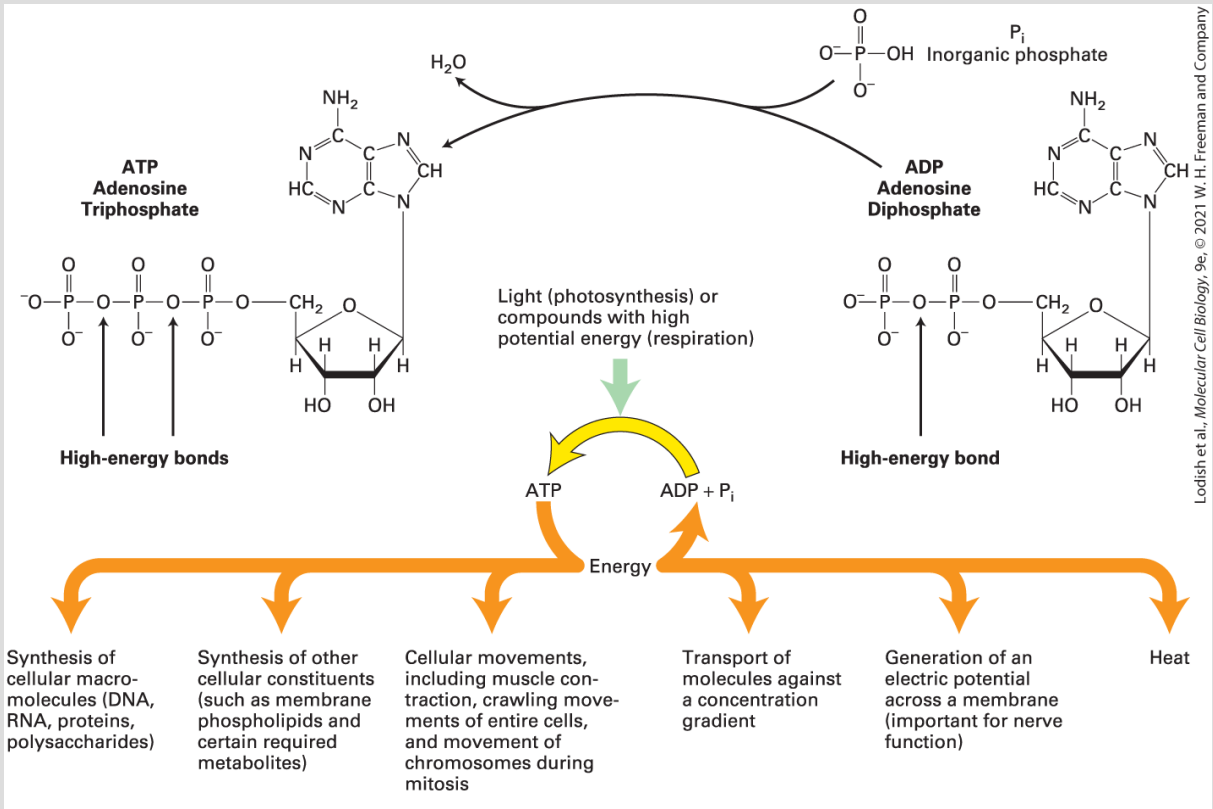


FIGURE 1-7 Adenosine triphosphate (ATP) is the most common molecule used by cells to capture, store, and transfer energy. ATP is formed from adenosine diphosphate (ADP) and inorganic phosphate (P_i) by photosynthesis in plants and photosynthetic algae and by the breakdown of sugars and fats in most cells. The energy released by the splitting (hydrolysis) of P_i from ATP drives many cellular processes.

Other small molecules (e.g., hormones such as adrenaline and histamine) act as signals that direct the activities of cells (see Chapters 15 and 16), and neurons (nerve cells) communicate with one another by releasing and sensing certain small signaling molecules such as serotonin (see Chapter 23). The powerful physiological effects of a frightening event, for example, come from the instantaneous flooding of the body with the small-molecule

hormone adrenaline, which mobilizes the fight-or-flight response (see Chapter 15).

Certain small molecules (**monomers**) can be joined to form **polymers** (also called **macromolecules**) through repetition of a single type of covalent chemical-linkage reaction. Cells produce three types of large macromolecules: polysaccharides, proteins, and nucleic acids. Sugars, for example, are the monomers used to form polysaccharides. Different polymers of D-glucose form cellulose, an important component of plant cell walls (see Chapter 20), and glycogen, a storage form of glucose found in liver and muscle (see Chapter 15). The cell is careful to provide the appropriate mix of small molecules needed as precursors for synthesis of macromolecules.

Proteins Give Cells Structure and Perform Most Cellular Tasks

Proteins, the workhorses of the cell, are the most abundant and functionally versatile of the cellular macromolecules. Cells string together 20 different **amino acids** (see Figure 2-14) in linear chains, each with a defined sequence, to form proteins, which commonly range in length from 100 to 1000 amino acids. During or just after its polymerization, a linear chain of amino acids folds into a complex shape, conferring a distinctive three-dimensional structure and

function on the protein (Figure 1-8). Humans obtain amino acids either by synthesizing them from other molecules or by breaking down proteins that we eat.

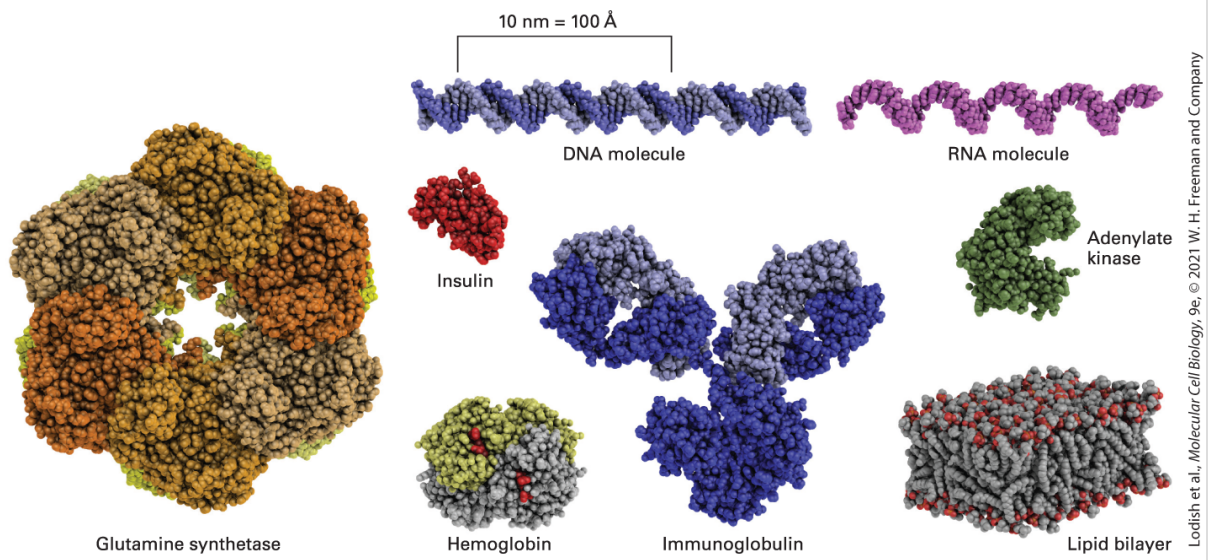


FIGURE 1-8 Models of some representative proteins drawn to a common scale and compared with a small portion of a lipid bilayer, a double-stranded DNA molecule, and a single-stranded RNA molecule. Each protein has a defined three-dimensional shape held together by numerous covalent and noncovalent chemical bonds. The illustrated proteins include enzymes (glutamine synthetase and adenylate kinase), an antibody (immunoglobulin), a hormone (insulin), and the oxygen carrier (hemoglobin) in red blood cells.

[Glutamine synthetase data from H. S. Gill and D. Eisenberg, 2001, *Biochemistry* **40**:1903–1912, PDB ID 1fpy. Insulin data from E. N. Baker et al., 1988, *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **319**:369–456, PDB ID 4ins. Hemoglobin data from G. Fermi et al., 1984, *J. Mol. Biol.* **175**:159–174, PDB ID 2hhb. Immunoglobulin data from L. J. Harris et al., 1998, *J. Mol. Biol.* **275**:861–872, PDB ID 1igy. Adenylate kinase data from G. Bunkoczi et al., PDB ID 2c9y.]

Proteins have a variety of functions in the cell. Many proteins are **enzymes**, which accelerate (catalyze) chemical reactions involving

small molecules or macromolecules (see Chapter 3). Certain proteins catalyze steps in the synthesis of all proteins; others catalyze synthesis of macromolecules such as DNA and RNA (see Chapter 5). Other proteins form long fibers that serve as structural components of a cell, for example, by forming an internal skeleton termed the *cytoskeleton*. Other proteins associated with the cytoskeleton power the movement of subcellular structures such as chromosomes, and even of whole cells, by using energy stored in the chemical bonds of ATP (see Chapters 17 and 18). Still other proteins bind adjacent cells together or form parts of the extracellular matrix (see Figure 1-4). Proteins can be sensors that change shape as temperature, ion concentrations, or other properties of the cell change. Many proteins that are embedded in the cell-surface (plasma) membrane import and export a variety of small molecules and ions (see Chapter 11). Some proteins, such as insulin, are hormones; others are hormone receptors that bind their target protein or small molecule and then generate a signal that regulates a specific aspect of cell function (see Chapters 15 and 16). Other important classes of proteins bind to specific segments of DNA, turning genes on or off (see Chapter 8). In fact, much of molecular cell biology consists of studying the function of specific proteins in specific cell types.

Nucleic Acids Carry Coded Information for Making Proteins at the Right Time

and Place

The macromolecule that garners the most public attention is deoxyribonucleic acid (DNA), whose functional properties make it the cell's master molecule. The three-dimensional structure of DNA was first proposed by James D. Watson and Francis H. C. Crick in 1953 based on the x-ray crystal analysis of DNA by Rosalind Franklin. The DNA molecule consists of two long helical strands that are coiled around a common axis to form a **double helix** (Figure 1-9). The double-helical structure of DNA, one of nature's most magnificent constructions, is critical to the phenomenon of **heredity**, the transfer of genetically determined characteristics from one generation to the next.

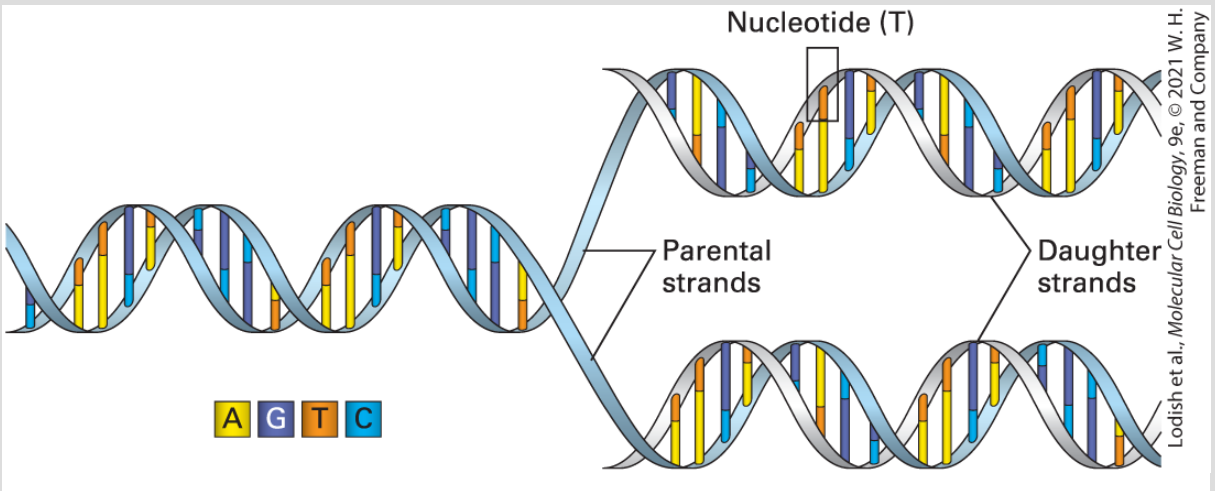


FIGURE 1-9 DNA consists of two complementary strands wound around each other to form a double helix. The double helix is stabilized by weak hydrogen bonds between the A and T bases and between the C and G bases as well as by interactions between the stacked bases. During replication, the two strands are unwound and used as templates to produce complementary strands. The outcome is two identical copies of the original double helix, each containing one of the original strands and one new daughter (complementary) strand.

DNA strands are composed of monomers called **nucleotides**; these monomers are often referred to as *bases* because they contain cyclic organic bases (see Chapter 5). Four different nucleotides, abbreviated A, T, C, and G, are joined to form a DNA strand, with the base parts projecting inward from the backbone of the strand. Two strands bind together via hydrogen bonding between the bases in the strands and twist to form a double helix. Each DNA double helix has a simple construction: wherever one strand has an A, the other strand has a T, and each C is matched with a G (see Figure 1-9). This **complementary** matching of the two strands is so strong that if complementary strands are separated under the right salt concentration and temperature conditions, they will spontaneously zip back together. This property is critical for DNA replication and inheritance, as we will learn in Chapter 5, and also underlies many of the techniques for studying DNA molecules that are detailed in Chapter 6.

The genetic information carried by DNA resides in its *sequence*, the linear order of nucleotides along a strand. Specific segments of DNA, termed *genes*, carry instructions for making specific proteins. Most bacteria have a few thousand protein-coding genes; yeasts and other unicellular eukaryotes have about 5000. Humans and other metazoans have between 13,000 and 23,000, while many plants have more (Table 1-2).

TABLE 1-2 • Genome Sizes of Organisms Used in Molecular Cell Biology Research That Have Been Completely Sequenced

	Base Pairs (Millions)	Approximate Number of Encoded Proteins[i]	Chromosomes[ii]	Reference
Eubacteria				
Mycoplasma genitalium	0.58	500	1	a
Helicobacter pylori	1.67	1,500	1	a
Haemophilus influenzae	1.83	1,600	1	a
Escherichia coli	4.64	4,100	1	a
Bacillus subtilis	4.22	4,200	1	a
Archaea				
Methanococcus jannaschii	1.74	1,800	1	a
Sulfolobus solfataricus	2.99	3,000	1	a
Single-Celled Eukaryotes				
Saccharomyces cerevisiae	12.16	6,700	16	b
Chlamydomonas reinhardtii	120.4	14,400	17	b

Plasmodium falciparum	23.26	5,400	14	b
Multicellular Eukaryotes (Metazoans)				
Drosophila melanogaster	168.74	13,900	6	b
Caenorhabditis elegans	100.29	20,500	6	b
Schmidtea mediterranea (planarian)	480	$>20,000$[iii]	4	c
Danio rerio (zebrafish)	1412.46	26,500	25	b
Gallus gallus (chicken)	1072.54	15,500	33	b
Mus musculus (mouse)	3480.96	23,100	21	b
Homo sapiens (human)	3326.74	20,800	24	b
Arabidopsis thaliana	135.67	27,400	5	b

[i]Numbers of encoded proteins are current estimates rounded to the nearest 100 based on genome DNA sequences. They will most likely change slightly in Eubacteria and Archaea because of the inclusion of newly discovered genes that code for very small proteins, and modestly in eukaryotes because of newly discovered small genes and because of pseudogenes that are not expressed.

[ii]Only nuclear chromosomes are counted in eukaryotes, including distinct sex chromosomes in metazoans.

[iii]Predicted value.

SOURCE: Table courtesy of Dr. Juan Alvarez-Dominguez. References: a, http://www.ncbi.nlm.nih.gov/genome/; b, http://ensemblgenomes.org/; c, http://www.genome.gov/12512286.

As we discuss later in this chapter, many of the genes in bacteria specify the sequences of proteins that catalyze reactions that occur universally, such as the metabolism of glucose and the synthesis of nucleic acids and proteins. These genes, and the proteins encoded by them, are conserved throughout all living organisms, and thus studies on the functions of the genes and proteins in bacterial cells have yielded profound insights into basic life processes. Similarly, many genes in unicellular eukaryotes, such as yeasts, encode proteins that are conserved throughout all eukaryotes but are not found in prokaryotes; we will see how yeasts have been used in studies of processes such as cell division that have yielded profound insights into human diseases such as cancer.

How is information stored in the DNA sequences converted into proteins? Cells use two processes in series (Figure 1-10). In the first process, called **transcription**, the protein-coding region of a gene is copied into a single-stranded **ribonucleic acid (RNA)** whose sequence is the same as one of the two strands of the double-stranded DNA. A large enzyme, **RNA polymerase**, catalyzes the

linkage of nucleotides into an RNA chain using DNA as a template. In eukaryotic cells, the initial RNA product is processed into a smaller messenger RNA molecule, which moves out of the nucleus to the **cytoplasm**, the region of the cell outside of the nucleus. Here the **ribosome**, an enormously complex molecular machine composed of both RNA and proteins, carries out the second process, called **translation**. During translation, the ribosome assembles and links together amino acids in the precise order dictated by the mRNA sequence according to the nearly universal **genetic code**. We examine the cell components that carry out transcription and translation in detail in Chapter 5.

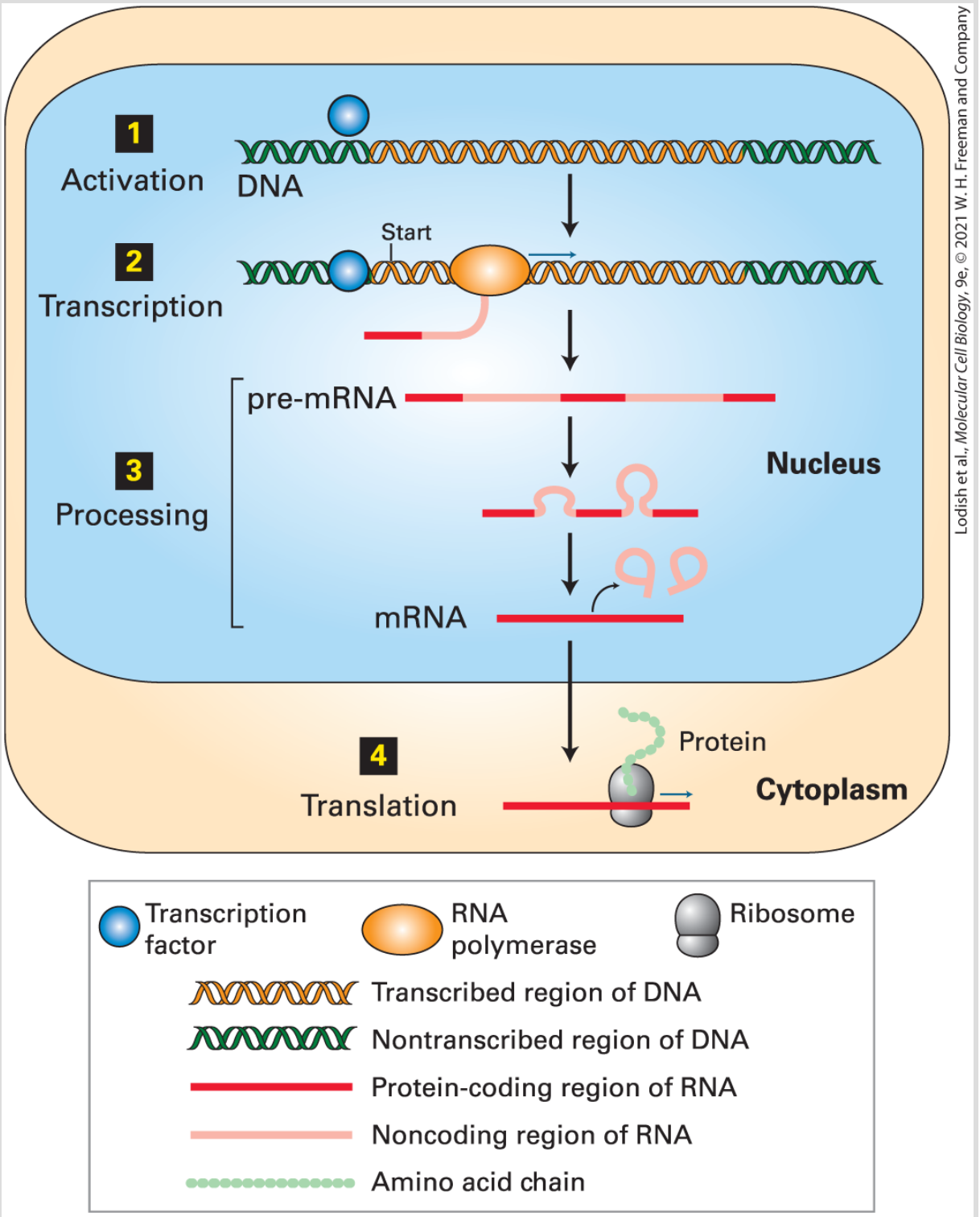


FIGURE 1-10 In eukaryotic cells the information encoded in DNA is converted into the amino acid sequences of proteins by a multistep process. Step 1 : Transcription factors and other proteins bind to the regulatory regions of the specific genes they control to

activate those genes. Step 2: RNA polymerase begins transcription of an activated gene at a specific location, the start site. The polymerase moves along the DNA, linking nucleotides into a single-stranded pre-mRNA transcript using one of the DNA strands as a template. Step 3: The transcript, whose nucleotide sequence is identical to that of the nontemplate strand, is processed to remove noncoding intron sequences. Step 4: The mature mRNA moves to the cytoplasm, where it is bound by ribosomes that read its sequence and assemble a protein by chemically linking amino acids one at a time into a linear chain.

In addition to its role in transferring information from nucleus to cytoplasm, RNA can serve as a framework for building a molecular machine. The ribosome, for example, is built of four RNA chains that bind to more than 50 proteins to make a remarkably precise and efficient mRNA reader and protein synthesizer. While most chemical reactions in cells are catalyzed by proteins, a few, such as the formation by ribosomes of the peptide bonds that connect amino acids in proteins, are catalyzed by RNA molecules.

Well before the entire human genome was sequenced, it was apparent that only about 10 percent of human DNA consists of protein-coding genes, and for many years the remaining 90 percent was considered "junk DNA"! In recent years, we've learned that much of the so-called junk DNA is actually copied into thousands of RNA molecules that, though they do not encode proteins, serve equally important purposes in the cell (see Chapter 9). At present, however, we know the function of only a very few of these abundant noncoding RNAs.

Like enzymes, certain RNA molecules, termed **ribozymes**, catalyze chemical reactions, as exemplified by the RNA inside a ribosome. Many scientists support the *RNA world* hypothesis, which proposes that RNA molecules that could replicate themselves were the precursors of current life forms; billions of years ago, the RNA world gradually evolved into the DNA, RNA, and protein world of today's organisms.

All organisms must control when and where their genes are transcribed. Nearly all the cells in our bodies contain the full set of human genes, but in each cell type only some of these genes are active, or turned on, and used to make proteins. For instance, liver cells produce some proteins that are not produced by muscle cells, and vice versa. Moreover, many cells respond to external signals or changes in external conditions by turning specific genes on or off, thereby adapting their repertoire of proteins to meet current needs. Commonly, protein-coding genes contain two parts: the coding region specifies the amino acid sequence of a protein, and the regulatory region binds specific proteins called **transcription factors**, which act as switches, either activating or repressing transcription of particular genes, as discussed in Chapter 8. By this means, the regulatory region controls when and in which cells the gene's encoded protein is made.

Phospholipids Are the Conserved Building Blocks of All Cellular Membranes

Life depends on cells being enclosed by a surface membrane that forms the interface of a cell with its environment. In eukaryotes, membranes also surround all of the internal compartments, termed *organelles*. Small molecules called *phospholipids* spontaneously pack together in such a way that they form the basic structure of these membrane sheets. Each of these bipartite molecules has a "water-loving" (hydrophilic) "head" and a "water-hating" (hydrophobic) "tail" composed of fatty acid chains. Cellular membranes are composed primarily of a bilayer (two layers) of phospholipid molecules. The two phospholipid layers of a membrane are oriented with all the hydrophilic heads directed toward the inner or outer surfaces of the membrane and the hydrophobic tails buried within its interior (Figure 1-11). Smaller amounts of other lipids, such as cholesterol, are inserted into this phospholipid framework. Cellular membranes are extremely thin relative to the size of a cell. If you magnify a bacterium or yeast cell about 10,000 times to the size of a soccer ball, the plasma membrane is about as thick as a sheet of paper!

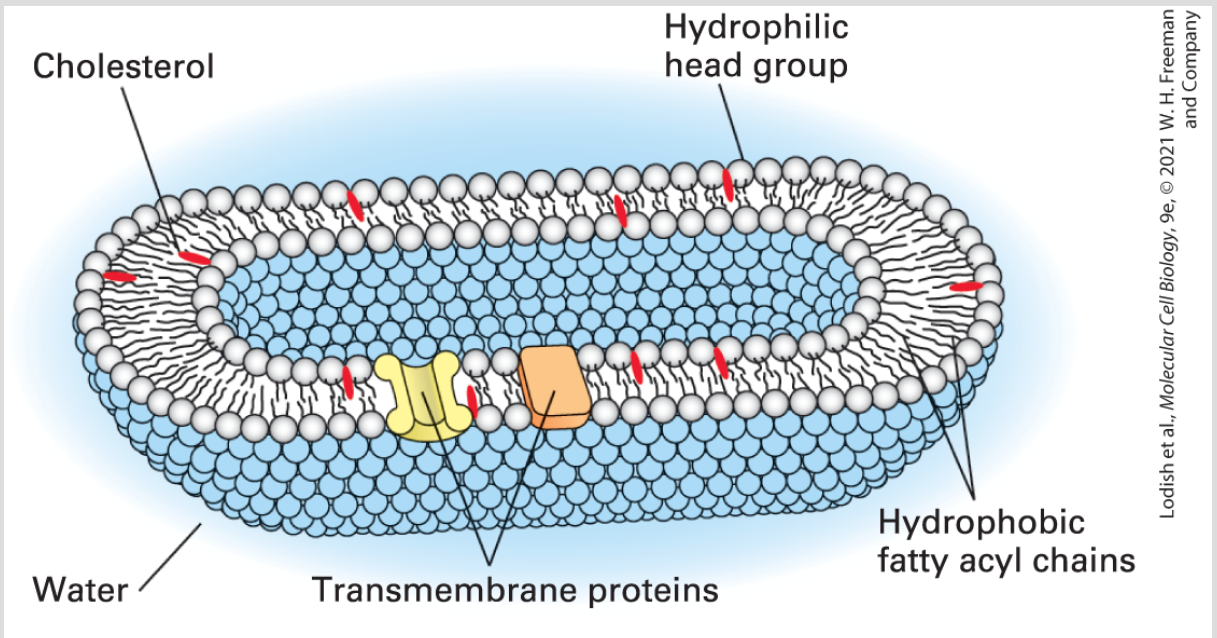


FIGURE 1-11 The watery interior of cells is surrounded by the plasma membrane, a two-layered shell of phospholipids. The phospholipid molecules are oriented with their hydrophobic fatty acyl chains (black squiggly lines) facing inward and their hydrophilic head groups (spheres) facing outward. Thus both sides of the membrane are lined by head groups, mainly charged phosphates, adjacent to the watery spaces inside and outside the cell. All biological membranes have the same basic phospholipid bilayer structure. Cholesterol (red) and various proteins are embedded in the bilayer. The interior space is vastly larger relative to the volume of the plasma membrane than is depicted here.

Phospholipid membranes are impermeable to water, all ions, and virtually all hydrophilic small molecules. Thus each membrane in each cell also contains groups of proteins that allow specific ions and small molecules to cross. Other membrane proteins serve to attach the cell to other cells or to polymers that surround it (see Figure 1-4); still others give the cell its shape or allow its shape to change. We will learn more about membranes and how molecules cross them in Chapters 10 and 11.

New cells are always derived from parental cells by cell division. We've seen that the two strands of the parental DNA are templates for the synthesis of new DNA molecules such that each daughter DNA molecule has the same sequence as the parental one. In parallel, new membranes are made by incorporation of phospholipids and proteins into existing membranes in the parental cell and divided between daughter cells by fission. Thus membrane synthesis, like DNA synthesis, is templated by a parental structure.

Quality Control of All Cellular Macromolecules Is Essential for Life

All cells are constantly bombarded by potentially hazardous chemicals and radiation that can damage the macromolecules within, and cells expend a significant amount of their energy to ensure that damaged molecules — especially DNA — are repaired or degraded. Proteins function only when they are folded into their proper three-dimensional structure (see Figure 1-8). Proteins that fail to fold, or become unfolded as a result of stresses such as an elevated temperature, are usually rapidly degraded by enzymes within the cell, releasing their constituent amino acids, which are reused for making new proteins. Occasionally, though, misfolded proteins aggregate together inside the cell, causing damage to the cell. As discussed in Chapter 3, such protein aggregates that form

over time in brain nerve cells are thought to be the cause of Alzheimer's and other neurodegenerative diseases.

Ultraviolet and gamma radiation, as well as many chemicals, can damage DNA, causing changes in the normal DNA sequence. Such mutations can cause errors in DNA replication and lead to inability of that sequence to direct the synthesis of the correct protein. Cells have several mechanisms, discussed in Chapter 5, to repair such damage to DNA, converting the mutation back to the normal sequence. Occasionally mutations are not repaired and cause death of the cell. Alternatively, unrepaired mutations can cause proteins to be made in the wrong cell or at the wrong time and lead to uncontrolled growth of the cell. Chapter 25 details how many types of mutations that occur over an individual's life cause cancer.

1.2 Prokaryotic Cell Structure and Function

The biological universe consists of two types of cells: prokaryotic and eukaryotic. Prokaryotic cells such as bacteria consist of a single closed compartment that is surrounded by a plasma membrane, lack a defined nucleus, and have a relatively simple internal organization (Figure 1-12). Eukaryotic cells contain a defined membrane-bounded nucleus and extensive internal membranes that enclose the organelles (Figure 1-13).

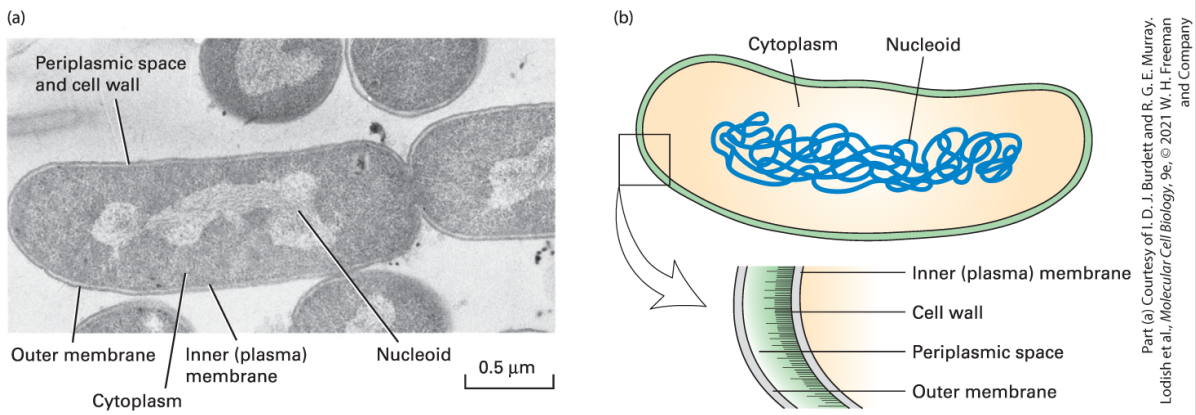
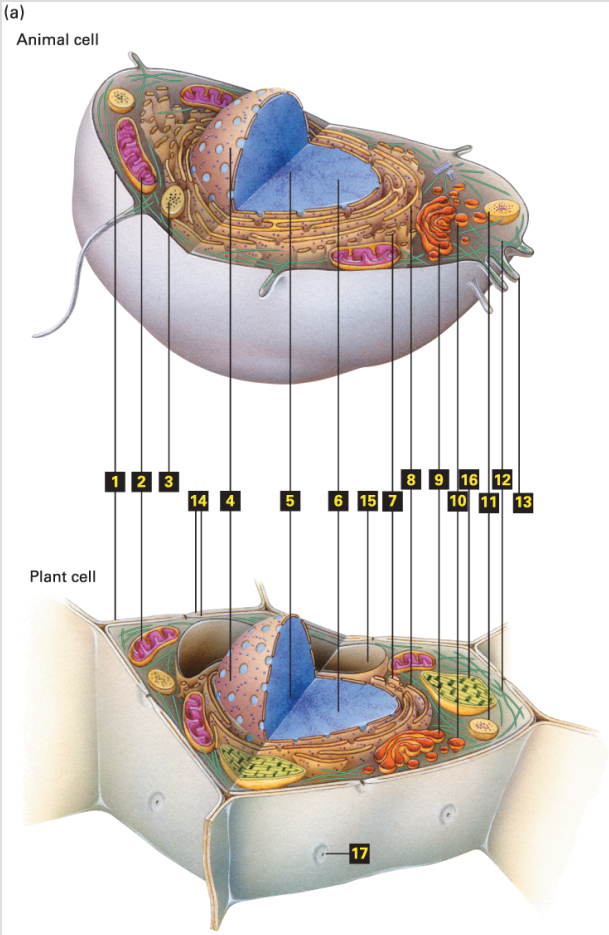


FIGURE 1-12 Prokaryotic cells have a relatively simple structure. (a) Electron micrograph of a thin section of *Escherichia coli*, a common intestinal bacterium. The nucleoid, consisting of the bacterial DNA, is not enclosed within a membrane. *E. coli* and other gram-negative bacteria are surrounded by two membranes separated by the periplasmic space. The thin cell wall is adjacent to the inner membrane. (b) This artist's drawing shows the nucleoid (blue) and a magnification of the layers that surround the cytoplasm in a gram-negative bacterium. Most of the cell is composed of water, proteins, ions, and other molecules that are too small to be depicted at the scale of this drawing.



1. Plasma membrane controls movement of molecules in and out of the cell and functions in cell-cell signaling and cell adhesion.
2. Mitochondria, which are surrounded by a double membrane, generate ATP by oxidation of glucose and fatty acids.
3. Lysosomes, which have an acidic lumen, degrade material internalized by the cell and worn-out cellular membranes and organelles.
4. Nuclear envelope, a double membrane, encloses the contents of the nucleus; the outer nuclear membrane is continuous with the rough ER.
5. Nucleolus is a nuclear subcompartment where most of the cell´s rRNA is synthesized.
6. Nucleus is filled with chromatin composed of DNA and proteins; site of mRNA and tRNA synthesis.
7. Smooth endoplasmic reticulum (ER) contains enzymes that synthesize lipids and detoxify certain hydrophobic molecules.
8. Rough endoplasmic reticulum (ER) functions in the synthesis, processing, and sorting of secreted proteins, lysosomal proteins, and certain membrane proteins.
9. Golgi complex processes and sorts secreted proteins, lysosomal proteins, and membrane proteins synthesized on the rough ER.
10. Secretory vesicles store secreted proteins and fuse with the plasma membrane to release their contents.
11. Peroxisomes contain enzymes that break down fatty acids into smaller molecules used for biosynthesis and also detoxify certain molecules.
12. Cytoskeletal fibers form networks and bundles that support cellular membranes, help organize organelles, and participate in cell movement.
13. Microvilli increase surface area for absorption of nutrients from surrounding medium.
14. Cell wall, composed largely of cellulose, helps maintain the cell´s shape and provides protection against mechanical stress.
15. Vacuole stores water, ions, and nutrients, degrades macromolecules, and functions in cell elongation during growth.
16. Chloroplasts, which carry out photosynthesis, are surrounded by a double membrane and contain a network of internal membrane-bounded sacs.
17. Plasmodesmata are tubelike cell junctions that span the cell wall and connect the cytoplasms of adjacent plant cells.

Part (b) Courtesy of I. D. J. Burdett and R. G. E. Murray.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

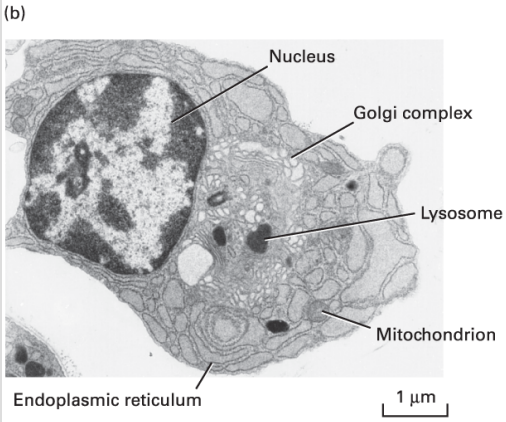


FIGURE 1-13 Subcellular organization of eukaryotic cells. (a) Schematic overview of a typical animal cell (*top*) and plant cell (*bottom*) and their major substructures. Not every cell type will contain all the organelles, granules, and fibrous structures shown here, and other substructures can be present in some cell types. Cells also differ considerably in shape and

in the prominence of various organelles and substructures. (b) Electron micrograph of a plasma cell, a type of white blood cell that secretes antibodies, showing some of the larger organelles. Not all of the organelles are labeled.

Prokaryotes Comprise Two Kingdoms: Archaea and Eubacteria

In recent years, detailed analysis of DNA sequences from a variety of prokaryotic organisms has revealed two distinct kingdoms: the Eubacteria, often simply called "bacteria," and the *Archaea* (see Figure 1-1). Eubacteria are single-celled organisms; they include the photosynthetic cyanobacteria, or "blue-green algae." Figure 1-12 illustrates the general structure of a typical eubacterial cell; archaeal cells have a similar structure. Bacterial cells are commonly 1–2 μm in size and consist of a single closed compartment containing the cytoplasm and bounded by the plasma membrane. The genome is composed of a single circular, double-stranded DNA molecule; many prokaryotes contain additional small circular DNA molecules called *plasmids*. Although bacterial cells do not have a defined nucleus, the DNA is extensively folded into the central region of the cell, called the *nucleoid*. Many proteins are precisely localized within the cytosol or nucleoid or in the plasma membrane or cell wall, indicating the presence of an elaborate structural and functional organization.

Most ribosomes are found in the cytoplasm. Unlike mRNAs in eukaryotes (see Figure 1-10), bacterial mRNAs undergo limited if any processing. Because there is no membrane barrier between bacterial DNA and cytoplasm, ribosomes are able to bind to an mRNA as soon as part of it has been synthesized by RNA polymerase; thus in prokaryotes, transcription and translation occur contemporaneously.

Bacterial cells possess a cell wall, which lies adjacent to the external side of the plasma membrane. The cell wall is composed of layers of peptidoglycan, a complex of proteins and oligosaccharides; it helps protect the cell and maintain its shape. Some bacteria (e.g., *E. coli*) have a thin inner cell wall and an outer membrane separated from the inner cell wall by the periplasmic space. Such bacteria are not stained by the Gram technique and thus are classified as gram-negative. Other bacteria (e.g., *Bacillus polymyxa*) that have a thicker cell wall and no outer membrane take the Gram stain and thus are classified as gram-positive.

In addition to DNA sequence distinctions that separate them from Eubacteria, Archaea have cell membranes that differ dramatically in composition from those of Eubacteria and eukaryotes. Many archaeans grow in unusual, often extreme, environments that may resemble the ancient conditions that existed when life first appeared on Earth. For instance, halophiles ("salt lovers") require high concentrations of salt to survive, and thermoacidophiles ("heat and

acid lovers") grow in hot (80 °C) sulfur springs, where a pH of less than 2 is common.

Many Bacteria Including *Escherichia coli* Are Widely Used in Biological Research

The bacterial lineage includes *Escherichia coli*, a favorite experimental organism, which in nature is common in soil and in animal intestines. *E. coli* and several other bacteria have a number of advantages as experimental organisms. They grow rapidly in a simple and inexpensive medium containing glucose and salts, in which they can synthesize all necessary amino acids, lipids, vitamins, and other essential small molecules. Like all bacteria, *E. coli* possesses elegant mechanisms for controlling gene activity that are now well understood. Over time, researchers have developed powerful systems for genetic analysis of this organism. These systems are facilitated by the small size of bacterial genomes, the ease of obtaining mutants, the availability of techniques for transferring genes into bacteria, an enormous wealth of knowledge about bacterial gene control and protein functions, and the relative simplicity of mapping genes relative to one another in the bacterial genome. In Chapter 6, we see that *E. coli* is widely used today in multiple types of recombinant DNA research.

Bacteria such as *E. coli* that grow in environments as diverse as the soil and the human gut have about 4000 genes, encoding about the same number of proteins (see Table 1-2). Certain disease-causing bacteria that only grow inside human cells, such as *Chlamydia,* have only ~1000 genes; they acquire amino acids and other nutrients from their host cells and thus have evolved to lack genes for enzymes that catalyze reactions in the synthesis of amino acids and certain lipids. Other parasitic bacteria such as *Mycoplasma* species have even fewer genes.

Many bacterial genes encoding proteins essential for DNA, RNA, and protein synthesis, and for membrane function are conserved in all organisms, and much of our knowledge of these important cellular processes was uncovered first by studies in *E. coli* and other bacteria. For example, certain *E. coli* membrane proteins that import amino acids across the plasma membrane are closely related in sequence, structure, and function to membrane proteins in mammalian brain cells that import small nerve-to-nerve signaling molecules called *neurotransmitters* (see Chapters 11 and 23).

Throughout this chapter, we will encounter eukaryotic organisms that have been chosen for study because, like *E. coli*, they are easy to grow and study. Of course, many bacteria cause serious diseases. Research on disease-causing bacteria is often focused on understanding their unique biology and on discovering antibiotics

that selectively kill them but not their human or animal hosts — for instance, drugs that block formation of parts of the cell wall.

1.3 Eukaryotic Cell Structure and Function

Eukaryotes comprise all members of the plant and animal kingdoms as well as protozoans (*proto*, "primitive"; *zoan*, "animal"), which are exclusively unicellular and include fungi and amoebae. Eukaryotic cells are commonly about 10–100 μm across, generally much larger than bacteria. A typical human fibroblast, a connective tissue cell, is about 15 μm across, with a volume and dry weight some thousands of times those of an *E. coli* cell. An amoeba, a single-celled protozoan, can have a cell diameter of approximately 0.5 mm, more than 30 times that of a fibroblast.

Eukaryotic cells, like prokaryotic cells, are surrounded by a plasma membrane. However, unlike prokaryotic cells, all eukaryotic cells (the human red blood cell is an exception) also contain extensive internal membranes that enclose specific subcellular compartments, the **organelles** (see Figure 1-13). The **cytosol**, the organelle-free part of the cytoplasm, contains water, dissolved ions, small molecules, and many types of proteins. Plant cells and most fungal cells are surrounded by a cell wall that gives the cell a rigid shape and also allows for rapid cell expansion.

All eukaryotic cells have many of the same organelles and other subcellular structures, attesting to their common evolutionary origin. Many organelles are surrounded by a single phospholipid

membrane, but the nucleus, mitochondrion, and chloroplast are enclosed by two membranes. Each organelle membrane and each space in the interior of an organelle has a unique and specific set of proteins that enable each type of organelle to have a distinct physiological function — for example, production of ATP from ADP and P_i occurs in mitochondria and chloroplasts; and many lipids, membrane proteins, and proteins to be secreted from cells are synthesized in the rough endoplasmic reticulum. Many outworn or misfolded proteins are degraded in lysosomes and their constituent amino acids recycled. Much like a factory building automobiles or computers, such compartmentalization of functions allows the cell as a unit to operate efficiently.

The membranes defining these subcellular compartments contain proteins that control their internal ionic composition so that the internal composition of each compartment generally differs from that of the surrounding cytosol as well as compartments in the other organelles. Here we describe the organelles common to all eukaryotic cells as well as several that are found only in certain types of eukaryotes. We begin with the proteins that give eukaryotic cells their shapes and organize the organelles.

The Cytoskeleton Has Many Important Functions

The cytoplasm contains an array of fibrous proteins collectively called the **cytoskeleton** (see Chapters 17 and 18). Three classes of fibers compose the cytoskeleton: **microtubules** (20 nm in diameter), built of polymers of the protein tubulin; **microfilaments** (7 nm in diameter), built of the protein actin; and **intermediate filaments** (10 nm in diameter). All of these fibers are long chains of multiple copies of one or more small protein subunits (Figure 1-14).

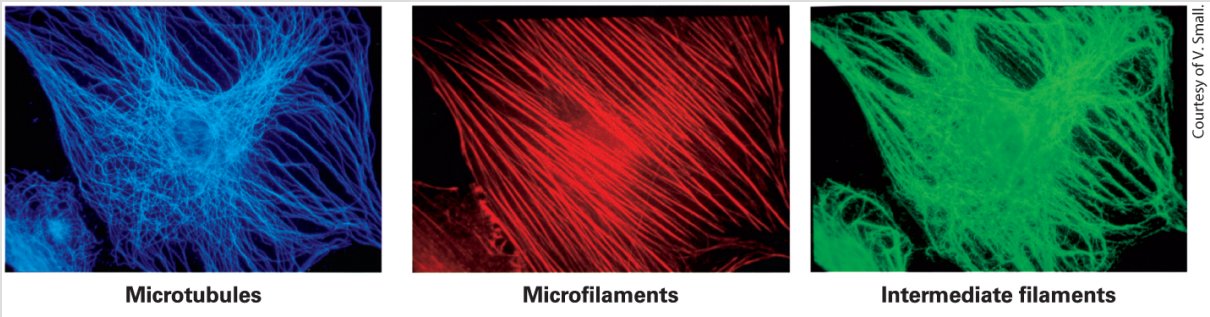


FIGURE 1-14 The three types of cytoskeletal filaments have characteristic distributions within mammalian cells. Three views of the same cell. A cultured fibroblast was treated with a detergent to make it permeable to large molecules and then treated with three different antibody preparations. Each antibody binds specifically to the protein monomers forming one type of filament and is chemically linked to a differently colored fluorescent dye (green, blue, or red). Visualization of the stained cell in a fluorescence microscope reveals the locations of filaments bound to a particular dye-antibody preparation. In this case, microtubules are stained blue; actin microfilaments, red; and intermediate filaments, green. All three fiber systems contribute to the shape and movements of cells.

The cytoskeleton has many crucial functions. It gives the cell strength and rigidity, thereby helping to maintain its shape; this is perhaps most obvious with neurons, in which microtubules and other fibers allow the formation of the long, slim protuberances — the axons and dendrites (see Figure 1-3e and Chapter 23) — that

emanate from the cell body and allow each neuron to carry out its specialized functions. Certain cytoskeletal fibers provide the tracks along which many organelles move, thus ensuring that they are localized to the correct place in the cytosol. Other fibers play key roles in cell motility and, in muscle cells, the ability to contract. Perhaps most important, cell division and the segregation of chromosomes and organelles into the two daughter cells could not occur without the organizational framework provided by the microtubule cytoskeleton and its associated proteins (Chapter 19).

Cilia and **flagella** are similar fingerlike extensions of the cell surface. Surrounded by the plasma membrane, they contain a bundle of microtubules that gives them their shape and, acting together with motor proteins, allows them to beat rhythmically. They propel materials across epithelial surfaces (Figure 1-15), enable sperm to swim, and push eggs through the oviduct (see Chapter 18). As detailed in Chapter 16, most vertebrate cells contain at least one cilium that plays a key role in cell-cell signaling.

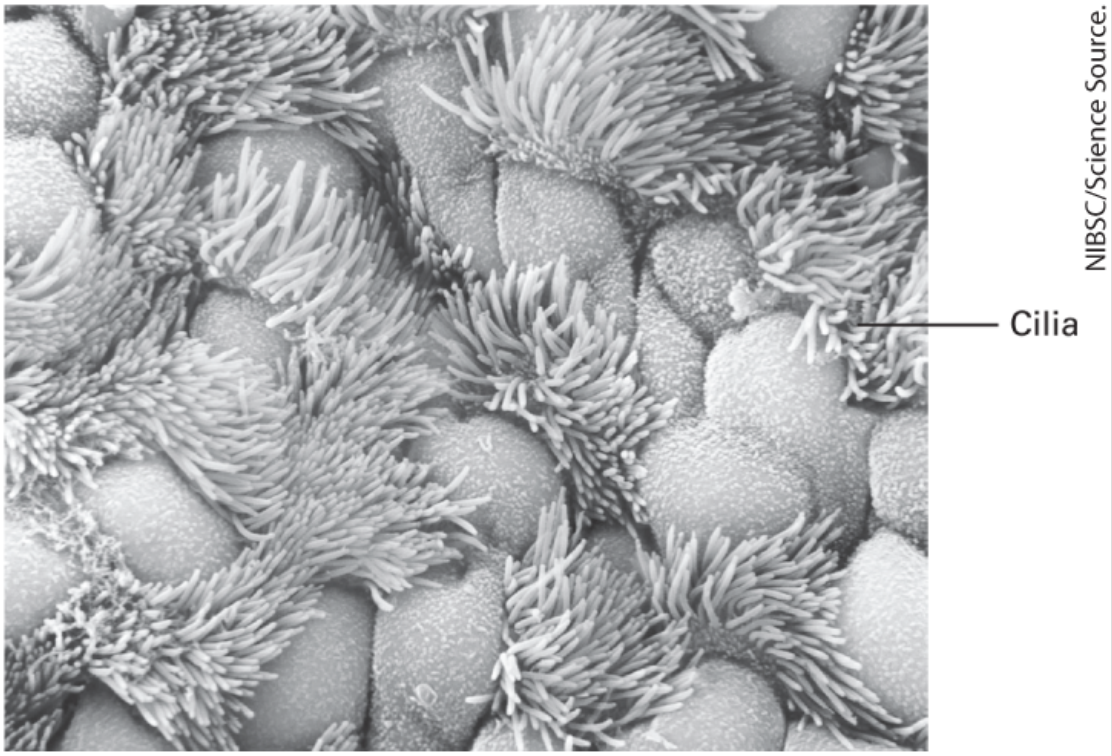


FIGURE 1-15 Surface of the ciliated epithelium lining a mammalian trachea viewed in a scanning electron microscope. Beating cilia, which have a core of microtubules, propel mucus and foreign particles out of the respiratory tract, keeping the lungs and airways clear.

The Nucleus Contains the DNA Genome, Apparatuses for Synthesis of DNA and RNA, and a Fibrous Matrix

In a growing or differentiating cell, the **nucleus**, the largest organelle in animal cells, is the site of DNA replication and the synthesis of ribosomal RNA, mRNA, and a large variety of noncoding RNAs (see Chapters 5 and 8). The nucleus is surrounded

by two membranes, each one a phospholipid bilayer containing many different types of proteins (Figure 1-16). The *inner nuclear membrane* defines the nucleus itself. In most cells, the *outer nuclear membrane* is continuous with the endoplasmic reticulum, and the space between the inner and outer nuclear membranes is continuous with the lumen — the central cavity — of the endoplasmic reticulum (Figure 1-16a). The two nuclear membranes appear to come together at **nuclear pore complexes**, ringlike structures composed of specific membrane proteins embedded in the inner and outer membranes. RNAs and proteins move between the nucleus and the cytosol through these nuclear pores (red arrow in Figure 1-16a). The structure of the nuclear pores and the regulated transport of material through them are detailed in Chapters 9 and 13. Intermediate-filament proteins called **lamins** form a two-dimensional network, called the **nuclear lamina**, that adheres to the inner surface of the inner membrane, giving it shape and rigidity. The lamina breaks down early in cell division and is reassembled after the two daughter cells have separated, as we detail in Chapter 19.

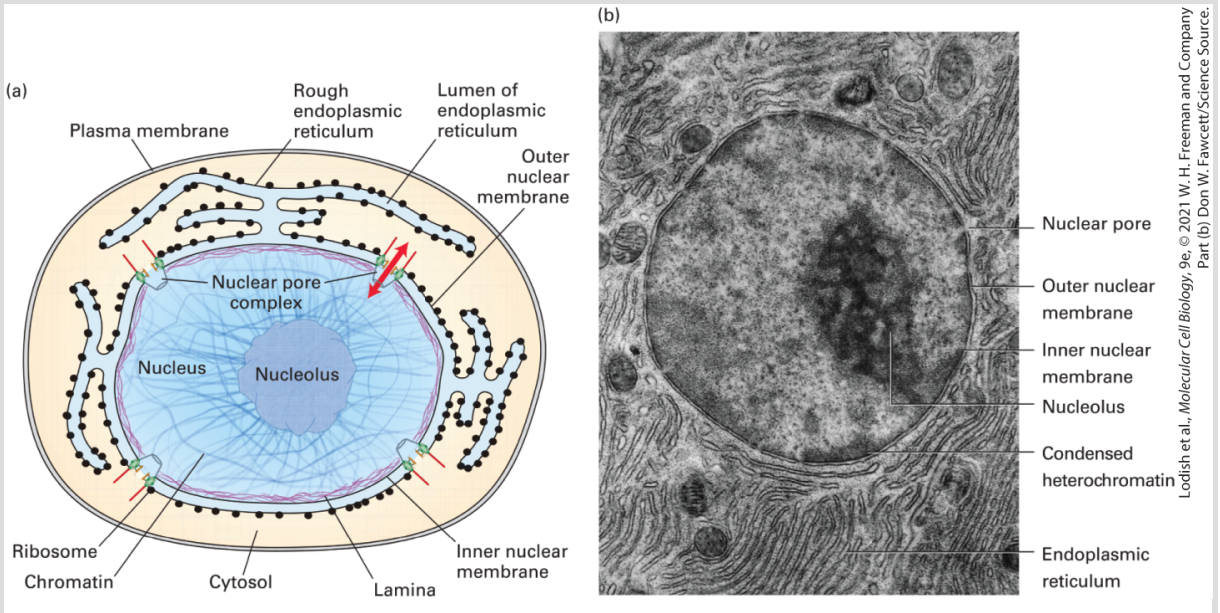


FIGURE 1-16 Structure of the nucleus. (a) Schematic diagram of the structure of a typical cell nucleus and the connection of the outer nuclear membrane with the rough endoplasmic reticulum. The small black dots attached to the membrane of the rough endoplasmic reticulum are ribosomes that are synthesizing membrane and secreted proteins. The red arrow depicts the bidirectional movement of molecules between the nucleus and cytosol through the nuclear pores. For simplicity, many of the organelles in the cytoplasm (see Figure 1-13) are not shown. (b) Electron micrograph of a pancreatic acinar cell from the bat *Myotis lucifugus*. The nucleolus is a subcompartment of the nucleus and is not surrounded by a membrane; ribosomal RNA is produced in the nucleolus and ribosomes are assembled there from ribosomal RNA and ribosomal proteins. Darkly staining areas in the nucleus outside the nucleolus are regions of heterochromatin, regions of chromatin that are compact and unable to be transcribed into messenger RNA.

The total DNA in an organism is referred to as its **genome**. In prokaryotic cells, most or all of the genetic information resides in a single circular DNA molecule about a millimeter in length; this molecule lies, folded back on itself many times, in the central region of the micrometer-sized cell (see Figure 1-12 and Table 1-2). In contrast, DNA in the nuclei of eukaryotic cells is distributed among

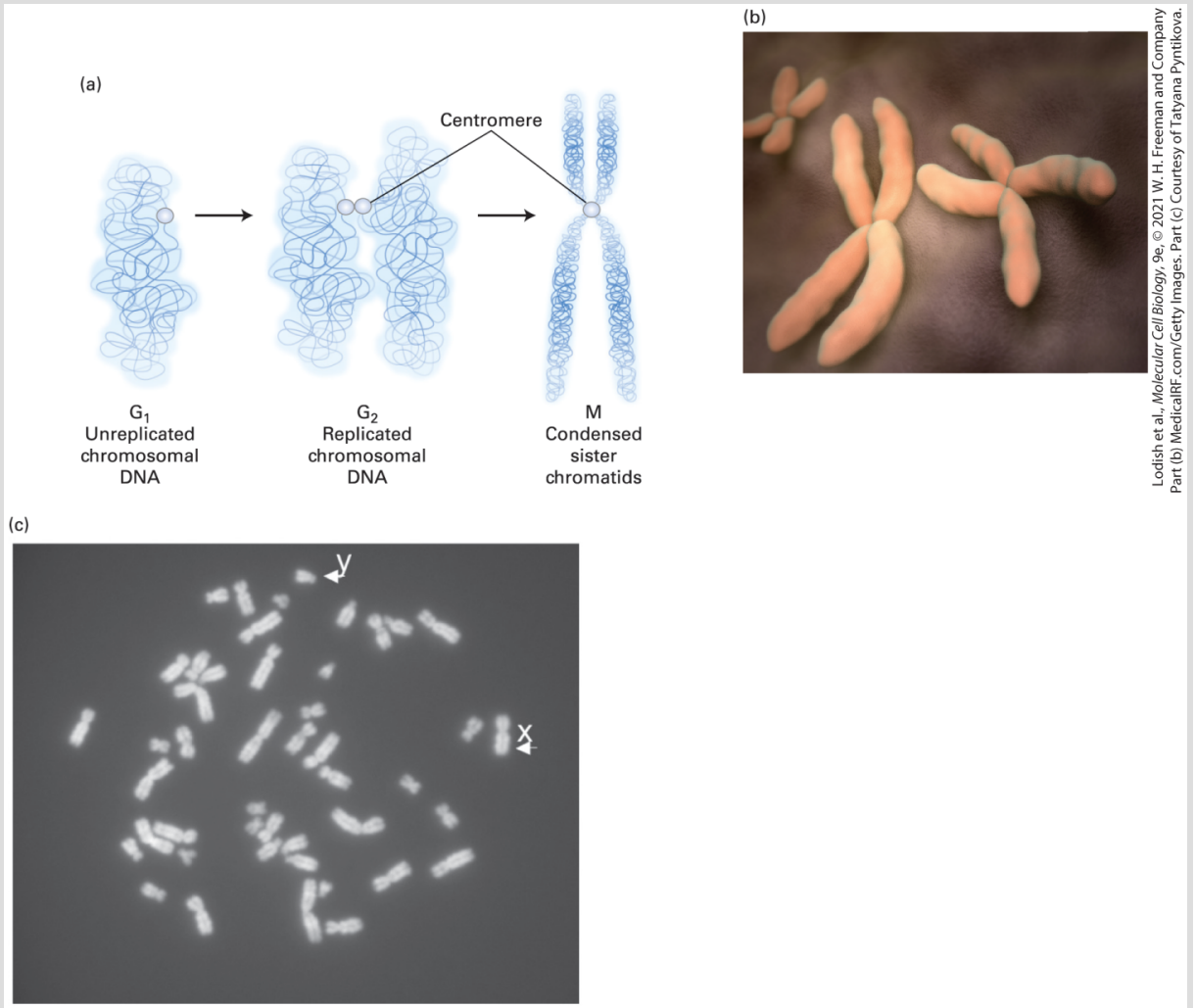
multiple long linear structures called **chromosomes**. The length and number of chromosomes are the same in all cells of a particular species, but vary among different species (see Table 1-2). As detailed in Chapter 7, each chromosome comprises a single DNA molecule associated with numerous histones and other proteins. In a nucleus that is not dividing, the chromosomes are dispersed and are not dense enough to be observed in the light microscope.

Typical nuclei contain several subcompartments, not bound by membranes, that have specific functions. The largest, the *nucleolus*, gathers together the parts of several different chromosomes containing the genes that direct the synthesis of ribosomal RNA and is the site of assembly of the ribosomal RNAs and ribosomal proteins into ribosomes.

When nondividing cells are visualized in an electron microscope, the non-nucleolar regions of the nucleus, called the *nucleoplasm*, can be seen to have dark- and light-staining areas. The dark areas, which are often closely associated with the nuclear membrane, contain condensed, concentrated DNA that cannot be transcribed into RNA, called **heterochromatin** (see Figure 1-16b).

Chromosomes, which stain intensely with basic dyes, are visible in light and electron microscopes only during cell division, when the DNA becomes tightly compacted (Figure 1-17). Although the large genomic DNA molecule in prokaryotes is also associated with

proteins, the arrangement of DNA within a bacterial chromosome differs greatly from that within the linear chromosomes of eukaryotic cells; bacterial chromosomes are circular and are associated with different types of proteins than are eukaryotic chromosomes.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Part (b) MedicalRF.com/Getty Images. Part (c) Courtesy of Tatyana Pyntikova.

FIGURE 1-17 Individual chromosomes can be seen in cells during cell division (mitosis). (a) During the S phase of the cell cycle (see Figure 1-22) the chromosomal DNA together with its attached proteins (not shown) is duplicated, and the daughter *sister chromatids*, each with a complete copy of the chromosomal DNA, remain attached at the centromere. (b) During the actual cell division process (mitosis), the chromosomal DNA becomes highly compacted,

and the pairs of sister chromatids can be seen in the electron microscope, as depicted here. (c) Light-microscope image of a chromosomal spread from a cultured human male lymphoid cell arrested in the metaphase stage of mitosis by treatment with the microtubule-depolymerizing drug colcemid. There is a single copy of the duplicated X and Y chromosomes and two copies of each of the others.

The Endoplasmic Reticulum Is the Site of Synthesis of Most Membrane and Secreted Proteins as Well as Many Lipids

We noted earlier that each type of organelle membrane, and the interior space of each organelle, contains a discrete set of proteins that enable it to carry out its specific function. Indeed, about a third of the proteins in a typical eukaryotic cell are found in the plasma membrane, in a specific organelle membrane, or in the interior space of an organelle. Thus there are a large number of proteins (and lipids) that have to be synthesized and sorted to different locations — about one-third of the proteome, the totality of proteins made by a eukaryotic cell. The synthesis of lipids, secreted proteins, and many types of membrane proteins occurs on the organelle with the largest membrane surface in a eukaryotic cell, the **endoplasmic reticulum (ER)**.

The ER is an extensive network of closed, flattened, membrane-bounded sacs called **cisternae** (Figure 1-18; see also Figure 1-16a),

and is generally subdivided into regions termed the **smooth** and **rough endoplasmic reticulum**. The smooth ER is smooth because it lacks ribosomes; it is the site of synthesis of fatty acids and phospholipids. In contrast, the cytosolic side of the rough endoplasmic reticulum is studded with ribosomes; these ribosomes synthesize certain membrane and organelle proteins and virtually all proteins that are to be secreted from the cell (see Chapter 13). Newly made membrane proteins become associated with the rough ER membrane, and proteins to be secreted accumulate in the **lumen**, the aqueous interior of the organelle.

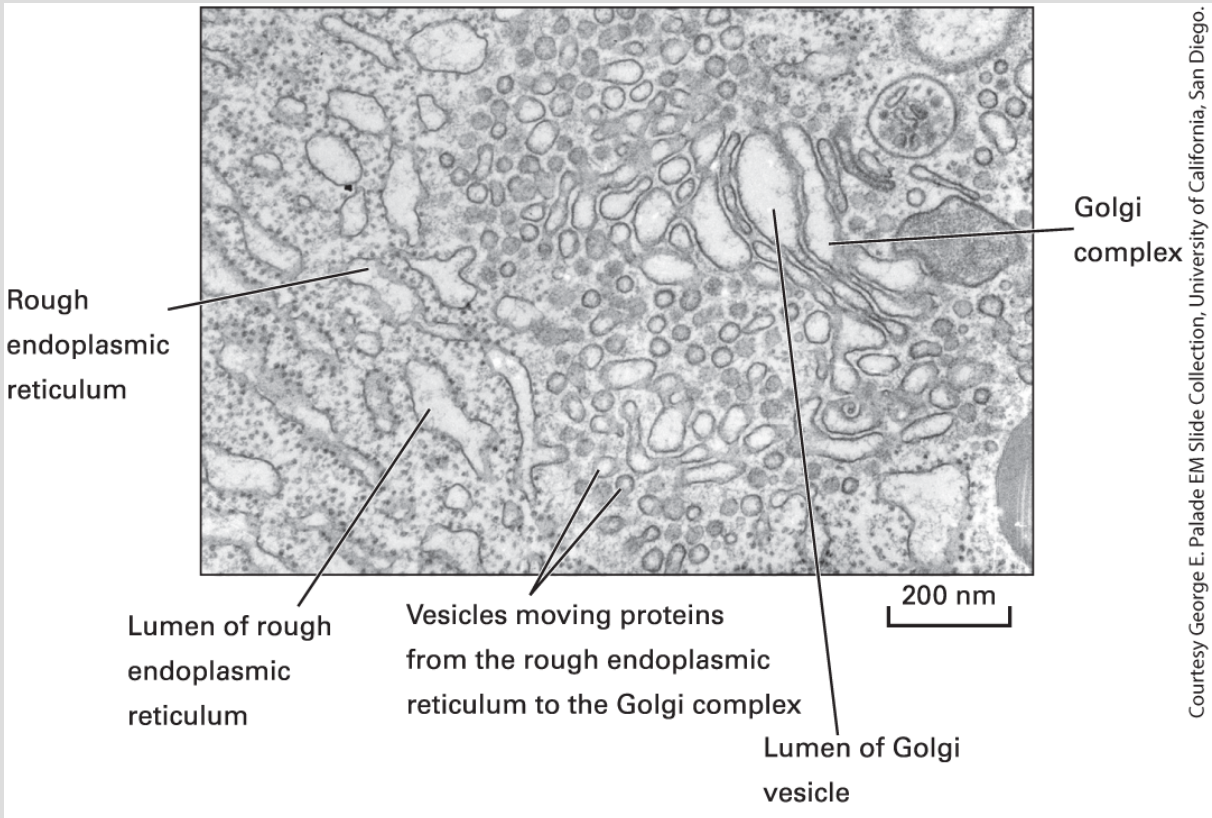


FIGURE 1-18 The Golgi complex and rough endoplasmic reticulum. An electron micrograph of a section of a human liver cell shows the abundant ribosome-studded rough endoplasmic reticulum and the Golgi complex, as well as many ribosomes free in the cytosol.

The Golgi Complex Sorts Secreted Proteins and Many Membrane Proteins to Their Final Destinations in the Cell

Several minutes after secreted proteins and membrane proteins are synthesized in the rough ER, most of them leave the organelle within small, membrane-bounded transport vesicles about 50 nanometers (0.05 μm) in diameter. These vesicles, which bud from regions of the rough ER not coated with ribosomes, carry the proteins to another membrane-bounded organelle, the Golgi complex (see Figure 1-18). As detailed in Chapter 14, secreted proteins and membrane proteins undergo a series of enzyme-catalyzed chemical modifications in the Golgi complex that are essential for these proteins to function normally.

After these proteins are modified in the Golgi complex, they are transported out of the complex by a second set of small, ~0.05-μm diameter vesicles, which bud from one side of the Golgi complex. These vesicles carry proteins to the plasma membrane or to lysosomes or to other organelles, or they release proteins from the cell into the extracellular space. How intracellular transport vesicles are formed, how they target specific proteins to specific membranes, and how they fuse with these membranes to deliver their contents is also discussed in Chapter 14.

Endosomes Bring Proteins and Particles from the Outside into Cells

Although transport proteins in the plasma membrane mediate the movement of ions and small molecules into the cell across the phospholipid bilayer, proteins, particles such as viruses, and some other soluble macromolecules in the extracellular milieu are brought into the cell by **endocytosis**. In this process, a segment of the plasma membrane buds inward — a process termed *invagination* — into a *coated pit*, whose cytosolic face is lined by a specific set of proteins that cause these vesicles to form. The pit pinches from the membrane into a small, ~0.05-μm-diameter membrane-bounded vesicle that contains the extracellular material. The vesicle is delivered to and fuses with an **endosome**, a sorting station formed of a vesicle and membrane-lined tubules (Figure 1-19). From this compartment, some membrane proteins are recycled back to the plasma membrane; other membrane proteins are transported in vesicles that eventually fuse with **lysosomes** for degradation. The entire endocytic pathway is described in detail in Chapter 14.

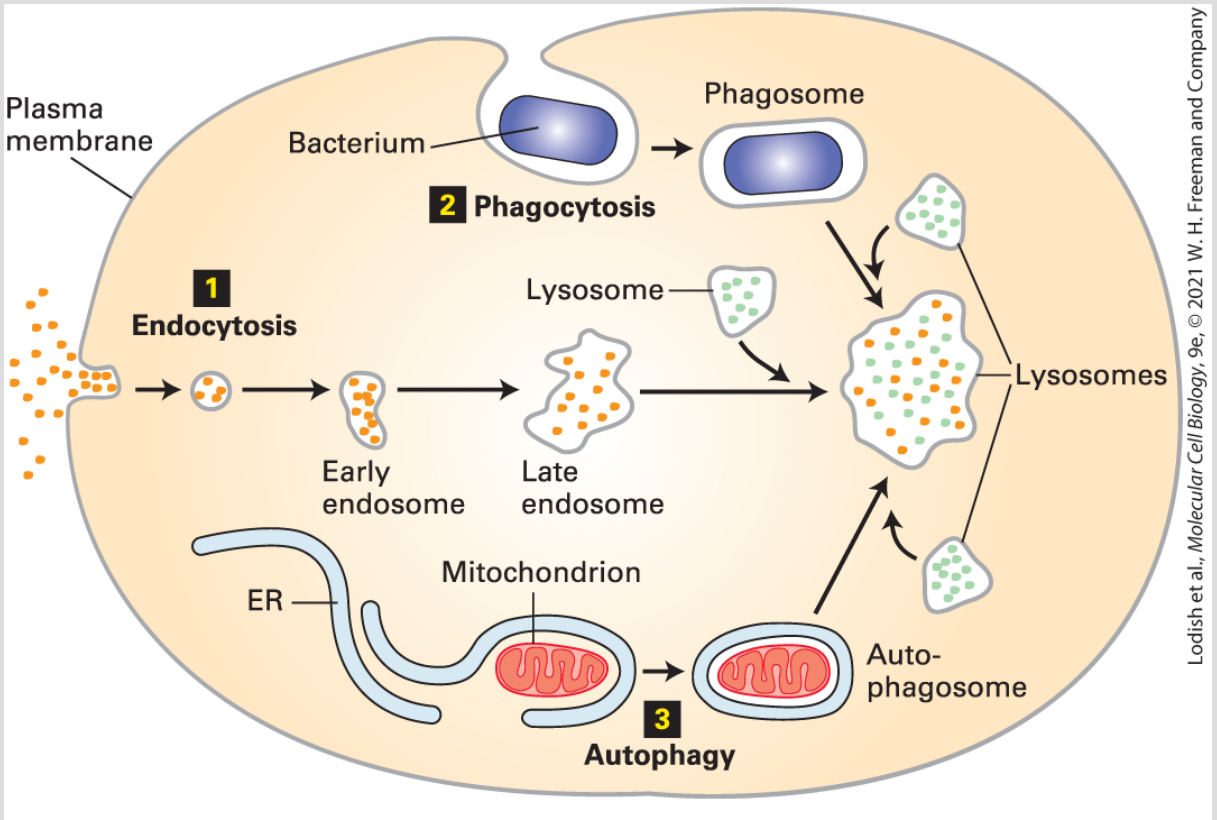


FIGURE 1-19 Endosomes and other cellular structures deliver materials to lysosomes. Schematic overview of three pathways by which materials are moved to lysosomes. Within the acidic lumen of a lysosome, hydrolytic enzymes degrade proteins, nucleic acids, lipids, and other large molecules.

1 Soluble macromolecules and molecules bound to proteins on the cell surface are taken into the cell by invagination of segments of the plasma membrane and delivered to lysosomes through the endocytic pathway; endosomes fuse with lysosomes, intermixing their contents.

2 Whole cells and other large, insoluble particles move from the cell surface to lysosomes through the phagocytic pathway; phagosomes fuse with lysosomes, intermixing their contents.

3 Worn-out organelles and bulk cytoplasm are delivered to lysosomes through the autophagic pathway.

Lysosomes Are Cellular Recycling Centers

Lysosomes contain a group of enzymes that degrade polymers into their monomeric subunits. Found exclusively in animal cells, lysosomes are responsible for degrading many components that have become obsolete for the cell or organism. As examples, nucleases degrade RNA and DNA into their mononucleotide building blocks; proteases degrade a variety of proteins and peptides into amino acids; phosphatases remove phosphate groups from mononucleotides, phospholipids, and other compounds; still other enzymes degrade complex polysaccharides and lipids into smaller units. The small molecules so generated are usually reutilized by the cell to make new proteins, nucleic acids, and carbohydrates. All of these lysosomal enzymes, collectively termed *acid hydrolases*, work most efficiently at acidic pH values. The acidic pH $\sim$5.0 found within lysosomes helps to denature proteins, making them accessible to the action of the lysosomal hydrolases. These enzymes are less active at the neutral pH of cells and most extracellular fluids. Thus if a lysosome releases its enzymes into the cytosol, where the pH is between 7.0 and 7.3, they cause little degradation of cytosolic components. The low pH of lysosomes is generated by proteins in the lysosomal membrane that pump protons (H^+ ions) from the cytosol into the lumen of the lysosome, as detailed in Chapter 11. Lysosomes provide an excellent example of the ability of intracellular membranes to form closed compartments in which the

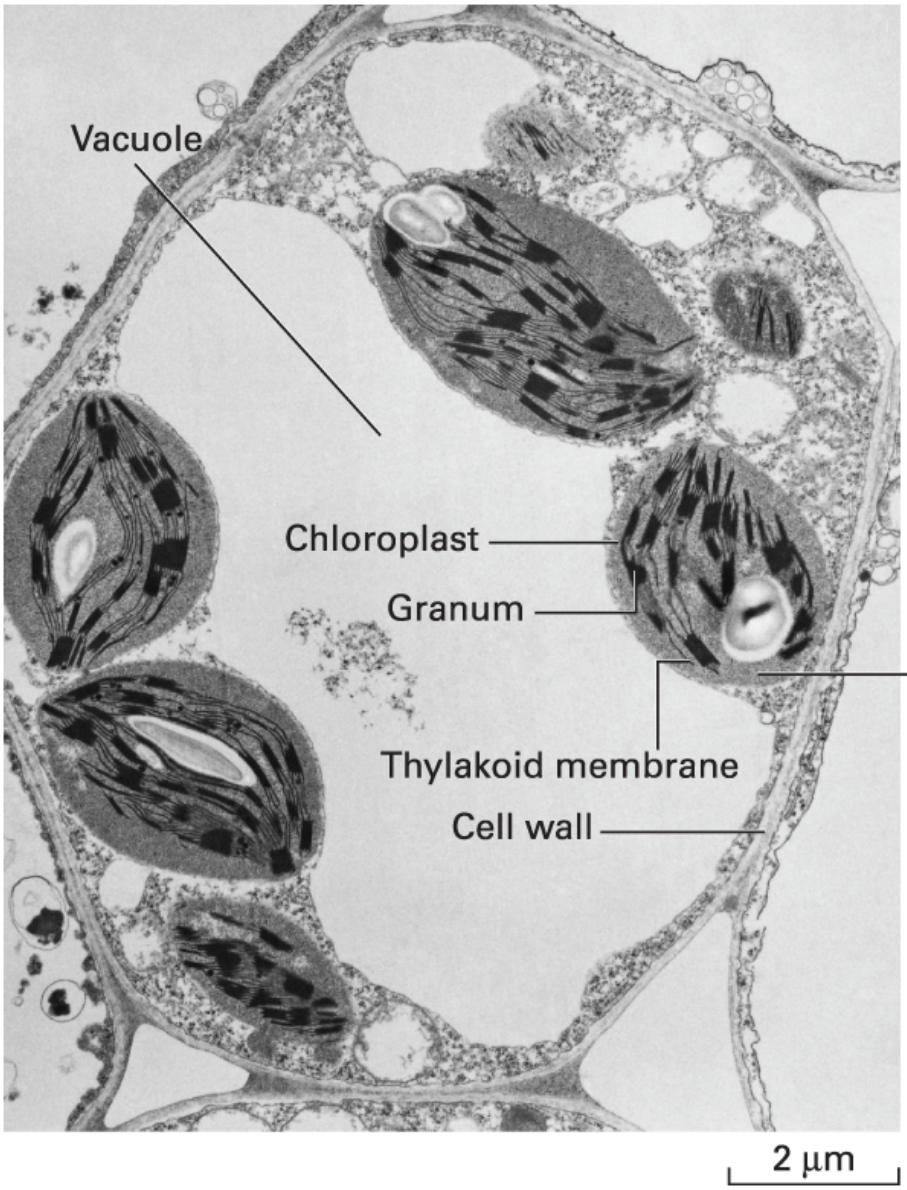
composition of the lumen (the aqueous interior of the compartment) differs substantially from that of the surrounding cytosol.

Aged organelles are degraded in a lysosome by a process called **autophagy** ("eating oneself"). Materials taken into a cell by endocytosis or phagocytosis may also be degraded in lysosomes (see Figure 1-19). In **phagocytosis**, large, insoluble particles (e.g., bacteria) are enveloped by the plasma membrane and internalized, and the particles are quickly delivered to lysosomes for degradation. Cytosolic and nuclear proteins generally are not degraded in lysosomes, but rather in proteasomes, large multiprotein complexes in the cytosol (see Chapter 3).

Plant Vacuoles Store Water, Ions, and Small-Molecule Nutrients Such as Sugars and Amino Acids

Most plant cells contain at least one membrane-bounded **vacuole**. A variety of membrane proteins in the vacuolar membrane transport water, ions, and small-molecule nutrients such as sugars and amino acids from the cytosol into the vacuole lumen. The number and size of vacuoles depend on both the type of cell and its stage of development; a single vacuole may occupy as much as 80 percent of a mature plant cell (Figure 1-20). Like that of a lysosome, the lumen

of a vacuole contains a battery of degradative enzymes and has an acidic pH, which is maintained by similar transport proteins in the vacuolar membrane. Thus plant vacuoles may also have a degradative function similar to that of lysosomes in animal cells. Similar storage vacuoles are found in green algae and in many microorganisms such as fungi.



Biophoto Associates/Science Source.

FIGURE 1-20 Electron micrograph of a thin section of a leaf cell. In this cell, a single large vacuole occupies much of the cell volume. Parts of five chloroplasts and the cell wall are also visible. Note the internal subcompartments in the chloroplasts.

Peroxisomes and Plant Glyoxisomes Metabolize Fatty Acids and Other Small Molecules Without Producing ATP from ADP and P_i

All animal cells (except erythrocytes) and many plant and fungal cells contain **peroxisomes**, a class of roughly spherical organelles 0.2–1.0 μm in diameter. A major function of peroxisomes is to catabolize (degrade) long-chain fatty acids to CO_2. As discussed in Chapter 12, mitochondria also contain enzymes that degrade fatty acids to CO_2, but in contrast to mitochondria, peroxisomes do not synthesize ATP from ADP and P_i. Several of the enzymes in peroxisomes are *oxidases* that use molecular oxygen to add oxygen atoms — a process termed *oxidation* — to a variety of organic substances and in the process form hydrogen peroxide (H_2O_2), a corrosive substance. Peroxisomes also contain copious amounts of the enzyme *catalase*, which degrades hydrogen peroxide to yield water and oxygen (see Chapter 12).

Plant seeds contain *glyoxisomes*, small organelles that oxidize stored lipids as a source of carbon and energy for growth. They are similar to peroxisomes and contain many of the same types of enzymes as well as additional ones used to convert fatty acids into glucose precursors.

Mitochondria Are the Principal Sites of ATP Production in Aerobic Cells

The activities of all parts of a cell are powered by the energy stored in ATP molecules. The main sites of ATP production during aerobic metabolism (metabolism requiring oxygen) are the **mitochondria** (Figure 1-21), which can occupy up to 25 percent of the volume of the cytoplasm. These complex organelles are generally exceeded in size only by the nucleus, vacuoles, and chloroplasts. The two membranes that are found in all mitochondria differ in composition and function. The **outer mitochondrial membrane** contains proteins that allow many molecules to move from the cytosol to the **intermembrane space** between the inner and outer membrane. The **inner mitochondrial membrane**, which is much less permeable, is about 20 percent lipid and 80 percent protein — a proportion of protein that is higher than those in other cellular membranes. The surface area of the inner membrane is greatly increased by a large number of infoldings, or **cristae**, that protrude into the **matrix**, the central aqueous space inside the inner membrane. The large

number of proteins in the inner membrane are needed to complete the aerobic degradation of fuels to form ATP from ADP and inorganic phosphate (see Figure 1-7 and Chapter 12).

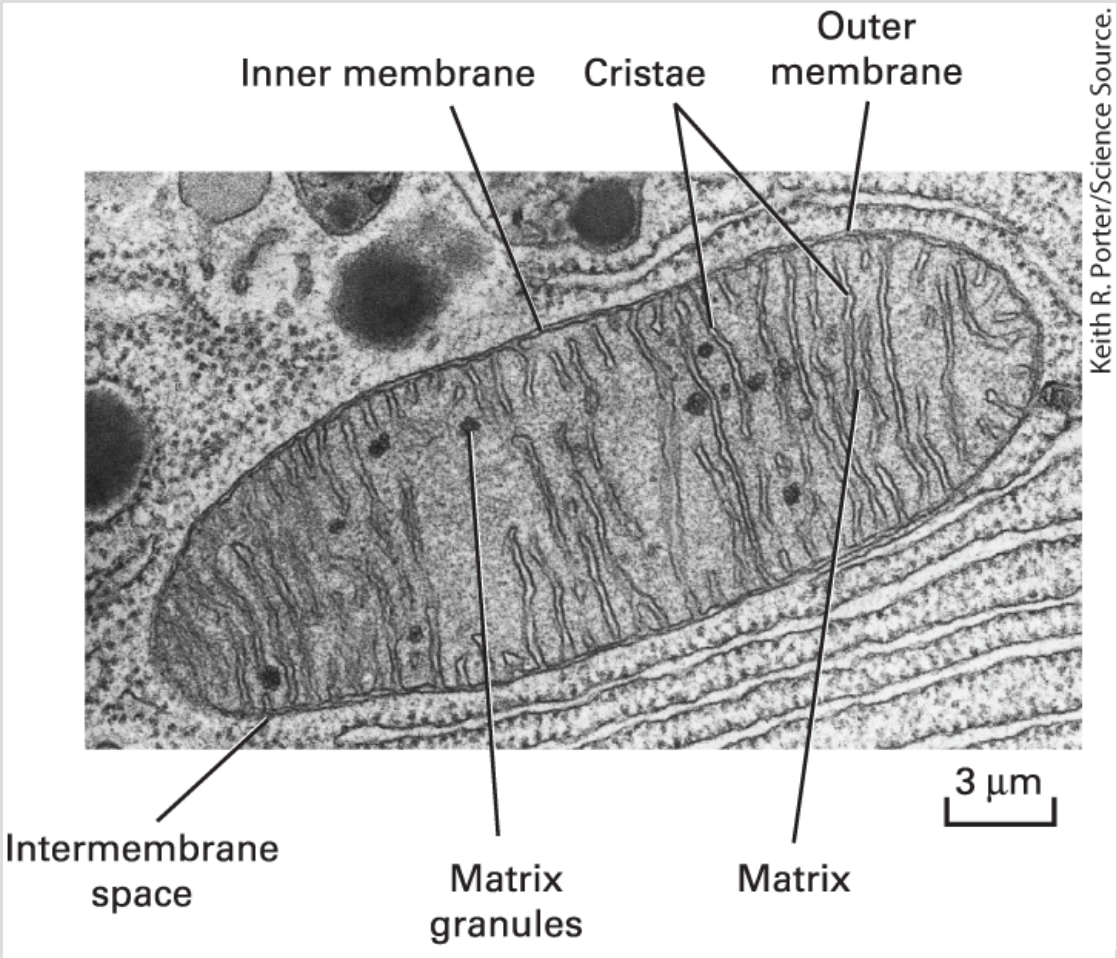


FIGURE 1-21 Electron micrograph of a mitochondrion in a pancreas cell. The smooth outer membrane forms the outside boundary of the mitochondrion. The inner membrane is distinct from the outer membrane and is highly invaginated to form sheets and tubes called *cristae*; ATP is produced by proteins embedded in the membranes of the cristae. The aqueous space between the inner and outer membranes (the intermembrane space) and the space inside the inner membrane (the matrix) each contain specific proteins important for the metabolism of sugars, lipids, and other molecules.

In non-photosynthetic cells, the principal fuels for ATP synthesis are fatty acids and glucose. The complete aerobic degradation of 1 molecule of glucose to carbon dioxide and water is coupled to the synthesis of as many as 30 molecules of ATP from ADP and inorganic phosphate (see Figure 1-7). In eukaryotic cells, the initial stages of glucose degradation take place in the cytosol, where 2 ATP molecules per glucose molecule are generated. The terminal stages of oxidation and ATP synthesis are carried out by enzymes in the mitochondrial matrix and inner membrane (see Chapter 12); as many as 28 ATP molecules per glucose molecule are generated in mitochondria. Similarly, virtually all the ATP formed in the oxidation of fatty acids to carbon dioxide is generated in mitochondria. Thus mitochondria can be regarded as the power plants of the cell.

As discussed in Chapter 12, the popular *endosymbiont hypothesis* postulates that mitochondria originated by endocytosis of an ancient aerobic bacterium by the precursor of a eukaryotic cell; the bacterial plasma membrane evolved to become the inner mitochondrial membrane. Indeed, mitochondria contain small DNA molecules that encode a small number of mitochondrial proteins and that are thought to have descended from the DNA genome of this ancient bacterium.

Chloroplasts Contain Internal Compartments in Which Photosynthesis Takes Place

Except for vacuoles, **chloroplasts** are the largest and the most characteristic organelles in the cells of plants and green algae (see Figure 1-20). The endosymbiont hypothesis (see Chapter 12) posits that these organelles originated by endocytosis of a primitive photosynthetic bacterium. Chloroplasts can be as long as 10 μm and are typically 0.5–2 μm wide, but they vary in size and shape in different cells, especially among the algae. In addition to the inner and outer membranes that bound a chloroplast, this organelle also contains an extensive internal system of interconnected membrane-bounded vesicles called **thylakoids**, which are flattened to form disks. Thylakoids often form stacks called *grana* and are embedded in an aqueous matrix termed the *stroma*. The thylakoid membranes contain green pigments (chlorophylls) and other pigments that absorb light, as well as enzymes that generate ATP from ADP and inorganic phosphate during photosynthesis. Some of the ATP is used to convert carbon dioxide into three-carbon intermediates by enzymes located in the stroma; the intermediates are then exported to the cytosol and converted into sugars.

The molecular mechanisms by which ATP is formed in mitochondria and chloroplasts are very similar, as explained in Chapter 12. Besides being surrounded by two membranes,

chloroplasts and mitochondria have other features in common: both often migrate from place to place within cells, and both contain their own DNA, which encodes some of the key organelle proteins (see Chapter 12). The proteins encoded by mitochondrial or chloroplast DNA are synthesized on ribosomes within the organelles. However, most of the proteins in each organelle are encoded in nuclear DNA and are synthesized in the cytosol; these proteins are then incorporated into the organelles and directed to the correct membrane or space by processes described in Chapter 13.

Many Organelle-Like Structures Are Unbounded by a Membrane

In addition to membrane-enveloped organelles, eukaryotic cells contain a variety of substructures in which important biochemical functions are concentrated but that are not surrounded by a membrane. These membraneless organelles range in size from 0.1–3 micrometers in diameter. One ubiquitous example is the **nucleolus**, a subcompartment of the nucleus where ribosomal RNA is synthesized and ribosomes are assembled (see Figure 1-16b and Chapter 9). Other subcompartments of the nucleus have recently been made visible by the new microscopic techniques detailed in Chapter 4. These include **nuclear speckles**, sites where RNA

polymerase and other gene regulatory proteins are concentrated and where messenger RNAs are synthesized and processed.

Subcompartments of the cytosol include **P bodies**, sites where many messenger RNAs are degraded, and **cytosolic stress granules**, membraneless organelles where some messenger RNAs are stored awaiting a signal to be translated into a protein.

Although much remains to be learned about the organization of these membraneless organelles, recent evidence suggests that they are formed by many weak but specific interactions between proteins and nucleic acids that exclude the surrounding aqueous phase. Their properties are characteristic of liquid droplets suspended in another, chemically distinct liquid, and they are often referred to as **biomolecular condensates** (see Chapter 3). Their structures are highly dynamic, in that proteins can diffuse into and out of these condensates. Clearly much needs to be learned about their formation, structure, and biological functions.

All Eukaryotic Cells Use a Similar Cycle to Regulate Their Division

Unicellular eukaryotes, animals, and plants all use essentially the same **cell cycle**, the series of events that prepares a cell to divide, and the same actual division process, called **mitosis**. The eukaryotic cell cycle is commonly divided into four phases (Figure 1-22). The

chromosomes and the DNA they carry are duplicated during the **S (synthesis) phase**. The replicated chromosomes condense (see Figure 1-17) and separate during the **M (mitotic) phase** during which the cell divides, and each daughter cell gets one copy of each chromosome. The M and S phases are separated by two gap phases, the **G_1 phase** and the **G_2 phase**, during which mRNAs, proteins, lipids, and other cell constituents are made and the cell increases in size.

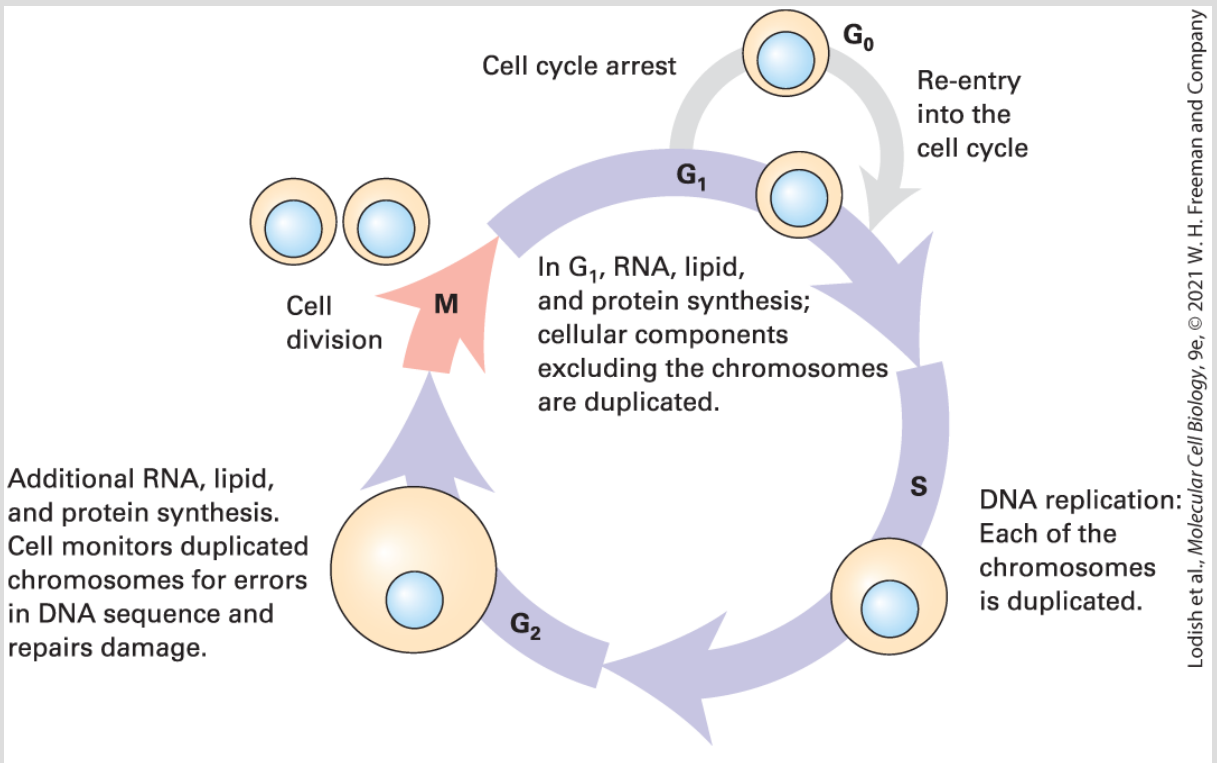


FIGURE 1-22 During growth, all eukaryotic cells continually progress through the four phases of the cell cycle. In proliferating cells, the four phases of the cell cycle proceed successively. In humans, the cycle takes from 10 to 20 hours depending on cell type and developmental state. Yeasts divide much faster. During interphase, which consists of the G_1, S, and G_2 phases, the cell roughly doubles its mass. Replication of DNA during the S phase leaves the cell with two copies of each type of chromosome (see Figure 1-17c). In the mitotic (M) phase, the chromosomes are evenly partitioned into two daughter cells, and in

most cases the cytoplasm divides roughly in half by division along the midline. Under certain conditions, such as starvation or when a tissue has reached its final size, cells will stop cycling and remain in a waiting state called G_0. Some types of cells in G_0 can re-enter the cell cycle if conditions change.

Under optimal conditions, some bacteria, such as *E. coli*, can divide to form two daughter cells once every 20 minutes. Most eukaryotic cells take considerably longer to grow and divide, generally several hours. Moreover, the cell cycle in eukaryotes is highly regulated (see Chapter 19). This tight control prevents imbalanced, excessive growth of cells and tissues if essential nutrients or certain hormonal signals are lacking. Some highly specialized cells in adult animals, such as neurons and heart muscle cells, divide rarely, if at all. However, an organism usually replaces worn-out cells or makes more cells in response to a new need — for example, new muscle cells are generated from undifferentiated stem cells in response to exercise or damage. Another example is the formation of additional red blood cells when a person ascends to a higher altitude and needs more capacity to capture oxygen. The fundamental defect in cancer is loss of the ability to control the growth and division of cells. In Chapter 25, we examine the molecular and cellular events that lead to inappropriate, uncontrolled proliferation of cells.

1.4 Unicellular Eukaryotic Organisms Widely Used in Cell Biology Research

Our current understanding of the molecular functioning of eukaryotic cells largely rests on studies of just a few types of organisms, often termed *model organisms* (Figure 1-23). Because of the evolutionary conservation of genes, proteins, organelles, cell types, and so forth, discoveries about biological structures and functions obtained with one experimental organism often apply to others. Thus researchers generally conduct studies with the organism that is most suitable for rapidly and completely answering the question being posed, knowing that the results obtained in one organism are likely to be broadly applicable.

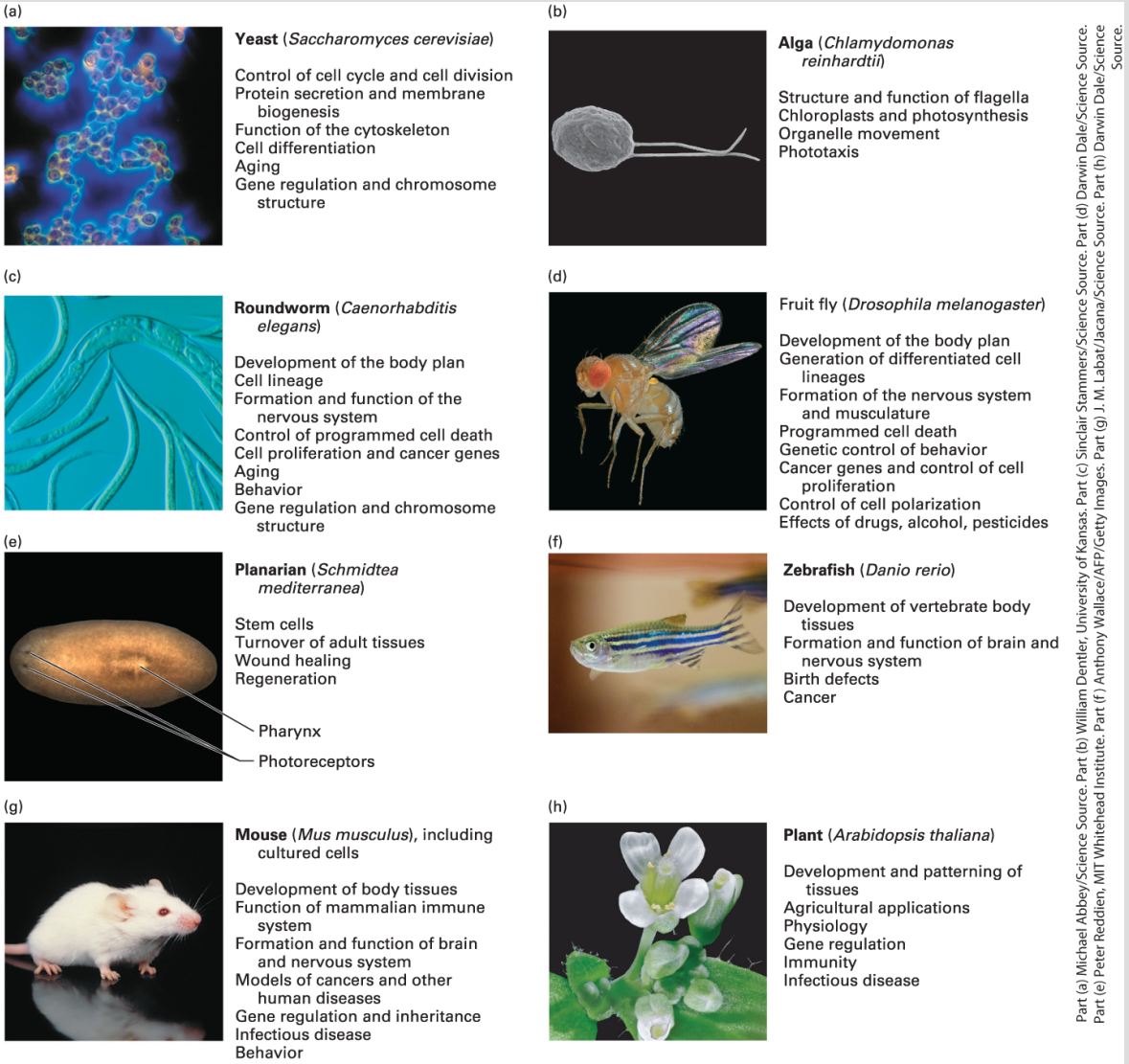


FIGURE 1-23 Many eukaryotic organisms used in cell biology research have advantages for certain types of studies. The yeast *Saccharomyces cerevisiae* (a) has the cellular organization of a eukaryote but is a relatively simple single-celled organism that is easy to grow and to manipulate genetically. The green alga *Chlamydomonas reinhardtii* (b) is widely used to study photosynthesis and the structure and function of flagella. In the roundworm *Caenorhabditis elegans* (c), which has a small number of cells arranged in a nearly identical way in every worm, the formation of each individual cell can be traced. The fruit fly *Drosophila melanogaster* (d), first used to discover the properties of chromosomes, has been especially valuable in identifying genes that control embryonic development. Many of these genes are evolutionarily conserved in humans. Planaria (e) are flatworms that can regenerate any part of the body that is cut off, including the head and the photoreceptors.

The stem cells that give rise to their new cells and tissues are widely studied. The zebrafish *Danio rerio* (f) is used for rapid genetic screens to identify genes that control vertebrate development and organogenesis. Of the experimental animal systems, mice (*Mus musculus*) (g) are evolutionarily the closest to humans and have thus provided models for studying numerous human genetic and infectious diseases. The mustard-family weed *Arabidopsis thaliana* (h) has been used for genetic screens to identify genes involved in nearly every aspect of plant life.

As we have seen, bacteria are excellent experimental systems for studies of several cellular functions, but they lack the organelles found in eukaryotes. Unicellular eukaryotes such as yeasts are used to study many fundamental aspects of eukaryotic cell structure and function. Metazoan organisms such as the roundworm, fruit fly, and mouse are required to study more complex tissue and organ systems and development. As we will see in this section and the next, these and other metazoans are widely used to understand complex cell systems and mechanisms.

Yeasts Are Used to Study Fundamental Aspects of Eukaryotic Cell Structure and Function

One group of single-celled eukaryotes, the yeasts, has proven exceptionally useful in molecular and genetic analysis of eukaryotic cell formation and function. Yeasts and their multicellular cousins, the molds, are fungi. Fungi have important ecological roles in

breaking down plant and animal remains for reuse. They also make numerous antibiotics and are used in the manufacture of bread, beer, and wine.

The common yeast used to make bread and beer, *Saccharomyces cerevisiae* (Figure 1-24a), appears frequently in this book because it has proved to be an extremely useful experimental organism. *S. cerevisiae* is a spherical cell that divides by budding of a daughter cell from a side of the mother cell, eventually breaking free (Figure 1-24a, b). Homologs of many of the approximately 6600 different proteins expressed in an *S. cerevisiae* cell (see Table 1-2) are found in most, if not all, eukaryotes and are important for cell division or for the functioning of individual eukaryotic organelles. Much of what we know of the proteins in the endoplasmic reticulum and Golgi complex that promote protein secretion was elucidated first in genetic studies of yeasts (see Chapter 14). Yeasts were also essential for the identification of many proteins that regulate the cell cycle and catalyze DNA replication and transcription.

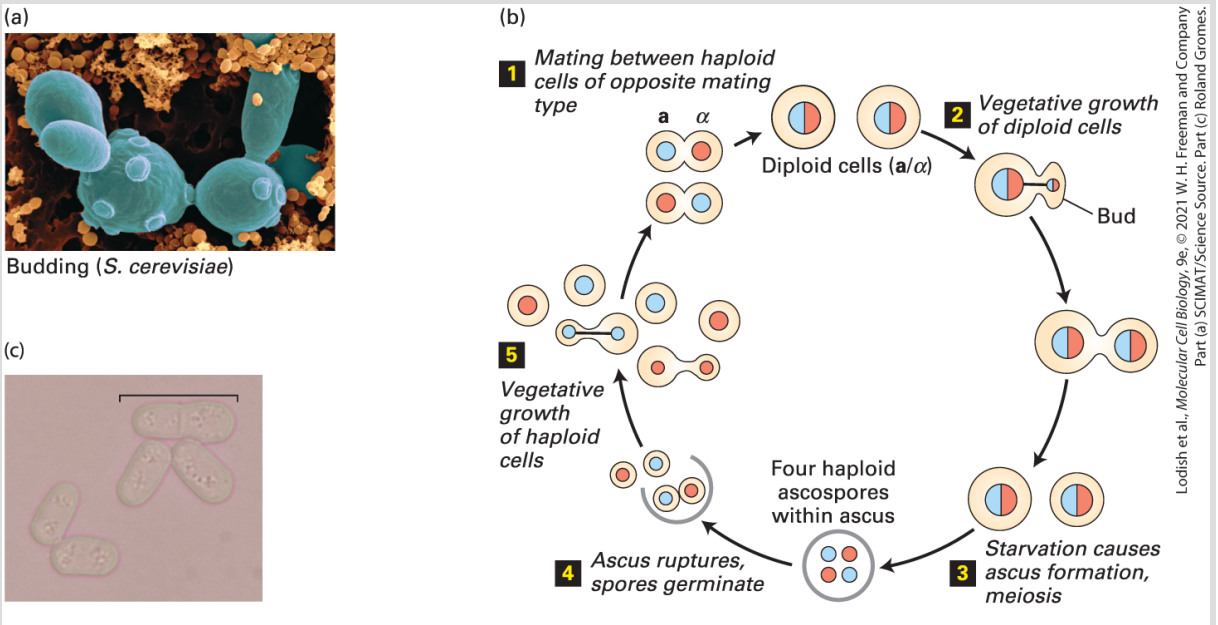


FIGURE 1-24 The yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* can be haploid or diploid and can reproduce sexually or asexually. (a) Scanning electron micrograph of the budding yeast *Saccharomyces cerevisiae*. These cells grow by an unusual type of mitosis termed *Smitotic budding*. One daughter nucleus remains in the mother cell; the other daughter nucleus is transported into the bud, which grows in size and soon is released as a new cell. After each bud cell breaks free, a scar is left at the budding site, so the number of previous buds on the parent cell can be counted. The orange-colored cells are bacteria. (b) Haploid *S. cerevisiae* cells can have different mating types, called **a** (blue) and α (orange). Both types contain a single copy of each yeast chromosome, half the usual number, and grow by mitotic budding. Two haploid cells that differ in mating type, one **a** and one α, can fuse together to form an **a**/α diploid cell. An **a** haploid cell secretes a pheromone protein termed the **a** *mating factor* that binds to a receptor protein on the surface of an α cell, and conversely an α cell secretes the α factor protein that binds to a different surface receptor on the **a** cell. Thus each type of cell recognizes the mating factor produced by the opposite type, and these hormones trigger fusion of the **a** and α cell. The **a**/α diploid cell contains two copies of each chromosome; diploid cells can multiply by mitotic budding in a manner similar to that of the haploid cells. Under starvation conditions, a diploid cell can undergo meiosis, a special type of cell division, to form four haploid ascospores. Rupture of an ascus releases four haploid spores, which can germinate into haploid **a** and α cells. These cells can also multiply asexually. (c) Both haploid and diploid *Schizosaccharomyces pombe* cells have a mitotic division cycle different from that of *S.*

cerevisiae, one that more resembles that of human cells than that of *S. cerevisiae*. Bracket indicates a cell undergoing mitotic division.

The yeast species *Schizosaccharomyces pombe* diverged from *S. cerevisiae* between 300 and 600 million years ago and is also widely used in research. Unlike *S. cerevisiae, S. pombe* is a rod-shaped cell that grows by elongation at the ends; like human cells it divides by division at the cell midpoint (Figure 1-24c; see also Figure 1-22). Like human cells, but unlike *S. cerevisiae*, most genes in *S. pombe* have introns that are removed from RNA transcripts by processing in the nucleus (see Figure 1-10). *S. cerevisiae, S. pombe*, and other yeasts offer many advantages to molecular and cellular biologists:

- Vast numbers of yeast cells can be grown easily and cheaply in culture from a single cell; the cells in such **clones** are genetically identical and have the same biochemical properties. Individual proteins or multiprotein complexes can be purified from large amounts of cells and then studied in detail.
- Yeast cells may be either haploid (containing one copy of each chromosome) or diploid (containing two copies of each chromosome), and both forms can divide by mitosis; this ability makes isolating and characterizing mutations in genes encoding essential yeast cell proteins relatively straightforward, as we see in the next section.
- Yeasts, like many organisms, have a sexual cycle that allows exchange of genes between cells. Under starvation conditions, diploid cells undergo meiosis (see Chapter 19) to form haploid

daughter cells. Haploid *S. cerevisiae* cells are of two types, **a** and α cells. If haploid **a** and α cells come close to one another, they secrete hormones that induce the two cells to fuse, forming an **a**/α diploid cell that contains two copies of each chromosome, one from each parent cell (Figure 1-24b).

With the use of a single species such as *S. cerevisiae* as an experimental organism, results from studies carried out by tens of thousands of scientists worldwide, using multiple investigational techniques, can be combined to yield a deeper level of understanding of a single type of cell. As we will see many times in this book, conclusions based on studies of *S. cerevisiae* and *S. pombe* have often proved true for all eukaryotes. These conclusions have provided a basis for exploring the evolution of more complex processes in multicellular animals and plants.

Mutations in Yeast Led to the Identification of Key Cell Cycle Proteins

Biochemical studies can tell us much about an individual protein, but they cannot prove that it is required for cell division or any other cell process. The importance of a protein is demonstrated most firmly if a mutation that prevents its synthesis or makes it nonfunctional adversely affects the process under study.

In a classical genetics approach, scientists isolate and characterize mutants that lack the ability to do something a normal organism can do. Often large genetic screens are done to look for many different mutant individuals (e.g., fruit flies, yeast cells) that are unable to complete a certain process, such as cell division or muscle formation. Mutations are usually produced by treatment with a *mutagen*, a chemical or physical agent that promotes mutations in a largely random fashion. However, inactivating genes that encode proteins for an essential cell process will kill the organism. How then can we isolate and maintain mutant organisms or cells that are defective in some process, such as cell division or protein secretion, that is essential for survival?

One way is to isolate organisms with a **temperature-sensitive mutation**. These mutants are able to grow at the *permissive temperature*, but not at another, usually higher temperature, the *nonpermissive temperature*. Normal cells can grow at either temperature. In most cases, a temperature-sensitive mutant produces an altered protein that works at the permissive temperature but unfolds and is nonfunctional at the nonpermissive temperature. Screens for temperature-sensitive mutations are most readily done with haploid organisms such as yeasts because they have only one copy of each gene, and thus a mutation in it will immediately have a consequence.

By analyzing the effects of numerous different temperature-sensitive mutations that altered the division of haploid yeast cells, geneticists discovered most of the genes necessary for cell growth and division without knowing anything, initially, about which proteins they encode or how these proteins participate in the process. In general, the great power of genetics is to reveal the existence and relevance of all proteins required for a particular cell function without prior knowledge of their biochemical identity or molecular function. These mutation-defined genes can be isolated using recombinant DNA techniques discussed in Chapter 6. With the isolated genes in hand, the encoded proteins can be produced in a test tube or in engineered bacteria or cultured cells. In this way, biochemists can investigate whether the genes necessary for cell division encode proteins that associate with other proteins or DNA or catalyze particular chemical reactions during cell division (see Chapter 19).

Most of these yeast cell cycle genes are found in human cells as well, and the encoded proteins have similar amino acid sequences. Proteins from different organisms, but with similar amino acid sequences, are said to be **homologous**; such proteins often have the same or similar functions. Remarkably, it has been shown that a human protein that regulates the cell cycle, when expressed in a mutant yeast defective in the homologous yeast protein, is able to rescue the defect of the mutant yeast (i.e., to allow the cell to grow normally), thus demonstrating the protein's ability to function in a very different type of eukaryotic cell. This experimental result,

which together with identification and characterization of many other cell cycle genes garnered a Nobel Prize for Paul Nurse, was especially notable because the common ancestor of present-day yeasts, plants, and humans is thought to have lived over a billion years ago. Clearly the eukaryotic cell cycle and many of the genes and proteins that catalyze and regulate it evolved early in biological evolution and have remained quite constant over a very long period of evolutionary time. Subsequent studies showed that mutations in many yeast cell cycle proteins that allow uncontrolled cell growth also frequently occur in human cancers (see Chapter 25), again attesting to the important conserved functions of these proteins in all eukaryotes.

Studies in the Alga *Chlamydomonas reinhardtii* Led to the Development of a Powerful Technique to Study Brain Function

The green unicellular alga *Chlamydomonas reinhardtii* (see Figure 1-23b), which swims using its two long flagella, is widely used in studies of the structure, function, and assembly of this organelle. In part because of the powerful genetic techniques now available, *Chlamydomonas* is also used in studies of chloroplast formation and photosynthesis. The *Chlamydomonas* genome (see Table 1-2) encodes many more proteins than do those of yeasts, including flagellar

proteins and proteins needed to build a chloroplast, organelles not found in yeasts.

One important outcome of the use of this experimental organism came from studies of phototaxis, the behavior in which an organism moves toward or away from a source of light. *Chlamydomonas* needs to move toward light to undergo photosynthesis and thus generate the energy it needs to grow and divide, but light that is too intense repels it, as it causes damage to the chloroplast. Studies of *Chlamydomonas* phototaxis led to the discovery of two proteins in its plasma membrane that, when they absorb light, open a channel in the membrane that allows ions such as Ca^{2+} to flow from the extracellular medium into the cytosol, triggering phototactic responses. As detailed in Chapter 23, recombinant DNA techniques have been used to express one such protein in specific neurons in the mouse brain, allowing investigators to activate just one or a few cells in the brain using a point source of light. Thus studies on this humble alga have led to the development of an important experimental system — optogenetics — for the study of brain function.

The Parasite That Causes Malaria Has Novel Organelles That Allow It to Undergo a Remarkable Life Cycle

Whereas yeasts are used in the manufacture of bread, beer, wine, and cheese, some unicellular eukaryotes cause major human diseases and are widely studied in an attempt to develop vaccines or drugs that will kill them but not injure their human host. *Entamoeba histolytica* causes dysentery; *Trichomonas vaginalis*, vaginitis; and *Trypanosoma brucei*, sleeping sickness. Each year the worst of these protozoans, *Plasmodium falciparum* and related species, cause more than 200 million new cases of malaria, a disease that kills 1.5 million to 3 million people annually. These protozoans inhabit mammals and mosquitoes alternately, changing their morphology and behavior in response to signals in each of these environments.

The complex life cycle of *Plasmodium* dramatically illustrates how a single cell can adapt to multiple environments (Figure 1-25a). Additionally, the *merozoite* form that infects human red blood cells contains several organelles, not found in most eukaryotes, that enable the parasite to invade a red blood cell, including the rhoptries, polar ring, and micronemes, as well as a fuzzy surface coat on the outside of the cell (Figure 1-25b, c). Entry of the parasite into a red blood cell is initiated by binding of certain parasite cell-surface proteins to proteins on the red blood cell surface, accompanied by the secretion of proteins stored in the micronemes and rhoptries. A tight junction is formed between the two plasma membranes, and the fuzzy coat is lost as parasites enter the host cell (Figure 1-25c).

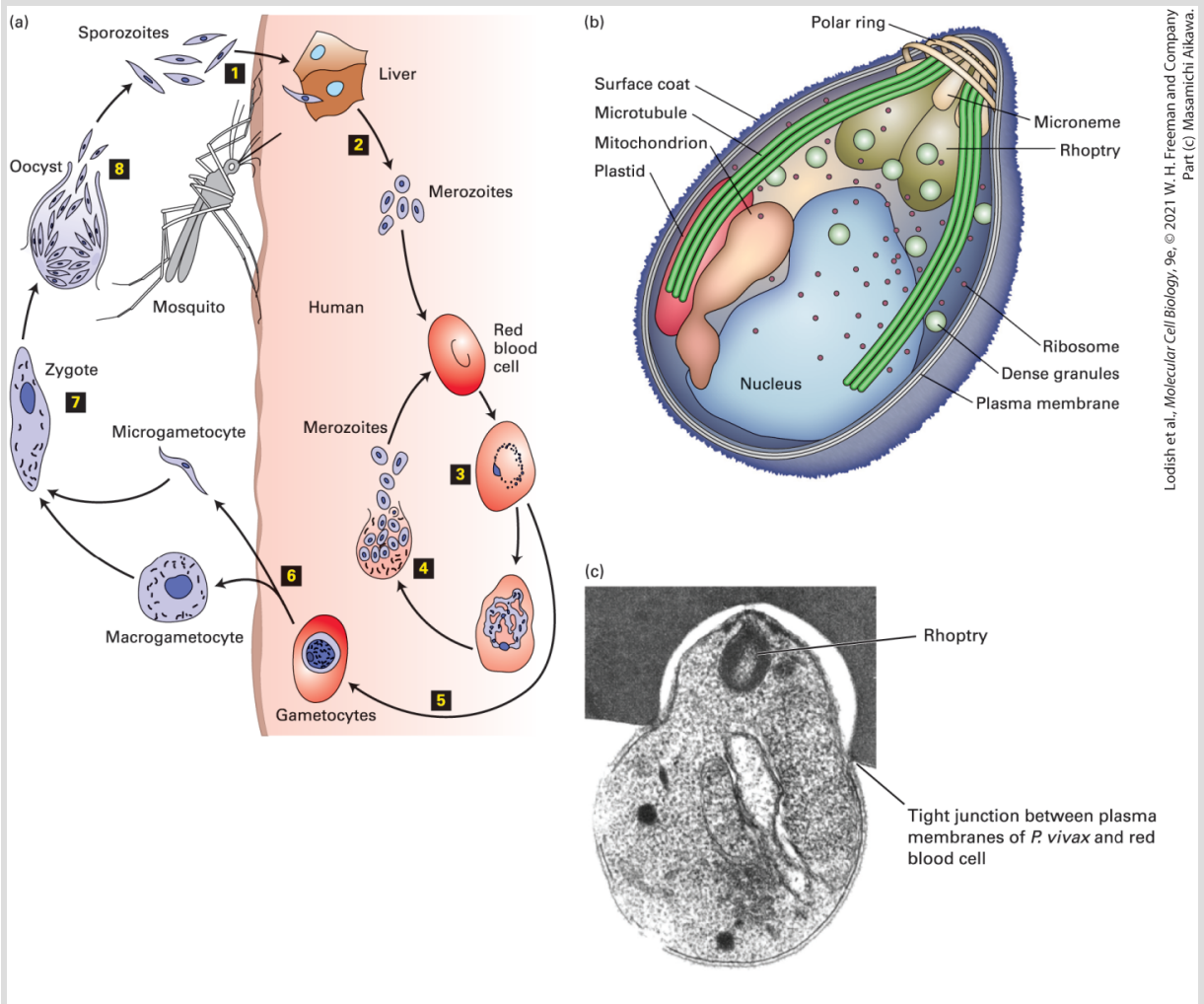


FIGURE 1-25 *Plasmodium* species, the parasites that cause malaria, are single-celled protozoans with a remarkable life cycle. Many *Plasmodium* species are known, and they can infect a variety of animals, cycling between insect and vertebrate hosts. The four species that cause malaria in humans undergo several dramatic transformations within their human and mosquito hosts. (a) Diagram of the life cycle. Like many yeasts, *Plasmodium* spends most of its life cycle as a haploid organism, with a single copy of most genes. Step 1: Haploid sporozoites enter a human host when an infected *Anopheles* mosquito bites a person. Step 2: They migrate to the liver, where they develop into merozoites, which are released into the blood. Merozoites differ substantially from sporozoites, so this transformation is a metamorphosis (Greek, “to transform” or “many shapes”). Step 3: Circulating merozoites invade red blood cells (RBCs) and reproduce within them. Proteins produced by some *Plasmodium* species move to the surface of infected RBCs, causing the cells to adhere to the walls of blood vessels. This prevents infected RBCs from circulating to the spleen, where cells of the immune system would destroy the RBCs and the *Plasmodium*

organisms they harbor. Step 4: After growing and reproducing in RBCs for a period of time characteristic of each *Plasmodium* species, the merozoites suddenly burst forth in synchrony from large numbers of infected cells. It is this event that brings on the fevers and shaking chills that are the well-known symptoms of malaria. Some of the released merozoites infect additional RBCs, creating a cycle of reproduction and infection. Step 5: Eventually, some merozoites develop into male and female gametocytes, another metamorphosis. These cells do not replicate further, waiting to be transferred to an *Anopheles* mosquito during a bloodmeal. Step 6: In the mosquito's stomach, the gametocytes are transformed into microgametocytes (sperm) or macrogametocytes (eggs), yet another metamorphosis marked by development of slender motile sperm. Step 7: Fusion of sperm and eggs generates the diploid zygotes, which implant into the lining of the stomach and grow into oocysts. Within the oocysts, parasites undergo meiosis, returning to a haploid chromosome content, followed by several rounds of mitosis that essentially are factories for producing sporozoites. Step 8: Rupture of an oocyst releases thousands of sporozoites, which migrate to the salivary glands, setting the stage for infection of another human host. (b) Organelles of the *Plasmodium vivax* merozoite. Some of these organelles are found only in *Plasmodium* and related eukaryotic parasitic microorganisms. (c) Section of a *Plasmodium vivax* merozoite invading a human red blood cell. See A. Cowman and B. Crabb, 2006, *Cell* **124**:755–766.

All the cellular transformations that occur during the *Plasmodium* life cycle are governed by instructions encoded in the genetic material of this parasite (see Table 1-2). The *Plasmodium* genome has about the same number of protein-coding genes as the yeast *S. cerevisiae*, but about two-thirds of the *Plasmodium* genes appear to be unique to this and related parasites. Their many unique genes attest to the great evolutionary distance between these parasites, members of the genus **Apicomplexa** (see Figure 1-1), and most other eukaryotes as well as the presence of unusual organelles required for their complex life cycles.

1.5 Metazoan Structure, Function, Evolution, and Differentiation

The evolution of multicellular organisms most likely began when cells remained associated in small colonies after division instead of separating into individual cells. A few prokaryotes and the colonial algae *Volvox* (see Figure 1-3d), as well as many fungi and slime molds, exhibit such rudimentary social behavior. The full flowering of multicellularity, however, occurred in the eukaryotic organisms called *metazoans* whose cells became differentiated and organized into groups, or tissues, in which the different cells performed specialized functions. In this section, we describe some of the principal features of metazoans: adhesion of cells to each other, communication between cells, and regulation of development such that the correct cells are produced at the right time in the right place in the organism.

Multicellularity Requires Cell-Cell and Cell-Matrix Adhesions

The cells of higher plants are encased in a network of chambers formed by the interlocking cell walls surrounding the cells and are connected by cytoplasmic bridges called **plasmodesmata** (see Figure 1-13a). Animal cells are often "glued" together into a chain, a ball, or a sheet by cell-adhesion proteins on their surfaces, often

called **cell-adhesion molecules**, or **CAMs** (see Figure 1-4d and Chapter 20). Some CAMs bind cells to one another; other types bind cells to the fibrous proteins and sugar polymers (polysaccharides) that form the **extracellular matrix**, forming a cohesive unit. In animals, the matrix cushions cells and allows nutrients to diffuse toward them and waste products to diffuse away. A specialized, especially tough matrix called the **basal lamina**, made up of polysaccharides and multiple proteins such as collagen, forms a supporting layer underlying cell sheets and prevents the cell aggregates from ripping apart (see Figure 1-4). Many CAMs and extracellular-matrix proteins found in humans also occur in invertebrates, indicating their importance during metazoan evolution.

Similarly, many of the proteins and small molecules used by metazoans as signaling molecules are conserved in humans and many invertebrates, as are their **receptors**, the cellular proteins that bind to these signaling molecules and trigger an effect in the receiving cell.

Epithelia Originated Early in Evolution

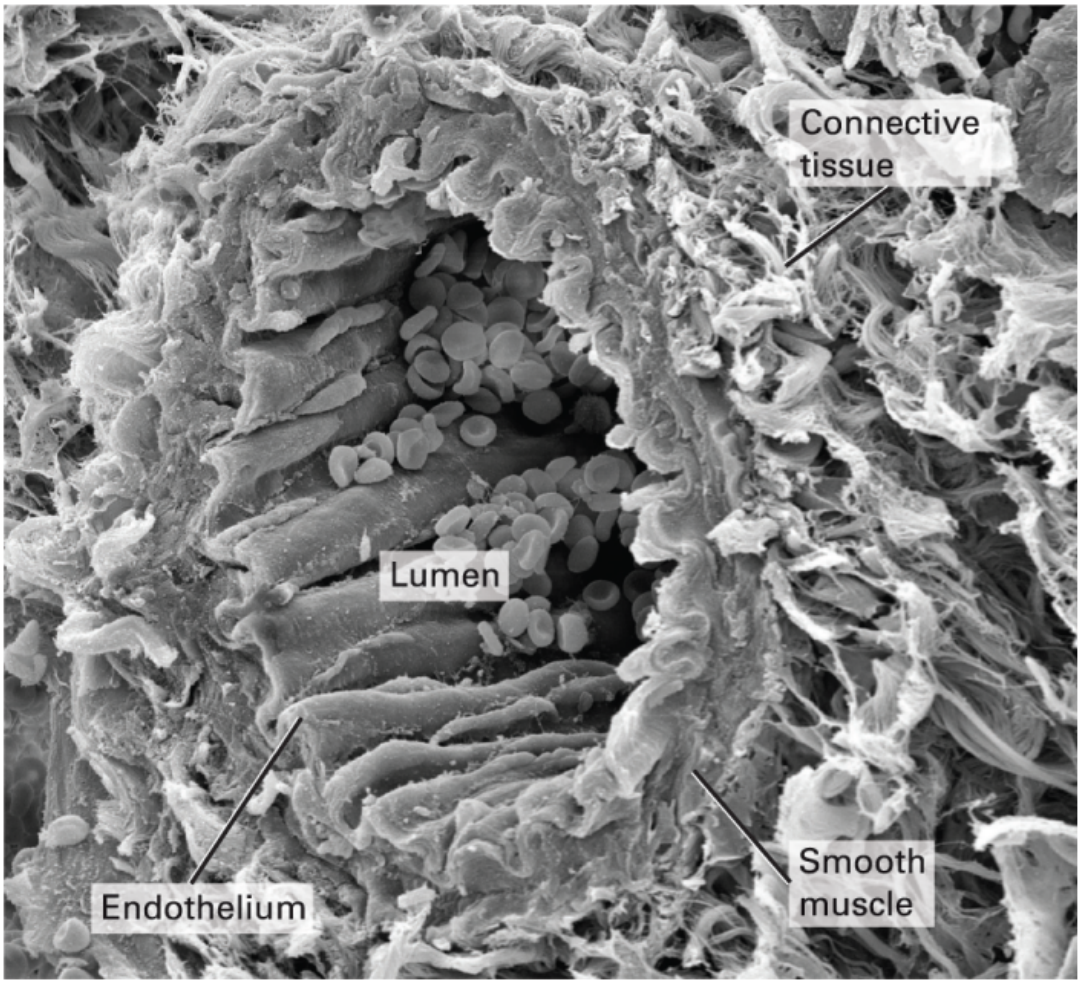
Metazoans, which are thought to have evolved in an ocean-like, saline environment, had to solve a fundamental problem: separating the inside of the organism from the outside. The external surfaces of all metazoan animals, as well as the surfaces of their internal

organs, are covered by a sheet-like layer of tissue called an **epithelium**. Epithelia commonly serve as barriers and protective surfaces, as exemplified by the sheets of epidermal cells that form the skin (see Figure 1-4). Other epithelia are one cell-layer thick and line internal organs such as the small intestine, where they are crucial for uptake and transport of the products of digestion (e.g., glucose and amino acids) into the blood (see Chapter 11). As discussed in Chapter 20, epithelia in different body locations have characteristic morphologies and functions.

Cells that form epithelial tissues are said to be *polarized* because their plasma membranes are organized into at least two discrete regions. Typically, the distinct surfaces of a polarized epithelial cell are the **apical** surface — the "top" of the cell facing the external world — and the **basal** and **lateral** (collectively, **basolateral**) surfaces that face the organism's interior. As shown in Figure 1-4, the basal surface usually contacts an underlying extracellular matrix, the basal lamina. As discussed in Chapter 20, specialized junction proteins in the basolateral plasma membrane link adjacent cells together and also bind the cells to the basal lamina.

Cells Are Organized into Tissues and Tissues into Organs

Cells in metazoans do not work in isolation; specialized groups of differentiated cells often form tissues, which are themselves the major components of organs. For example, a blood vessel is composed of layers of different types of tissue. The lumen of a small blood vessel is lined with a sheet-like layer of endothelial cells, or **endothelium**, which prevents blood cells from leaking out (Figure 1-26). A layer of smooth muscle tissue encircles the endothelium and basal lamina and contracts to limit blood flow. During times of fright, constriction of smaller peripheral vessels forces more blood to the vital organs. The muscle layer of a blood vessel is wrapped in an outer layer of connective tissue, a network of fibers and cells that encases the vessel walls and protects them from stretching and rupture.



SPL/Science Source.

FIGURE 1-26 All organs are organized arrangements of various tissues, as illustrated in this cross section of a small artery (arteriole). Blood flows through the vessel lumen, which is lined by a thin sheet of endothelial cells forming the endothelium and by the underlying basal lamina. This tissue adheres to the overlying layer of smooth muscle tissue; contraction of the muscle layer controls blood flow through the vessel. A fibrillar layer of connective tissue surrounds the vessel and connects it to other tissues.

This hierarchy of tissues is copied in other blood vessels, which differ mainly in the thickness of the layers. The wall of a major artery must withstand much stress and is therefore thicker than that of a minor vessel. The strategy of grouping and layering different

tissues is used to build other complex organs as well. In each case, the function of the organ is determined by the specific functions of its component tissues, and each type of cell in a tissue produces the specific groups of proteins that enable the tissue to carry out its functions.

Genomics Has Revealed Important Aspects of Metazoan Evolution and Cell Function

Metazoans — be they invertebrates such as the fruit fly *D. melanogaster* and the roundworm *C. elegans*, or vertebrates such as mice and humans — contain between 13,000 and 26,000 protein-coding genes, about two to four times as many as a yeast (see Table 1-2). Sequencing of entire genomes has shown that many of these genes are conserved among the metazoans, and genetic studies have shown that many of them are essential for the formation and function of specific tissues and organs. Thus many of the organisms listed in Table 1-2 are used to study the roles of these conserved proteins in cell development and function.

The human and mouse genomes encode about the same number of proteins as those of the roundworm *C. elegans*, frogs, and fish, yet mammalian cells contain about 30 times the DNA of a roundworm and two to three times the DNA of frogs and fish. Only about 10

percent of human DNA encodes proteins. We now know that much of the remaining 90 percent has important functions. As detailed in Chapter 8, many DNA segments, termed *enhancers*, bind proteins that regulate expression of nearby genes. The expression of many genes is controlled by large numbers of such regulatory sequences; control by multiple enhancers allows each mammalian gene to make the precise amount of mRNA and protein needed in each of many different types of cells.

Other segments of DNA are used to synthesize the thousands of RNA molecules that do not direct the synthesis of proteins and whose function in regulating gene expression is only now being uncovered. As an example, hundreds of different kinds of micro-RNAs, 20 to 25 nucleotides long, are abundant in metazoan cells. Each microRNA binds to and represses the activity of a specific set of target mRNAs, either by inhibiting the ability of these mRNAs to be translated into proteins or by triggering their degradation (see Chapter 9). By this means, each of these small RNAs can indirectly regulate the activity of many genes.

Some of this nonprotein-coding DNA probably regulates expression of genes that make us uniquely human. Indeed, fish and humans have about the same number of protein-coding genes — about 20,000 — yet as noted above, the human genome is over twice the size of that in fish (see Table 1-2). The human brain can perform complex mental processes such as reading and writing a textbook.

Somehow these 20,000 human genes are exquisitely regulated such that humans produce a brain with about 100,000,000,000 neurons, which communicate with one another at about 100,000,000,000,000 interaction sites termed *synapses*.

Genomics — the study of the entire DNA sequences of organisms — has shown us how close humans really are to Neanderthals and to our nearest nonhominid relatives, the great apes (Figure 1-27). Human DNA is 99 percent identical in sequence to that of chimpanzees and bonobos; the 1 percent difference is about 3,000,000 base pairs, but it somehow explains the obvious differences between our species, such as the evolution of human brains during the past ~5,000,000 years since we last shared a common ancestor.

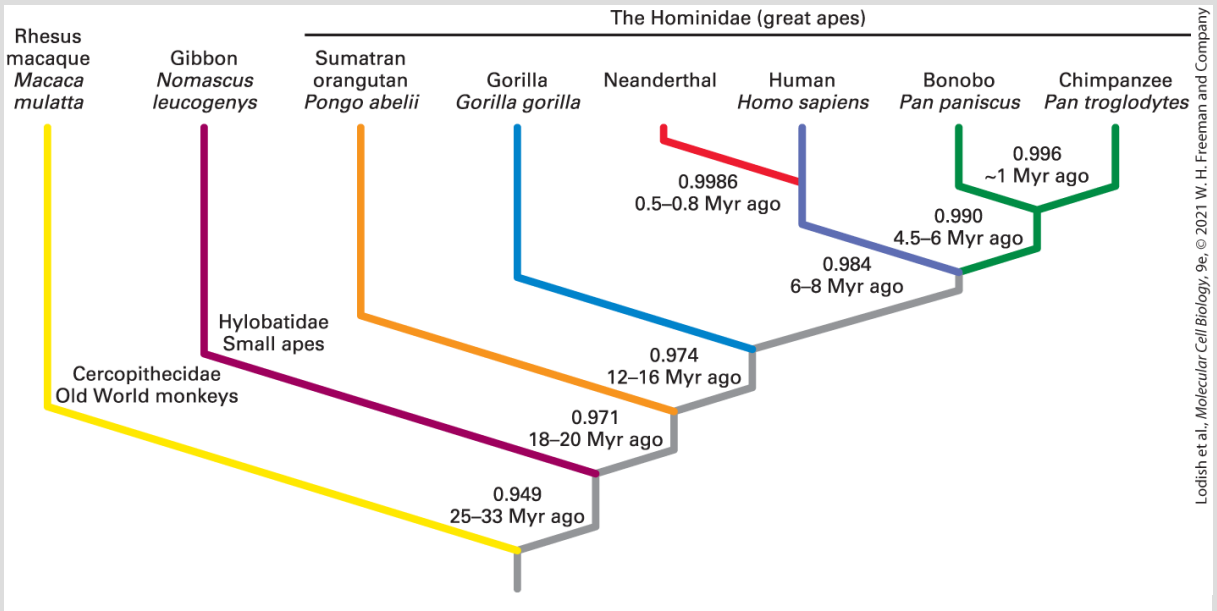


FIGURE 1-27 Evolutionary tree connecting monkeys, apes, and humans. This evolutionary tree of humans, Neanderthals, great apes, a small ape, and an Old World monkey was

estimated from the divergence among their genomic DNA sequences. Whole-genome DNA sequences were aligned, and the average nucleotide divergence in unique DNA sequences was estimated. Estimates of the times the different species diverged from each other, indicated at each node, were calculated in millions of years (Myr) based on DNA sequence identity; $\sim 1\ \mathrm{Myr}$ implies approximately $1\ \mathrm{Myr}$ or less.

[Data from D. P. Locke et al., 2011, *Nature* **469**:529–533 and K. Prüfer et al., 2014, *Nature* **505**:43.]

Genomics coupled with paleontological findings indicates that humans and mice descended from a common mammalian ancestor that probably lived about 75 million years ago. Nonetheless, both organisms contain about the same number of genes, and about 99 percent of mouse protein-coding genes have homologs in humans, and vice versa. Over 90 percent of mouse and human genomes can be partitioned into regions of **synteny** — that is, DNA segments that have the same order of unique DNA sequences and genes along a segment of a chromosome. This observation suggests that much of the gene order in the most recent common ancestor of humans and mice has been conserved in both species (Figure 1-28). Of course, mice are not people; relative to humans, mice have expanded families of genes related to immunity, reproduction, and olfaction, probably reflecting the differences between the human and mouse lifestyles.

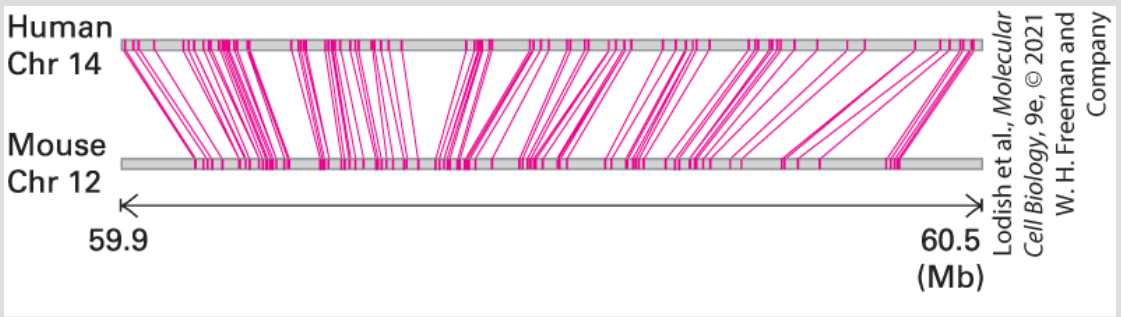


FIGURE 1-28 Conservation of synteny between human and mouse. Shown is a 510,000-base-pair (bp) segment of mouse chromosome 12 that shares common ancestry with a 600,000-bp section of human chromosome 14. Pink lines connect the conserved unique DNA sequences in the two genomes. Mb, 1 million base pairs.

[Data from Mouse Genome Sequencing Consortium, 2002, *Nature* **420**:520.]

It's not only human evolution that interests us! Polar bears live in the Arctic and eat a high-fat diet, mostly composed of seals. Recent genome sequencing allowed researchers to conclude that the most recent common ancestor of polar bears and their brown bear relatives, which live in temperate climates, was present about 500,000 years — or only about 20,000 bear generations — ago. During that rather short evolutionary period the polar bear genome acquired changes in many genes regulating cardiovascular function, fat metabolism, and heart development, allowing it to consume a diet very rich in fats.

Development Uses a Conserved Set of Master Transcription Factors and Involves Epigenetic Modifications to

DNA and Its Associated Histone Proteins

The astute reader will note a paradox in the previous discussion: if indeed most human protein-coding genes are shared with apes and mice, and many with flies and worms, how is it that these organisms look and function so differently? The answer to this question resides in the way genes are regulated during the development of all metazoans from a single cell, the fertilized egg. As we learn in Chapters 7 and 8, each protein-coding gene is associated with regulatory DNA sequences that differ in different organisms. Many of these regulatory sequences bind proteins that direct the expression of the gene, and thus the amount of a protein it makes, in specific types of cells. Some of these proteins are termed *master transcription factors*; these proteins bind to regulatory DNA sequences, are conserved throughout evolution, and control the development of specific types of cells by activating or repressing groups of genes, often at different or successive stages of development.

During development, many segments of each chromosome undergo enzyme-catalyzed chemical modifications — termed **epigenetic modifications** — that affect the ability of genes within these segments to be transcribed into RNA. As one example, addition of methyl (CH_3) groups to multiple cytosine bases in a segment of DNA triggers the formation of heterochromatin — condensed,

concentrated DNA that cannot be transcribed into RNA. Histone proteins bound to DNA can undergo multiple kinds of enzyme-catalyzed modifications, some enhancing and others inhibiting the ability of RNA polymerase to transcribe the adjacent gene into RNA. As detailed in Chapter 8, many of these modifications are replicated and passed on to daughter cells, a type of cell heredity that does not involve changes in the primary nucleotide sequence of the DNA genome.

The early stages in the development of a human embryo are similar to those in the mouse. They are characterized by rapid cell divisions (Figure 1-29) followed by the differentiation of cells into tissues. In all organisms, the embryonic body plan — the spatial pattern of cell types (tissues) and body parts — emerges from two influences: a program of genes that specifies the pattern of the body, and local cell interactions that induce different parts of the program.

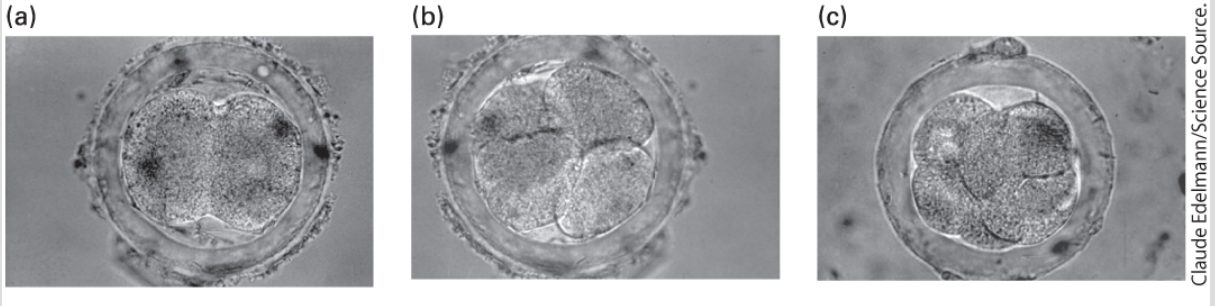


FIGURE 1-29 The first few cell divisions of a fertilized egg set the stage for all subsequent development. A developing mouse embryo is shown at the (a) two-cell, (b) four-cell, and (c) eight-cell stages. The embryo is surrounded by supporting membranes. The corresponding steps in human development occur during the first few days after fertilization.

With only a few exceptions, animals display axial symmetry; that is, their left and right sides mirror each other. This most basic of patterns is encoded in the genome. Developmental biologists have divided bilaterally symmetric animal phyla into two large groups depending on where the mouth and anus form in the early embryo. **Protostomes** develop a mouth close to a transient opening in the early embryo (the **blastopore**) and have a ventral nerve cord; protostomes include all worms, insects, and mollusks. **Deuterostomes** develop an anus close to this transient opening in the embryo and have a dorsal central nervous system; they include echinoderms (such as sea stars and sea urchins) and vertebrates. The bodies of both protostomes and deuterostomes are divided into discrete segments that form early in embryonic development. Protostomes and deuterostomes probably evolved from a common ancestor, termed *Urbilateria*, that lived approximately 600 million years ago (Figure 1-30a).

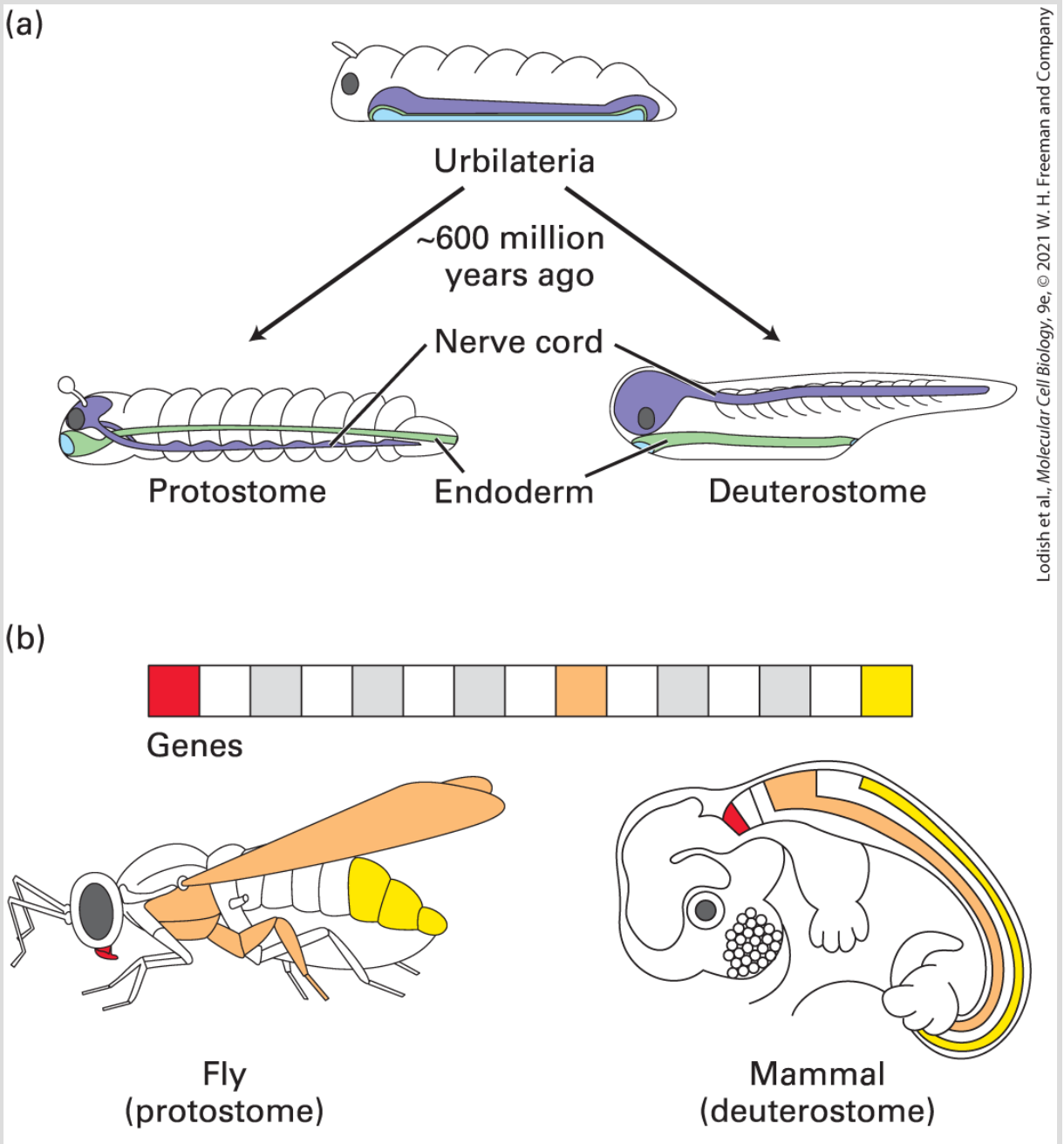


FIGURE 1-30 Similar master transcription factors, conserved during evolution, regulate early developmental processes in diverse animals. (a) Urbilateria is the presumed ancestor of all protostomes and deuterostomes that existed about 600 million years ago. The positions of its nerve cord (violet), surface ectoderm (mainly skin; white), and endoderm (mainly digestive tract and internal organs; light green) are shown. (b) Highly conserved master transcription factors called Hox proteins, which determine the identity of body segments during embryonic development, are found in both protostomes and deuterostomes. Hox genes are found in clusters on the chromosomes of most or all animals,

and they encode related master transcription factors that control the activities of other genes. In many animals, different Hox genes direct the development of different segments along the head-to-tail axis, as indicated by corresponding colors. Each Hox gene is activated (transcriptionally) in a specific region along the head-to-tail axis and controls the growth and development of tissues in that segment. For example, in the mouse, a deuterostome, certain Hox genes are responsible for the distinctive shapes of vertebrae. Mutations affecting Hox genes in the fruit fly, a protostome, cause body parts to form in the wrong locations, such as legs in lieu of antennae on the head. In both organisms, these genes provide a head-to-tail address and serve to direct the formation of structures in the appropriate places.

Many **patterning genes** encode master transcription factors that control expression of other genes and specify the general organization of an organism, beginning with the major body axes — anterior-posterior (head-to-tail), dorsal-ventral (back-to-belly), and left-right — and ending with body segments such as the head, chest, abdomen, and tail. The conservation of axial symmetry from the simplest worms to mammals is explained by the presence of conserved patterning genes in their genomes. Other patterning genes encode proteins that are important in cell adhesion or in cell signaling. This broad repertoire of patterning genes permits the integration and coordination of events in different parts of the developing embryo and gives each segment in the body its unique identity.

Remarkably, many patterning genes encoding master transcription factors are highly conserved in both protostomes and deuterostomes (Figure 1-30b). This conservation of body plan reflects evolutionary pressure to preserve the commonalities in the

molecular and cellular mechanisms controlling development in different organisms.

1.6 Metazoan Organisms Widely Used in Cell Biology Research

Many invertebrate and vertebrate organisms have been and continue to be immensely valuable for cell biology research. Invertebrate organisms with large-celled embryos that develop outside the mother's body (e.g., sea urchins and sea stars) are extremely useful for tracing the fates of cells as they form different tissues, as well as for making extracts for biochemical studies. For instance, a key protein in regulating cell division in all eukaryotes, including humans, was first identified in studies with sea star and sea urchin embryos and subsequently purified from extracts prepared from these embryos (see Chapter 19). Other organisms are investigated because of their adaptation to specific ecological niches. An example is the tardigrades, tiny 0.5-mm-long, 8-legged invertebrates (see the photo on the first page of the chapter). Tardigrades are widely studied because of their ability to survive exposures to extreme temperatures, high and low pressures, dehydration, and starvation. Indeed, they are the only organisms known to survive exposure to the deadly combination of low pressure and intense radiation of outer space.

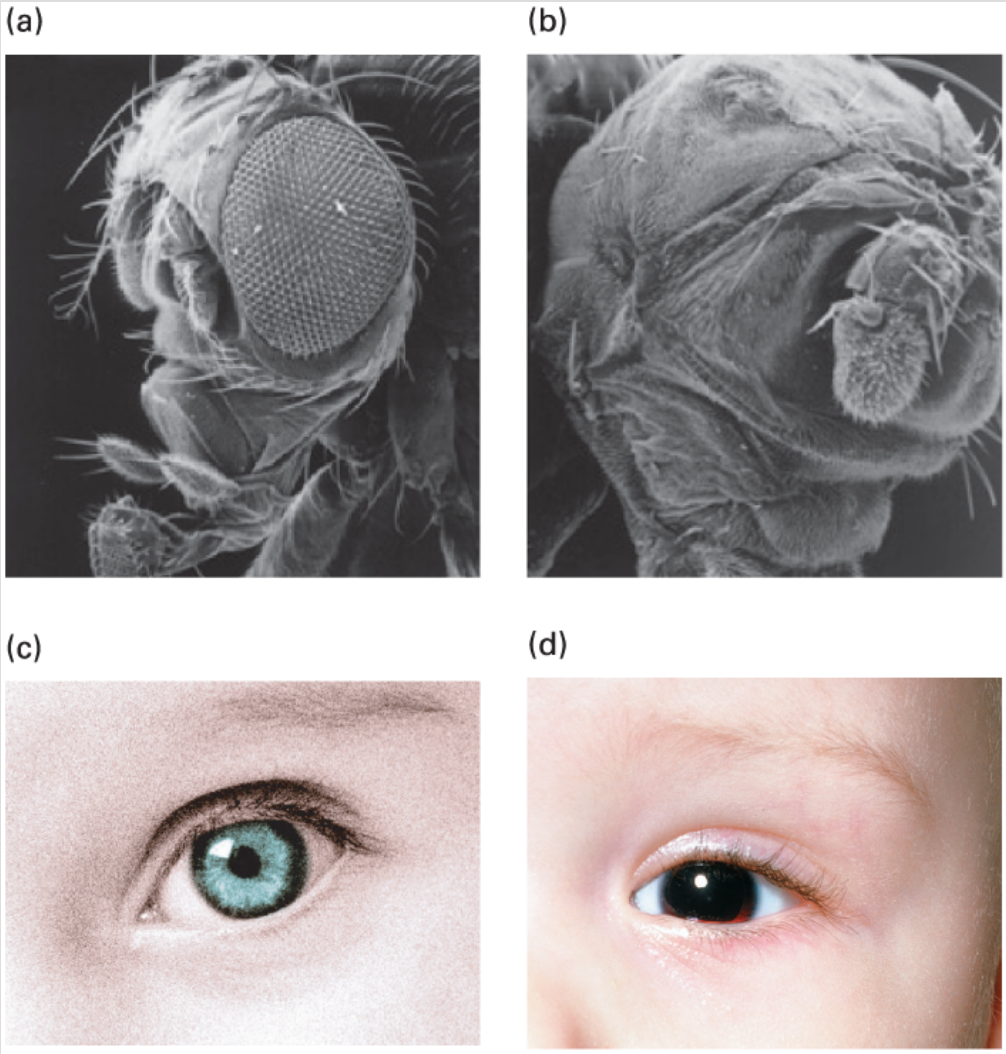
A few invertebrates — notably the roundworm *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and the planarian *Schmidtea mediterranea* — are especially amenable to a variety of

experimental approaches. These invertebrates have been extensively studied by thousands of research groups using multiple experimental technologies. Studies on these and other invertebrates led to important insights into animal development and cell function, and also to understanding of many human diseases.

Drosophila melanogaster and *Caenorhabditis elegans* Are Used to Identify Genes That Regulate Animal Development

D. melanogaster has been used as an experimental organism since 1910, when Thomas Hunt Morgan established the Fly Room laboratory at Columbia University. It is still extensively used in research because it can be readily reared in the laboratory, has only four pairs of chromosomes, breeds quickly, and lays many eggs (see Figure 1-23d). Each step in the development of this organism — from fertilized egg through the embryonic larval and pupal stages to the mature adult — is well understood. Most importantly, mutations in many genes important for development have been characterized in detail and these have provided important information about the functions of the encoded proteins. As an example, many genes important for development of the *Drosophila* eye have been identified; one of them, *eyeless*, is the master transcription factor that initiates eye development. Fly eyes and human eyes are very

different in their structure, function, and nerve connections. Nonetheless, *eyeless* in the fly and *Pax6* in the human are highly related proteins that regulate the activities of other genes and are descended from the same ancestral gene. Mutations in both the *eyeless* and the *Pax6* genes cause major defects in eye formation (Figure 1-31).



Parts (a) and (b) Courtesy Andreas Hefti, Interdepartmental Electron Microscopy (IEM), Biocenter of the University of Basel. Part (c) © Simon Fraser/Science Source. Part (d) © Mediscan/Alamy.

FIGURE 1-31 Homologous genes regulate eye development in *Drosophila* and humans. (a) Development of the large compound eyes in fruit flies requires a gene called *eyeless* (named for the mutant phenotype). (b) Flies with inactivated *eyeless*

genes lack eyes. (c) Normal human eyes require the gene *Pax6*, the homolog of *eyeless*. (d) People lacking adequate *Pax6* function have the genetic disease *aniridia*, a lack of irises in the eyes. *Eyeless* and *Pax6*, which encode highly related master transcription factors that regulate the activities of other genes, are homologs and presumably descended from the same ancestral gene.

The roundworm *Caenorhabditis elegans* was developed as an experimental organism by Sydney Brenner only in the 1960s. The adult worm (see Figure 1-23c) — about 1 mm long and 70 μm in diameter — can be either a hermaphrodite (a worm with both male and female organs that can fertilize itself and produce viable offspring) or a male (which can fertilize a hermaphrodite); this unusual reproductive behavior has many advantages for the rapid isolation and characterization of mutations affecting development or cell function. The hermaphrodite has precisely 959 somatic (non-germ) cells, whereas the male has 1031 (see Figure 22-26d). These small numbers of cells are arranged in a nearly identical way in every worm, and scientists have traced the lineage of each somatic cell in *C. elegans* from the fertilized egg to the mature worm using microscopy to follow the development of live worms (see Figure 22-26c). As we learn in Chapter 22, many of the cells born during development of *C. elegans* are destined to die by a process termed **programmed cell death**, or **apoptosis**. A similar process utilizing similar proteins occurs in many human cells, and the absence of programmed cell death contributes to the development of many human cancers.

Planaria Are Used to Study Stem Cells and Tissue Regeneration

In single-celled organisms, both daughter cells usually (though not always) resemble the parent cell. Similarly, in multicellular organisms, when many types of cells divide, the daughter cells look a lot like the parent cell — liver cells, for instance, divide to generate liver cells with the same characteristics and functions as their parent, as do insulin-producing cells in the pancreas. In contrast, **stem cells** and certain other undifferentiated cells can generate multiple types of differentiated descendant cells; these cells often divide in such a way that the two daughter cells are different. Such **asymmetric cell division** is characteristic of stem cells and is critical to the generation of different cell types in the body (see Chapter 22). Often one daughter cell resembles its parent in that it remains undifferentiated and retains its ability to give rise to multiple types of differentiated cells. The other daughter cell divides many times, and each of its daughter cells differentiates into one or several specific types of cells.

The planarian *Schmidtea mediterranea* is best known for its capacity to regenerate complete individuals — with a normal head — from minuscule body parts formed by dissection (see Figure 1-23e). Planaria contain stem cells, termed *neoblasts*, that replace cells lost to normal turnover. In portions of a dissected animal, neoblasts will, after several cell divisions, generate any cell type needed during

regeneration. These stem cells have served as a potent experimental system to discover how heads and tails, each built of many types of cells, are formed (see Chapters 16 and 22). The hormones that instruct stem cells in different parts of the body to generate specific types of cells are similar to those used in mammals, including humans, in development (see Chapter 16), and thus future studies on planarian regeneration may inform scientists how to regenerate human body parts such as a hand or an eye.

Studies on Fish, Mice, and Other Vertebrate Organisms Inform the Study of Human Development and Disease

The zebrafish *Danio rerio* (see Figure 1-23f) is widely used to study the development and functions of organs not found in invertebrates, especially the brain and the central nervous system. Its generation time — from fertilized egg to daughter fertilized eggs — is only 3 months, short for a vertebrate. The fertilized egg rapidly becomes transparent, allowing visualization of developing organs, and the embryo develops rapidly, such that precursors to all major organs appear within 36 hours of fertilization.

Mice (see Figure 1-23g) have one enormous advantage over other experimental organisms: they are the most closely related to

humans of any animal for which powerful genetic approaches have been available for many years. Mice and humans have shared living structures for millennia, have similar nervous systems, have similar immune systems, and are subject to infection by many of the same pathogens. As noted, both organisms contain about the same number of genes, and about 99 percent of mouse protein-coding genes have homologs in the human genome, and vice versa.

Using the Cas9 and other experimental systems described in Chapter 6, researchers can inactivate any desired gene and thus abolish production of its encoded protein. This approach is being used extensively to produce animal versions of human genetic diseases, in mice as well as in other animals. As an example, people with autism spectrum disorder often have mutations in specific protein-coding genes that regulate development of specific types of nerve cells. To understand the role of these mutations, these genes have been inactivated in mice and more recently in macaque monkeys. These animals often exhibit symptoms of the human disease, including repetitive actions such as excessive grooming, strongly suggesting that the human mutation indeed has a role in triggering the disorder. Such approaches can be useful in uncovering the role of specific genes in higher order brain tasks such as learning and memory, or in studies of viruses that infect only humans and nonhuman primates. Once animal models of a human disease are available, investigators can perform further studies on the molecular defects causing the disease and can test

new treatments, thereby minimizing the testing of new drugs on humans.

Human Genetic Diseases Elucidate Important Aspects of Cell Function

We have emphasized how analysis of mutant organisms missing a specific protein has informed the normal function of that gene and encoded protein. Although we cannot induce mutations in humans as we might in other organisms, we can study naturally occurring mutations in humans. Many genetic diseases are caused by mutations in a single protein; thus studies on humans with these diseases also can shed light on the normal function of those proteins. As an example, consider Duchenne muscular dystrophy (DMD), the most common of the hereditary muscle-wasting diseases, collectively called *muscular dystrophies*. DMD is a disorder caused by mutations in the dystrophin gene localized on the X chromosome. It affects 1 in 3300 boys and results in cardiac or respiratory failure and death, usually in the late teens or early twenties.

The first clue to understanding the molecular basis of this disease came from the cloning of the dystrophin gene. As detailed in Chapters 17 and 20, the very large protein encoded by this gene binds to other proteins in muscle cells to form a large multiprotein

complex. One end of the long dystrophin molecule binds to actin microfilaments that are a component of the contractile apparatus (see Figure 1-14); the other binds to a complex of proteins in the muscle plasma membrane (see Figure 20-41). The resulting large multiprotein assemblage, the dystrophin glycoprotein complex (DGC), links the extracellular matrix protein laminin to the cytoskeleton within muscle cells and prevents the muscle cell from ripping apart during contractions. Mutations in dystrophin or other DGC components can disrupt the DGC-mediated link between the exterior and interior of muscle cells and cause muscle weakness and eventual muscle deterioration and death. The first step in identifying the entire dystrophin glycoprotein complex was to clone the dystrophin-encoding gene using DNA from normal individuals and from patients with Duchenne.

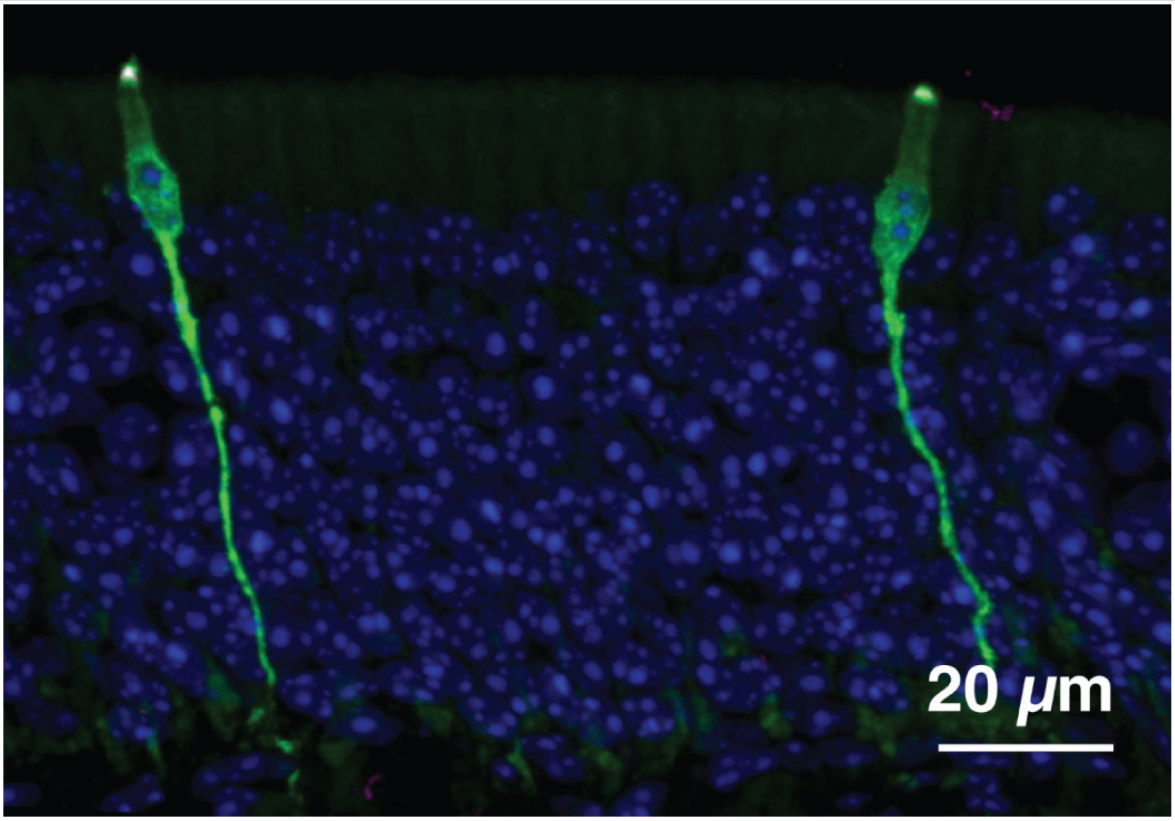
Unbiased Single Cell Sequencing Experiments Identify Altogether New Cell Types

How many different types of cells make up a mouse or a human body? Certain types of cells may carry out very important functions, but because they are present in miniscule amounts in a specific organ or tissue their very existence may have gone unnoticed!

The new technique of single cell RNA sequencing (Chapter 6) allows researchers to identify cell types present in only small amounts. The technique identifies the ~5000 types of messenger RNAs that are expressed in the largest numbers in a single isolated cell. If researchers apply the technique to many individual cells isolated from a tissue, they can identify new types of cells whose components of messenger RNAs, and thus of the proteins they encode, are different from those of known types of cells.

In a recent study using single-cell RNA sequencing, researchers analyzed tens of thousands of cells from mouse and human airways, mapping the physical locations of cell types and creating a cellular atlas of the tissue. Using a computer program to analyze the overall similarity between the types of mRNAs expressed in each cell, the researchers easily found six groups of cells that each expressed a similar population of mRNAs. These six groups of cells corresponded to the six well-known cell types present in the trachea — the abundant basal, club, and ciliated cells and the rarer tuft, neuroendocrine, and goblet cells. One extremely rare cell type, making up roughly 1 percent of the cell population, had a profile of mRNAs radically different from the 6 other known cells. The team named this new cell the *pulmonary ionocyte*. Strikingly, pulmonary ionocytes were found to express the *CFTR* gene at levels higher than any other cell type. The CFTR protein is critical for airway function; mutations in the *CFTR* gene cause cystic fibrosis in humans. For decades, researchers and clinicians assumed that the *CFTR* gene is

frequently expressed at low levels in ciliated cells, a cell type spread throughout the entire airway. According to the new data, the majority of CFTR is expressed in only a few cells (Figure 1-32), which researchers didn't even know existed until now.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 1-32 A new cell type in the mouse airway identified by unbiased single-cell sequencing of messenger RNAs. Ionocytes, a minor cell type in the upper respiratory tract of the mouse that express the CFTR protein, are stained green with an antibody that detects a protein expressed only in these cells. All cell nuclei are stained with a blue dye. By counting the number of blue nuclei in the tissue section, it can be determined that ionocytes comprise about 1 percent of the cells lining the mouse trachea (airway). These cells extend through the layer of neighboring epithelial cells to the airway surface. See text for an explanation of how these cells were identified.

[Republished with permission of Nature Publishing Group, from D. Montoro et. al., 2018. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* **560**:319;

permission conveyed through Copyright Clearance Center, Inc.]

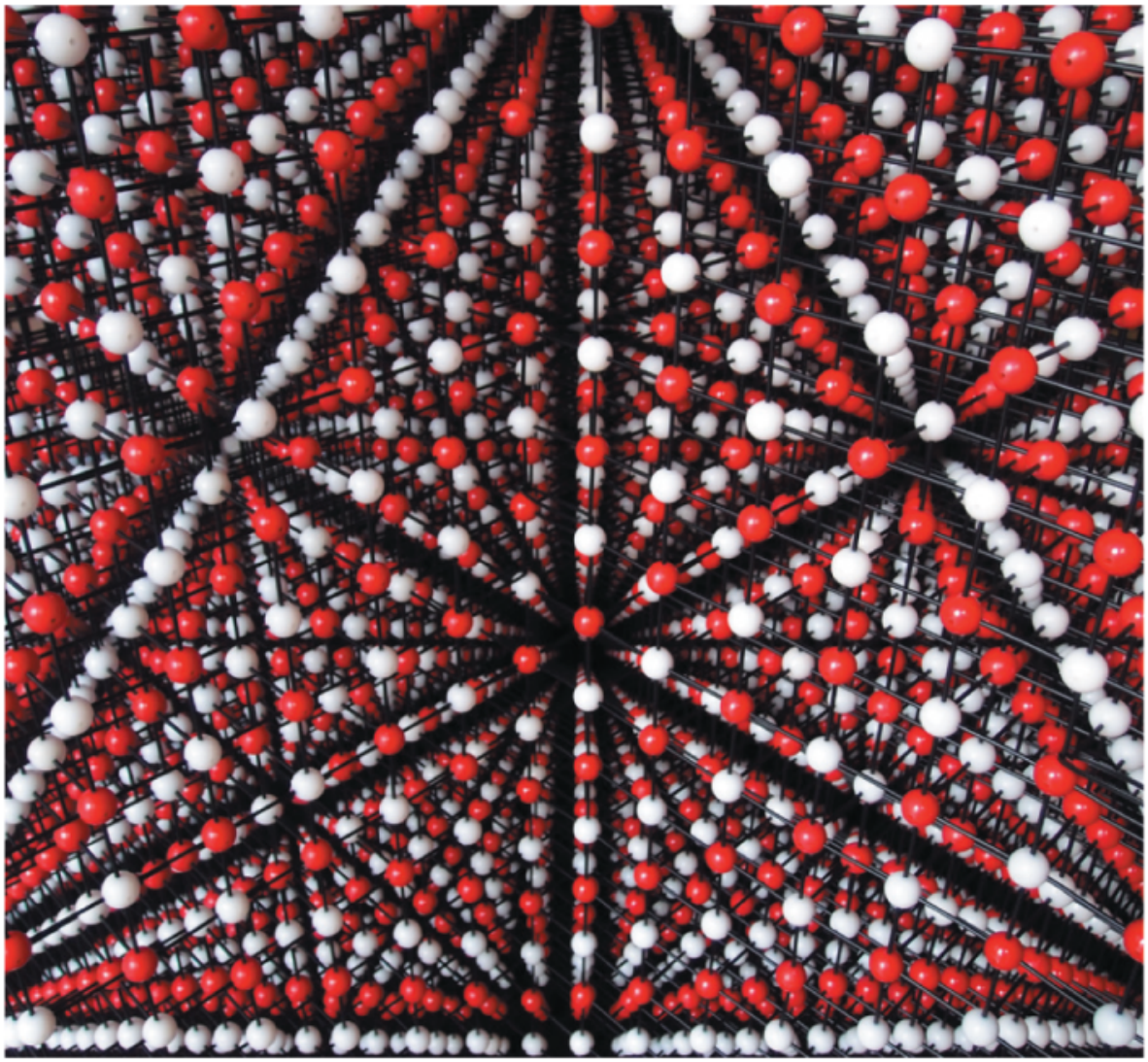
The Following Chapters Present Many Experimental Techniques and Much Experimental Data That Explains How We Know What We Know About Cell Structure and Function

In subsequent chapters of this book, we discuss cellular processes in much greater detail. We begin (in Chapter 2) with a discussion of the chemical nature of the building blocks of cells and the basic chemical processes required to understand the macromolecular processes discussed in subsequent chapters. We go on to discuss the structure and function of proteins (in Chapter 3). Chapter 4 discusses many of the techniques biologists use to culture and fractionate cells and to visualize specific proteins and structures within cells. Chapter 5 describes how DNA is replicated, how segments of DNA are copied into RNA, and how proteins are synthesized on ribosomes. Chapter 6 describes many of the techniques used to study genes, gene expression, and protein function, including the generation of animals with specific genetic mutations. Gene and chromosome structure and the regulation of gene expression are covered in Chapters 7, 8, and 9. Biomembrane structure is the topic of Chapter 10. The transport of ions and small

molecules across membranes is covered in Chapter 11, and Chapter 12 discusses cellular energetics and the functions of mitochondria and chloroplasts. Membrane biogenesis, protein secretion, and protein trafficking — the directing of proteins to their correct subcellular destinations — are the topics of Chapters 13 and 14. Chapters 15 and 16 discuss the many types of signals and signal receptors used by cells to communicate and regulate their activities. The cytoskeleton and cell movements are discussed in Chapters 17 and 18. Chapter 19 discusses the cell cycle and how cell division is regulated. The interactions among cells, and between cells and the extracellular matrix, that enable formation of tissues and organs are detailed in Chapter 20. A chapter new to this edition, Chapter 21, details how cells respond to changes in their environment — specifically, to changes in oxygen levels, temperature, day and night, and nutrient levels. Later chapters of the book discuss important types of specialized cells — stem cells (Chapter 22), neurons (Chapter 23), and cells of the immune system (Chapter 24). Chapter 25 discusses cancer and the multiple ways in which cell growth and differentiation can be altered by mutations.

CHAPTER 2

Chemical Foundations



Dr. Robert Krickl

Model of sodium chloride (NaCl) crystal. This photograph shows the largest (3.1 m tall) ball-and-stick physical model of a crystalline salt ever constructed, built from 40,110 balls by Robert Krickl.

OUTLINE

2.1 Covalent Bonds and Noncovalent Interactions
2.2 Chemical Building Blocks of Cells
2.3 Chemical Reactions and Chemical Equilibrium
2.4 Biochemical Energetics

The life of a cell depends on thousands of chemical interactions and reactions exquisitely coordinated with one another in time and space, influenced by the cell's genetic instructions and its environment. By understanding these interactions and reactions at a molecular level, we can begin to answer fundamental questions about cellular life: How does a cell extract nutrients and information from its environment? How does a cell convert the energy stored in nutrients into the work of movement or metabolism? How does a cell transform nutrients into the cellular components required for its survival? How does a cell transmit its genetic information to its progeny and use this information to survive and thrive? How does a cell link itself to other cells to form a tissue? How do cells communicate with one another so that a complex, efficiently functioning organism can develop and thrive? One of the goals of this textbook is to answer these and other questions about the structure and function of cells and organisms in terms of the properties of individual molecules and ions.

For example, the properties of one such molecule, water, control the evolution, structure, and function of all cells. An understanding of

biology is not possible without appreciating how the properties of water control the chemistry of life. Life first arose in a watery environment, and water constitutes 70–80 percent of most cells by weight. In fact, water is the most abundant molecule in biological systems. The remaining mass of organisms is made up of small molecules and ions and of combinations of these that form the larger macromolecules and macromolecular assemblies that make up a cell's machinery and architecture. The small molecules typically have molecular weights of less than 1000 daltons and include amino acids (the building blocks of proteins), nucleotides (the building blocks of DNA and RNA), lipids (the building blocks of biomembranes), and sugars (the building blocks of complex carbohydrates). The small molecules and ions together make up about 7 percent of the weight of living matter.

Many of the cell's biomolecules (such as sugars) readily dissolve in water; these molecules are referred to as **hydrophilic** ("water liking"). Others (such as cholesterol) are oily, fatlike substances that shun water; these molecules are said to be **hydrophobic** ("water fearing"). Still other biomolecules (such as phospholipids) contain both hydrophilic and hydrophobic regions; these molecules are said to be **amphipathic or amphiphilic** ("both liking"). The smooth functioning of cells, tissues, and organisms depends on all these molecules, from the smallest to the largest. Indeed, the chemistry of the simple proton (H^+) can be as important to the survival of a human cell as that of each gigantic DNA molecule (the mass of the

DNA molecule in human chromosome 1 is 8.6×10^{10} times that of a proton!). The chemical interactions of all these molecules, large and small, with water and with one another define the nature of life.

Luckily, although many types of biomolecules interact and react in numerous and complex pathways to form functional cells and organisms, a relatively small number of chemical principles are necessary to understand cellular processes at the molecular level (Figure 2-1).

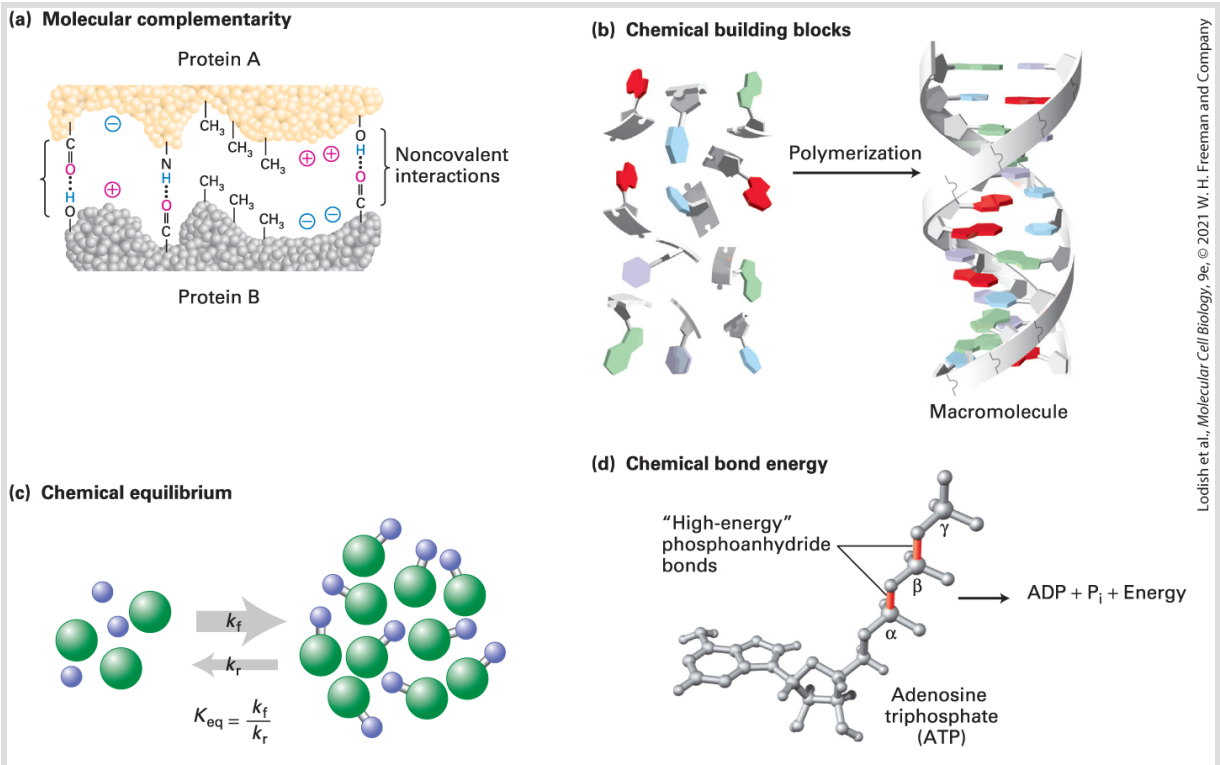


FIGURE 2-1 Chemistry of life: four key concepts. (a) Molecular complementarity lies at the heart of all biomolecular interactions (see Section 2.1), as when two proteins with complementary shapes and chemical properties come together to form a tightly bound complex. (b) Small molecules serve as building blocks for larger structures (see Section 2.2). For example, to generate the information-carrying macromolecule DNA, four small nucleotide building blocks are covalently linked into long strings (polymers), which then

wrap around each other to form the double helix. (c) Chemical reactions are reversible, and the distribution of the chemicals between starting reactants (*left*) and the products of the reactions (*right*) depends on the rate constants of the forward (k_{f}, upper arrow) and reverse (k_{r}, lower arrow) reactions. The ratio of these, K_{eq}, provides an informative measure of the relative amounts of products and reactants that will be present at equilibrium (see Section 2.3). (d) In many cases, the source of energy for chemical reactions in cells is the hydrolysis of the molecule ATP (see Section 2.4). This energy is released when a high-energy phosphoanhydride bond linking the β and γ phosphates in the ATP molecule (red) is broken by the addition of a water molecule, forming ADP and $\mathrm{P_i}$.

First, biological molecules are able to bind together where their surfaces have complementary shapes and chemical properties — negatively charged surfaces with positively charged surfaces, for example (Figure 2-1a). Second, as we have seen, chemical building blocks assemble to form macromolecules (Figure 2-1b). Third, these assemblages are possible because molecules are able to react to form new products in chemical reactions, which can proceed to chemical equilibrium (Figure 2-1c) or steady state. Finally, molecules release the energy stored in chemical bonds in a controlled fashion to power the work of the cell (Figure 2-1d). In this chapter, we review these key principles, some of which you already know well, because a little bit of chemistry can go a very long way in helping us to understand molecular cell biology. We begin with the covalent bonds that connect atoms into molecules and the noncovalent interactions that stabilize groups of atoms within and between molecules. We then consider the basic chemical building blocks of macromolecules and macromolecular assemblies. After reviewing those aspects of chemical equilibrium that are most

relevant to biological systems, we end the chapter with the basic principles of biochemical energetics, including the central role of ATP (adenosine triphosphate) in capturing and transferring energy in cellular metabolism.

2.1 Covalent Bonds and Noncovalent Interactions

Underpinning all of biology is the assembly of atoms into molecules that in turn interact with one another. Strong and weak attractive forces between atoms are the glue that holds individual molecules together and permits interactions between different molecules. When two atoms share a single pair of electrons, the result is a **covalent bond** — a type of strong force that links atoms together in molecules. When two atoms share multiple pairs of electrons, the result is multiple covalent bonds (e.g., double or triple bonds). The weak attractive forces of **noncovalent interactions** are equally important in determining the properties and functions of biomolecules such as proteins, nucleic acids, carbohydrates, and lipids. In this section, we first review covalent bonds and then discuss the four major types of noncovalent interactions: ionic bonds, hydrogen bonds, van der Waals interactions, and the hydrophobic effect.

The Electronic Structure of an Atom Determines the Number and Geometry of the Covalent Bonds It Can Make

Hydrogen, oxygen, carbon, nitrogen, phosphorus, and sulfur are the most abundant elements in biological molecules. These atoms, which rarely exist as isolated entities, readily form covalent bonds, sharing electrons in the outermost electron orbitals surrounding their nuclei (Figure 2-2). As a rule, each type of atom forms a characteristic number of covalent bonds with other atoms. These bonds have a well-defined geometry determined by the atom's size and by both the distribution of electrons around the nucleus and the number of electrons it can share. In some cases, the number of stable covalent bonds an atom can make is fixed; carbon, for example, always forms four covalent bonds. In other cases, the number of stable covalent bonds can vary; for example, sulfur can form two, four, or six stable covalent bonds.

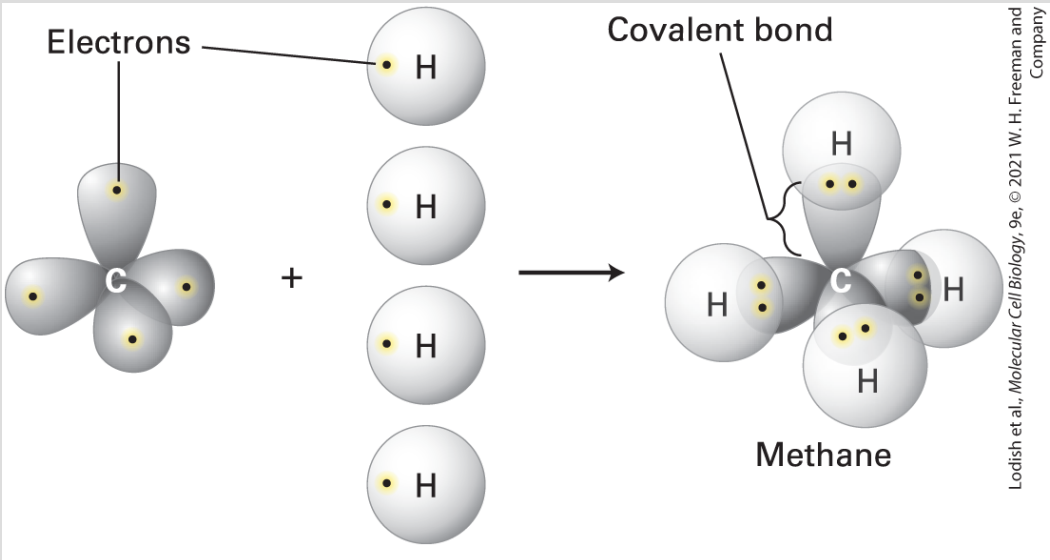


FIGURE 2-2 Covalent bonds form by the sharing of electrons. Covalent bonds, the strong forces that hold atoms together in molecules, form when atoms share electrons from their outermost electron orbitals. Each atom forms a defined number of covalent bonds having a distinct geometry.

All the biological building blocks are organized around the carbon atom. In these organic biomolecules, each carbon usually bonds to three or four other atoms. [Carbon can also bond to two other atoms, as in the linear molecule carbon dioxide, CO_2, which has two carbon-oxygen double bonds (O=C=O); however, such bond arrangements of carbon are not found in biological building blocks.] As illustrated in Figure 2-3a for formaldehyde, carbon can bond to three atoms, all in a common plane. The carbon atom forms two single bonds with two atoms and one double bond (two shared electron pairs) with the third atom. In the absence of other constraints, atoms joined by a single bond generally can rotate freely about the bond axis, whereas those connected by a double bond cannot. The rigid planarity imposed by double bonds has enormous significance for the shapes and flexibility of biomolecules such as phospholipids, proteins, and nucleic acids.

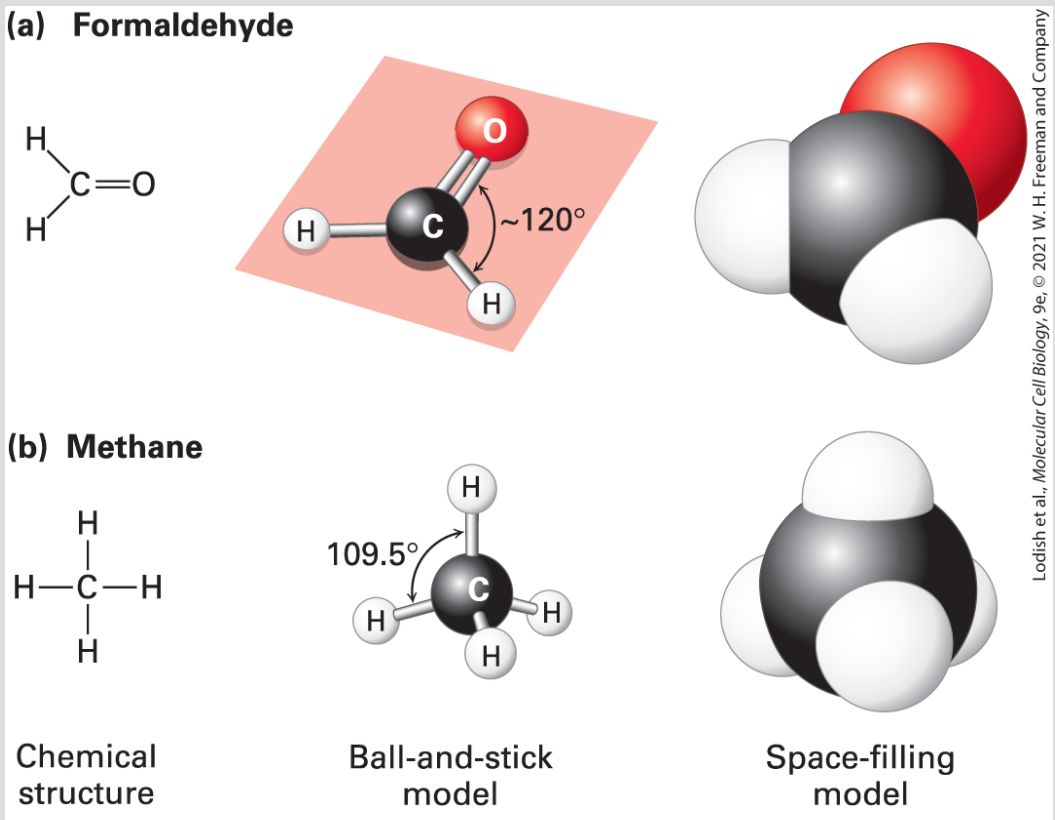


FIGURE 2-3 Geometry of bonds when carbon is covalently linked to three or four other atoms. (a) A carbon atom can be bonded to three atoms, as in formaldehyde (CH_2O). In this case, the carbon-bonding electrons participate in two single bonds and one double bond. All bonds lie in the same plane. Unlike atoms connected by a single bond, which usually can rotate freely about the bond axis, those connected by a double bond cannot. (b) When a carbon atom forms four single bonds, as in methane (CH_4), the bonded atoms (all H in this case) are oriented in space in the form of a tetrahedron. The letter representations on the left clearly indicate the atomic composition of each molecule and its bonding pattern. The ball-and-stick models in the center illustrate the geometric arrangement of the atoms and bonds, but the diameters of the balls representing the atoms and their nonbonding electrons are unrealistically small compared with the bond lengths. The sizes of the electron clouds in the space-filling models on the right more accurately represent the structure in three dimensions.

Carbon can also bond to four rather than three atoms. As illustrated by methane (CH_4), when carbon is bonded to four other atoms, the angle between any two bonds is 109.5°, and the positions of bonded atoms define the four points of a tetrahedron (Figure 2-3b). This geometry defines the structures of many biomolecules. A carbon (or any other) atom bonded to four dissimilar atoms or groups in a nonplanar configuration is said to be asymmetric. The four atoms connected to an *asymmetric carbon atom* in a tetrahedral orientation can be arranged in three-dimensional space in two different ways, producing molecules that are mirror images of each other, a property called *chirality* ("handedness," from the Greek word *cheir*, meaning "hand") (Figure 2-4). Such molecules are called *optical isomers*, or **stereoisomers**. Many molecules in cells contain at least one asymmetric carbon atom, often called a *chiral carbon* atom. The different stereoisomers of a molecule usually have completely different biological activities because the arrangement of atoms within their structures, and thus their ability to interact with other molecules, differs.

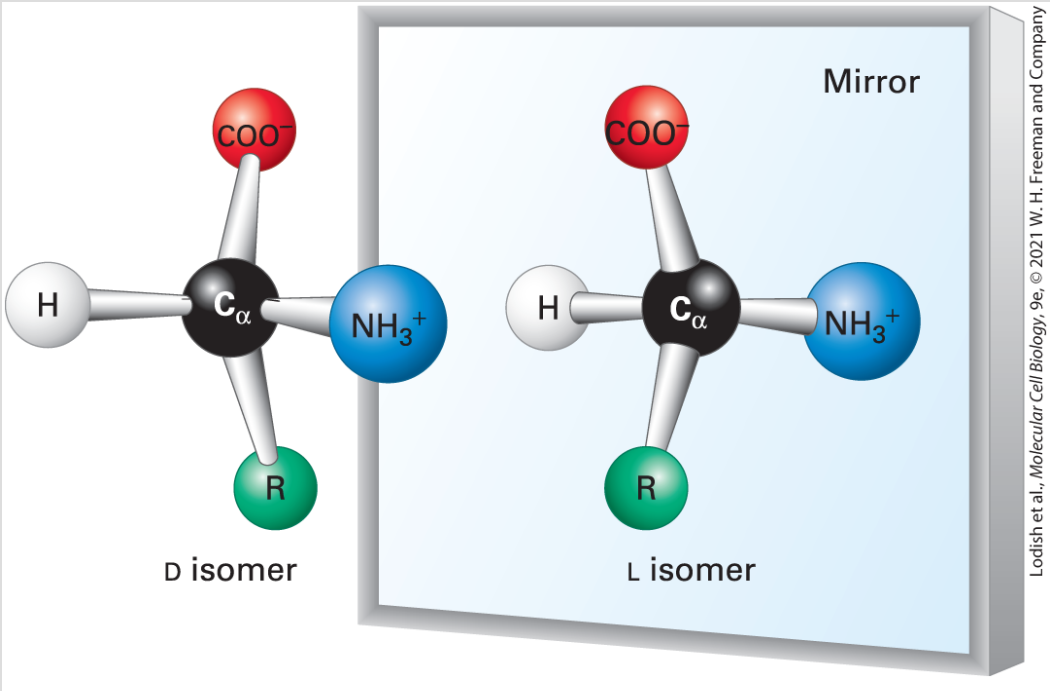


FIGURE 2-4 Stereoisomers. Many molecules in cells contain at least one asymmetric carbon atom. The tetrahedral geometry of bonds formed by an asymmetric carbon atom can be arranged in three-dimensional space in two different ways, producing molecules that are mirror images, or stereoisomers, of each other. Shown here is the common structure of an amino acid, with its central asymmetric carbon and four attached groups, including the R group, discussed in Section 2.2. Amino acids can exist in two mirror-image forms, designated L and D. Although the chemical properties of such stereoisomers are identical, their biological activities are distinct. Only L amino acids are found in proteins.

Some drugs are mixtures of the stereoisomers of a small molecule in which only one stereoisomer has the biological activity of interest, for example, inhibiting the activity of a specific protein molecule. Unfortunately, both the stereoisomer that has the biological activity of interest and the one without biological activity might have side effects that could be harmful. The use of a pure,

single stereoisomer of the chemical in place of the mixture may result in a more potent drug with fewer side effects. For example, one stereoisomer of the antidepressant drug *citalopram* (Celexa) is 170 times more potent than the other. Some stereoisomers have very different activities. Darvon is a pain reliever, whereas its stereoisomer, Novrad (*Darvon* spelled backward), is a cough suppressant. One stereoisomer of ketamine is an anesthetic, whereas the other causes hallucinations. ■

The typical numbers of covalent bonds formed by atoms common in biomolecules are shown in Table 2-1. A hydrogen atom forms only one covalent bond. An oxygen atom usually forms only two covalent bonds but has two additional pairs of electrons (indicated as dots in the table) that can participate in noncovalent interactions. These extra electrons that are not shared in a covalent bond are also called *nonbonding* electrons. As we will see in later chapters, the nonbonding electrons can play important roles in the chemical reactions between biomolecules. A sulfur atom forms two covalent bonds in hydrogen sulfide (H_2S) but can accommodate six covalent bonds in sulfuric acid (H_2SO_4) and its sulfate derivatives. Nitrogen and phosphorus atoms each have five electrons to share. In ammonia (NH_3), the nitrogen atom forms three covalent bonds; the pair of electrons around the atom not involved in a covalent bond can take part in noncovalent interactions. In the ammonium ion (NH_4^+), the nitrogen atom forms four covalent bonds, which have a tetrahedral geometry. A phosphorus atom commonly forms five

covalent bonds, as in phosphoric acid (H_3PO_4) and its phosphate derivatives, which form the backbone of nucleic acids. Phosphate groups covalently attached to proteins play a key role in regulating the activity of many proteins, and the central molecule in cellular energetics, ATP, contains three phosphate groups (see Section 2.4). A summary of common covalent linkages and functional groups, which confer distinctive chemical properties on the molecules of which they are a part, is provided in Table 2-2.

TABLE 2-1 • Bonding Properties of Atoms Most Abundant in Biomolecules

Atom and Outer Electrons	Usual Number of Covalent Bonds	Typical Bond Geometry
H	1	H
O	2	O
S	2, 4, or 6	S
N	3 or 4	N
P	5	P
C	4	C

TABLE 2-2 • Common Functional Groups and Linkages in Biomolecules

Functional Groups			
—OH Hydroxyl (alcohol)	O ‖ —C—R Acyl (triacylglycerol)	O ‖ —C— Carbonyl (ketone)	O ‖ —C—O⁻ Carboxyl (carboxylic acid)
—SH Sulfhydryl (thiol)	$-NH_2$ or $-\overset{+}{N}H_3$ Amino (amine)	O ‖ —O—P—O⁻ | O⁻ Phosphate (phosphorylated molecule)	O O ‖ ‖ ⁻O—P—O—P—O⁻ | | OH OH Pyrophosphate (diphosphate)
Linkages			
O ‖ —C—O—C— Ester		—C—O—C— Ether	O ‖ —N—C— Amide

All Covalent Bonds Are Not Equal: Electrons May Be Shared Equally or Unequally in Covalent Bonds

The extent of an atom's ability to attract an electron is called its *electronegativity*. In a bond between atoms with identical or similar electronegativities, the bonding electrons are essentially shared equally between the two atoms, as is the case for most carbon-carbon single bonds (C—C) and carbon-hydrogen single bonds (C—H). Such bonds are called **nonpolar**. In many molecules, however, the covalently bonded atoms sharing electrons have different electronegativities, resulting in unequal sharing of electrons. The bond between atoms that share electrons unequally is said to be **polar**.

One end of a polar bond has a partial negative charge (δ^-), and the other end has a partial positive charge (δ^+). In an O–H bond, for example, the oxygen atom is more electronegative than the hydrogen atom. As a result, the electrons forming the covalent bond spend more time around the oxygen atom than around the hydrogen. Thus the O–H bond possesses an electric **dipole**, a positive charge separated from an equal but opposite negative charge. The amount of δ^- charge on the oxygen atom of an O–H dipole is approximately 25 percent that of an electron, and there is an equivalent and opposite δ^+ charge on the H atom.

A common quantitative measure of the extent of charge separation, or strength, of a dipole is called the **dipole moment**, μ. For a chemical bond, the dipole moment is the product of the partial charge on each atom and the distance between the two atoms. Many molecules have multiple dipoles. In these cases, the amount of charge separation for the molecule as a whole — the molecule's dipole moment — depends in part on the dipole moments of all of its individual chemical bonds and in part on the geometry of the molecule (the relative orientations of the individual bond dipole moments).

Consider the example of water (H_2O). A water molecule has two O–H bonds and thus two individual bond dipole moments. If water were a linear molecule with the two bonds on exact opposite sides of the O atom, the two dipoles on each end of the molecule would be

identical in strength but would be oriented in opposite directions. The two dipole moments would cancel each other, and the dipole moment of the molecule as a whole would be zero. However, water is a V-shaped molecule, with the individual dipoles of its two O—H bonds both pointing toward the oxygen. One end of the water molecule (the end with the oxygen atom) has a partial negative charge and the other end (the one with the two hydrogen atoms) has a partial positive charge. As a consequence, the molecule as a whole is a dipole with a well-defined dipole moment (Figure 2-5). This dipole moment and the electronic properties of the oxygen and hydrogen atoms allow water to form electrostatic, noncovalent interactions with other water molecules and with molecules of other types. These interactions play a critical role in almost every biochemical interaction in cells and organisms, as we will see shortly.

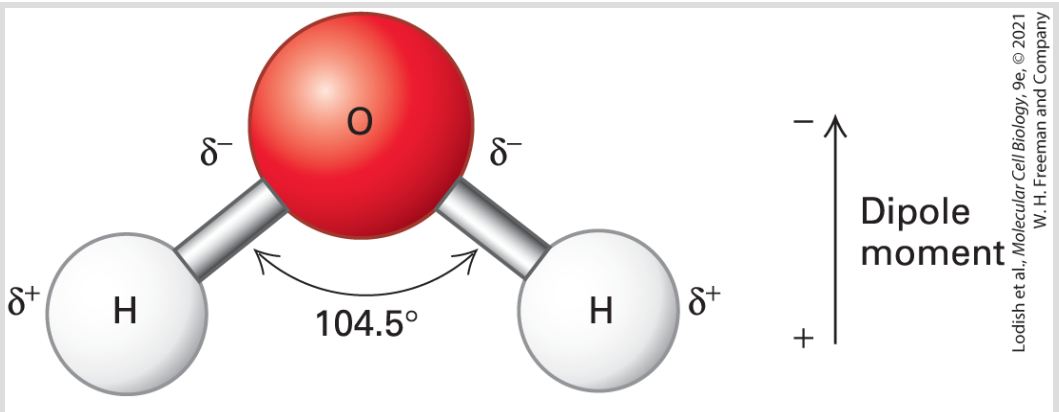
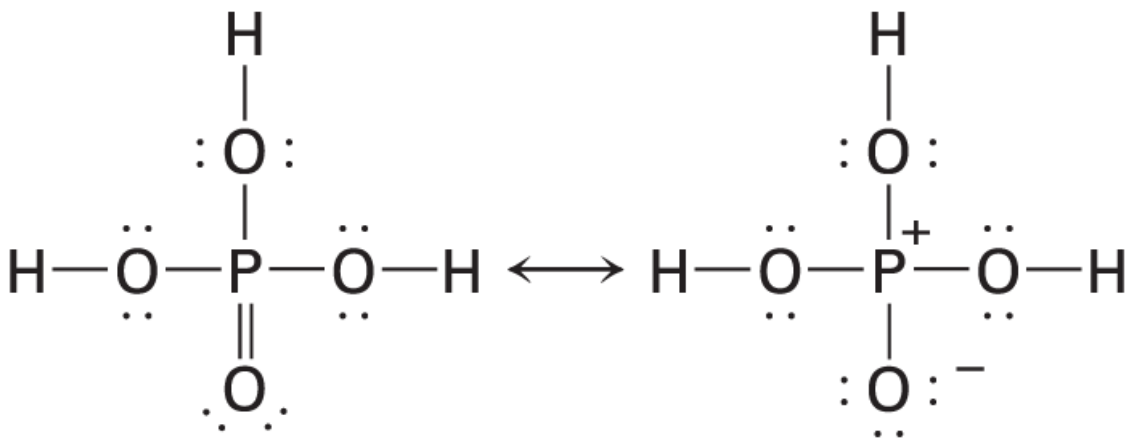


FIGURE 2-5 The dipole nature of a water molecule. The symbol δ represents a partial charge (a weaker charge than the one on an electron or a proton). Because of the difference in the electronegativities of H and O, each of the polar H—O bonds in water is a dipole. The sizes and directions of the dipoles of each

of the bonds determine the net distance between charges and amount of charge separation in the molecule — its dipole moment.

Another important example of polarity is the O=P double bond in H_3PO_4. In the structure of H_3PO_4 shown on the top of the next page, lines represent single and double bonds and nonbonding electrons are shown as pairs of dots (each dot represents one electron):



Because of the polarity of the O=P double bond, H_3PO_4 can also be represented by the structure on the right, in which one of the electrons from the P=O double bond has accumulated around the O atom, giving it a negative charge and leaving the P atom with a positive charge. These charges are important in noncovalent interactions. Neither of these two models precisely describes the electronic state of H_3PO_4. The actual structure can be considered to be an intermediate, or hybrid, between these two representations, as indicated by the double-headed arrow between them. Such intermediate structures are called *resonance hybrids*.

Covalent Bonds Are Much Stronger and More Stable Than Noncovalent Interactions

Covalent bonds are considered to be strong because the energies required to break them are much greater than the thermal energy available at room temperature ($25\ ^\circ\mathrm{C}$) or body temperature ($37\ ^\circ\mathrm{C}$). As a consequence, covalent bonds are stable at these temperatures. For example, the thermal energy available at $25\ ^\circ\mathrm{C}$ is approximately 0.6 kilocalorie per mole ($\mathrm{kcal/mol}$), whereas the energy required to break the $\mathrm{C{-}C}$ bond in ethane is about 140 times larger (Figure 2-6). Consequently, at room temperature ($25\ ^\circ\mathrm{C}$), fewer than 1 in 10^{12} ethane molecules are broken into pairs of CH_3 molecules, each containing an unpaired, nonbonding electron (called a *radical*).

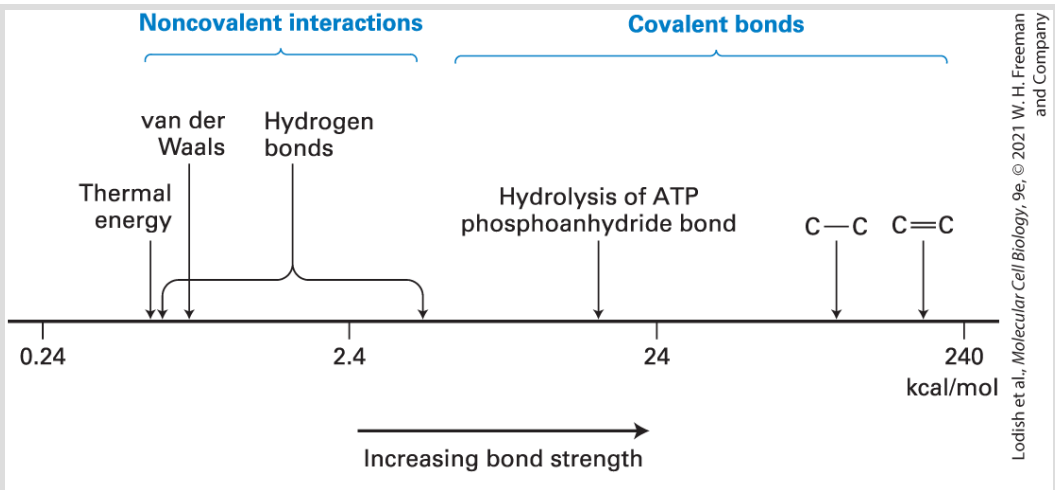


FIGURE 2-6 Relative energies of covalent bonds and noncovalent interactions. Bond energies are defined as the energy required to break a particular type of linkage. Shown here are the energies required to break a

variety of linkages, arranged on a log scale. Covalent bonds, including single (C−C) and double (C=C) carbon-carbon bonds, are one to two powers of 10 stronger than noncovalent interactions. Noncovalent interactions have energies somewhat greater than the thermal energy of the environment at normal room temperature (25 °C). Many biological processes are coupled to the energy released during hydrolysis of a phosphoanhydride bond in ATP.

Covalent single bonds in biological molecules have energies similar to the energy of the C−C bond in ethane. Because more electrons are shared between atoms in double bonds, they require more energy to break than single bonds. For instance, it takes 84 kcal/mol to break a single C−O bond but 170 kcal/mol to break a C=O double bond. The most common double bonds in biological molecules are C=O, C=N, C=C, and P=O.

In contrast, the energy required to break noncovalent interactions is only 1−5 kcal/mol, much less than the bond energies of covalent bonds (see Figure 2-6). Indeed, noncovalent interactions are weak enough that they are constantly being formed and broken at room temperature. Although these weak interactions have a transient existence at physiological temperatures (25−37 °C), multiple noncovalent interactions can, as we will see, act together to produce highly stable and specific associations between different parts of a large molecule or between different macromolecules. Protein-protein and protein-nucleic acid interactions are good examples of such noncovalent interactions. Below, we review the four main

types of noncovalent interactions and then consider their roles in the binding of biomolecules to one another and to other molecules.

Ionic Bonds Are Noncovalent Interactions Formed by the Electrostatic Attractions Between Oppositely Charged Ions

Ionic bonds (also called *ionic interactions*) result from the attraction between a positively charged ion — a **cation** — and a negatively charged ion — an **anion**. In sodium chloride (NaCl), for example, the bonding electron contributed by the sodium atom is completely transferred to the chlorine atom (Figure 2-7a). Unlike covalent bonds, ionic interactions do not have fixed or specific geometric orientations because the electrostatic field around an ion — its attraction for an opposite charge — is uniform in all directions. In solid NaCl, oppositely charged ions pack tightly together in an alternating pattern, forming the highly ordered crystalline array, or lattice, that is typical of salt crystals (Figure 2-7b, and see the opening figure for this chapter). The energy required to break an ionic interaction depends on the distance between the ions and the electrical properties of the environment of the ions.

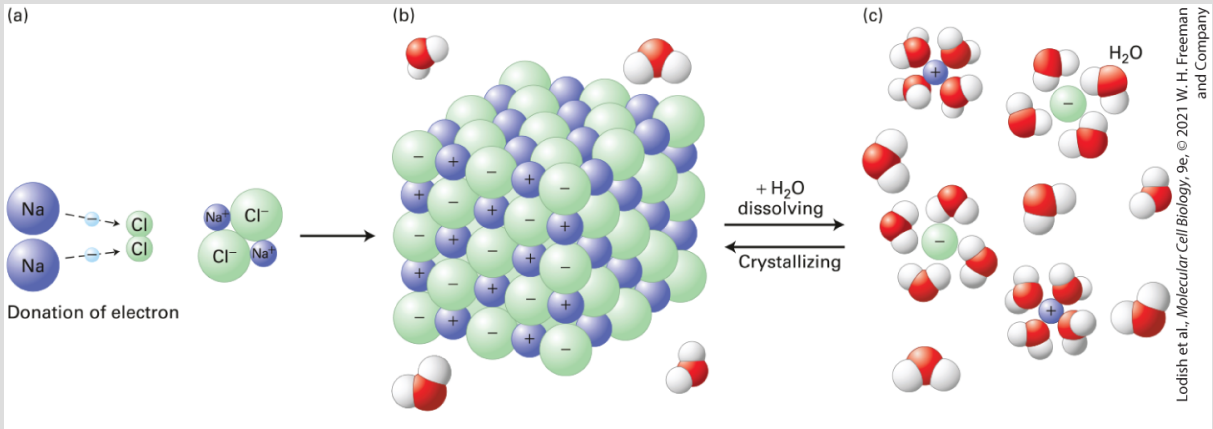


FIGURE 2-7 Electrostatic interactions of the oppositely charged ions of salt (NaCl) in crystals and in aqueous solution. (a) In crystalline table salt, sodium atoms are positively charged ions (Na^+) due to the loss of one electron each, whereas chloride atoms are correspondingly negatively charged (Cl^-) due to the gain of one electron each. (b) In solid form, ionic compounds form neatly ordered arrays, or crystals, of tightly packed ions in which the positively and negatively charged ions counterbalance each other. (c) When the crystals are dissolved in water, the ions separate, and their charges, no longer balanced by immediately adjacent ions of opposite charge, are stabilized by interactions with polar water. The water molecules and the ions are held together by electrostatic interactions between the charges on the ion and the partial charges on the water's oxygen and hydrogen atoms. In aqueous solutions, all ions are surrounded by a hydration shell of water molecules.

When solid salts dissolve in water, the ions separate from one another and are stabilized by their interactions with water molecules. In aqueous solutions, simple ions of biological significance, such as Na^+, K^+, Ca^{2+}, Mg^{2+}, and Cl^-, are *hydrated*, surrounded by a stable shell of water molecules held in place by electrostatic interactions between the ion at the center and the oppositely charged ends of the water molecules, which are dipoles (Figure 2-7c). Energy is released when ions tightly bind water molecules and spread out in an aqueous solution, due to the

increase in disorder, or entropy (see Section 2.4). Most ionic compounds dissolve readily in water because this energy of hydration is greater than the lattice energy that stabilizes the crystal structure. Parts or all of the aqueous *hydration shell* must be removed from ions in solution when they interact directly with proteins. For example, water of hydration is lost when ions pass through protein pores in the cell membrane during nerve conduction.

The relative strength of the interaction between two oppositely charged ions, A^- and C^+, depends on the concentration of other ions in a solution. The higher the concentration of other ions (e.g., Na^+ and Cl^-), the more opportunities A^- and C^+ have to interact ionically with those other ions, and thus the lower the energy required to break the interaction between A^- and C^+. As a result, increasing the concentrations of salts such as NaCl in a solution of biological molecules can weaken and even disrupt the ionic interactions holding the biomolecules together. This principle can be exploited to separate complex mixtures of interacting molecules, such as proteins, into their individual, pure components.

Hydrogen Bonds Are Noncovalent Interactions That Determine the

Properties of Water and the Water Solubility of Uncharged Molecules

A **hydrogen bond** is the interaction of a hydrogen atom with a partial positive charge in a dipole, such as water, with unpaired electrons from another atom, either in the same or in a different molecule. Normally, a hydrogen atom forms a covalent bond with a single other atom. However, a hydrogen atom covalently bonded to an electronegative donor atom D may form an additional weak association, the hydrogen bond, with an acceptor atom A, which must have a nonbonding pair of electrons available for the interaction:

$$D^{\delta-}\text{—}H^{\delta+} + : A^{\delta-} \rightleftharpoons D^{\delta-}\text{—}H^{\delta+} \underset{\textbf{Hydrogen bond}}{\cdot\!\square\!\cdot\!\square\!\cdot\!\square\!\cdot} : A^{\delta-}$$

The covalent D − H bond is a bit longer than it would be if there were no hydrogen bond because the acceptor pulls the hydrogen away from the donor. An important feature of all hydrogen bonds is directionality. In the strongest hydrogen bonds, the donor atom, the hydrogen atom, and the acceptor atom all lie in a straight line. Nonlinear hydrogen bonds are weaker than linear ones; still, multiple nonlinear hydrogen bonds help to stabilize the three-dimensional structures of many proteins.

Hydrogen bonds are both longer and weaker than covalent bonds between the same atoms. In water, for example, the distance between the nuclei of the hydrogen and oxygen atoms forming a hydrogen bond connecting two adjacent water molecules is about 0.27 nm, about twice the length of the covalent O–H bonds within a single water molecule (Figure 2-8a). A hydrogen bond between water molecules (approximately 5 kcal/mol) is much weaker than a covalent O–H bond (roughly 110 kcal/mol), although it is stronger than many other hydrogen bonds in biological molecules (1–2 kcal/mol). Extensive hydrogen bonding between water molecules accounts for many of water's key properties, including its unusually high melting and boiling points and its ability to dissolve many other molecules.

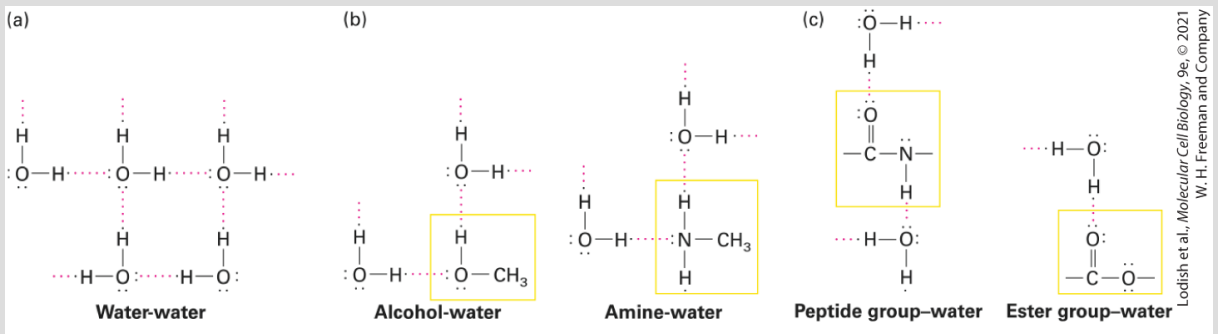


FIGURE 2-8 Hydrogen bonding of water with itself and with other compounds. Each pair of nonbonding outer electrons in an oxygen or a nitrogen atom can accept a hydrogen atom in a hydrogen bond. The hydroxyl and the amino groups can also form hydrogen bonds with water. (a) In liquid water, each water molecule forms transient hydrogen bonds with several others, creating a dynamic network of hydrogen-bonded molecules. (b) Water can also form hydrogen bonds with alcohols and amines, which accounts for the high solubility of these compounds. (c) The peptide group and the ester group, which are present in many biomolecules, commonly participate in hydrogen bonds with water or polar groups in other molecules.

The solubility of uncharged substances in an aqueous environment depends largely on their ability to form hydrogen bonds with water. For instance, the hydroxyl group (–OH) in alcohols (—CH_2OH) and the amino group (—NH_2) in amines (—CH_2NH_2) can form several hydrogen bonds with water, which allows these molecules to dissolve in water at high concentrations (Figure 2-8b). In general, molecules with polar bonds that easily form hydrogen bonds with water, as well as charged molecules and ions that interact with the dipole in water, can readily dissolve in water; that is, they are hydrophilic. In addition to hydroxyl and amino groups, many biological molecules contain peptide and ester groups, which form hydrogen bonds with water via otherwise nonbonded electrons on their carbonyl oxygens (Figure 2-8c). X-ray crystallography combined with computational analysis permits an accurate depiction of the distribution of the outermost unbonded electrons of atoms that can participate in hydrogen bonds as well as the electrons in covalent bonds, as illustrated in Figure 2-9.

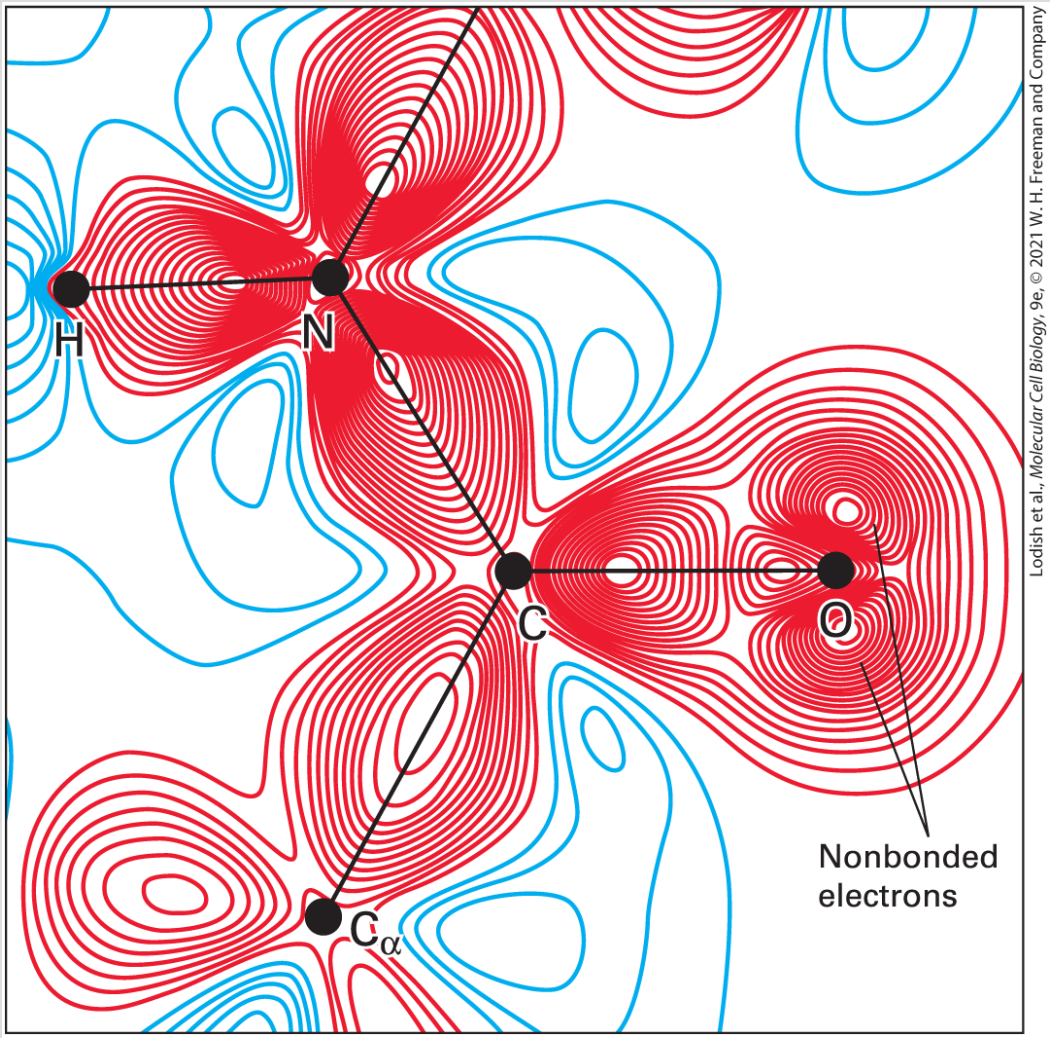


FIGURE 2-9 Distribution of bonding and outer nonbonding electrons in the peptide group. Shown here is a peptide bond linking two amino acids within a protein called *crambin*. No protein has been structurally characterized at higher resolution than crambin. The black lines represent the covalent bonds between atoms. The red (negative) and blue (positive) lines represent contours of charge determined using x-ray crystallography and computational methods. The greater the number of contour lines, the higher the charge. The high density of red contour lines between atoms represents the covalent bonds (shared electron pairs). The two sets of red contour lines emanating from the oxygen (O) and not falling on a covalent bond (black line) represent the two pairs of nonbonding electrons on the oxygen that are available to participate in hydrogen bonding. The high density of blue contour lines near the hydrogen (H)

bonded to nitrogen (N) represents a partial positive charge, indicating that this H can participate in hydrogen bonding.

[From *Proc. Nat'l. Acad. Sci. USA*, 2000, **97**(7):3171–3176, Fig. 3A. Accurate protein crystallography at ultra-high resolution: Valence electron distribution in crambin, by Christian Jelsch et al., copyright (2000) National Academy of Sciences, USA.]

Van der Waals Interactions Are Weak Attractive Interactions Caused by Transient Dipoles

When any two atoms approach each other closely, they create a weak attractive force called a **van der Waals interaction**. These interactions result from the momentary random fluctuations in the distribution of the electrons of any atom. The random fluctuations give rise to a transient unequal distribution of electrons, creating a temporary dipole. If two noncovalently bonded atoms are close enough, electrons of one atom will perturb the electrons of the other. This perturbation generates a transient dipole in the second atom, and the two dipoles attract each other weakly (Figure 2-10). Similarly, a polar covalent bond in one molecule attracts an oppositely oriented dipole in another.

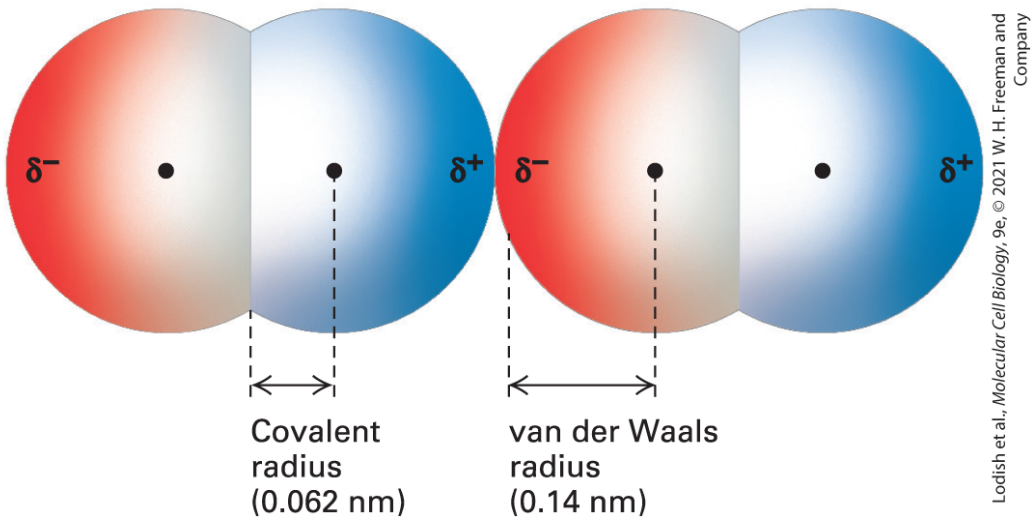


FIGURE 2-10 Two oxygen molecules in van der Waals contact. In this model, red indicates negative charge and blue indicates positive charge. Transient dipoles in the electron clouds of all atoms give rise to weak attractive forces, called *van der Waals interactions*. Each type of atom has a characteristic van der Waals radius at which van der Waals interactions with other atoms are optimal. Because atoms repel one another if they are close enough together for their outer electrons to overlap without being shared in a covalent bond, the van der Waals radius is a measure of the size of the electron cloud surrounding an atom. The covalent radius indicated here is for the double bond of O=O; the single-bond covalent radius of oxygen is slightly longer.

Van der Waals interactions, involving either transient or permanent dipoles, occur in all types of molecules, both polar and nonpolar. In particular, van der Waals interactions are responsible for the cohesion between nonpolar molecules such as heptane, $CH_3—(CH_2)_5—CH_3$, that cannot form hydrogen bonds or ionic interactions with each other. The strength of van der Waals interactions decreases rapidly with increasing distance; thus these noncovalent interactions can form only when atoms are quite close

to one another. However, if atoms get too close together, the negative charges of their electrons create a repulsive force. When the van der Waals attraction between two atoms exactly balances the repulsion between their two electron clouds, the atoms are said to be in van der Waals contact. The strength of the van der Waals interaction is about 1 kcal/mol, so it is weaker than typical hydrogen bonds, and its energy is only slightly higher than the average thermal energy of molecules at 25 °C. Thus multiple van der Waals interactions, a van der Waals interaction together with other noncovalent interactions, or both are required to form van der Waals–mediated, stable attractions within and between molecules.

The Hydrophobic Effect Causes Nonpolar Molecules to Adhere to One Another

Because nonpolar molecules do not contain charged groups, do not possess a dipole moment, and do not become hydrated, they are insoluble, or almost insoluble, in water; that is, they are hydrophobic. The covalent bonds between two carbon atoms and between carbon and hydrogen atoms are the most common nonpolar bonds in biological systems. **Hydrocarbons** — molecules made up only of carbon and hydrogen — are virtually insoluble in water. Large triacylglycerols (also known as triglycerides), which make up animal fats and vegetable oils, also essentially are insoluble

in water. As we will see later, the major part of these molecules consists of long hydrocarbon chains. After being shaken in water, triacylglycerols form a separate phase. A familiar example is the separation of oil from the water-based vinegar in an oil-and-vinegar salad dressing.

Nonpolar molecules or nonpolar parts of molecules tend to aggregate in water owing to a phenomenon called the **hydrophobic effect**. Because water molecules cannot form hydrogen bonds with nonpolar substances, they tend to form cages of relatively rigid hydrogen-bonded pentagons and hexagons around nonpolar molecules (Figure 2-11, *left*). This state is energetically unfavorable because it decreases the entropy, or randomness, of the population of water molecules. (The role of entropy in chemical systems is discussed in Section 2.4.) If nonpolar molecules in an aqueous environment aggregate with their hydrophobic surfaces facing each other, the net hydrophobic surface area exposed to water is reduced (Figure 2-11, *right*). As a consequence, fewer water molecules are needed to form the ordered cages surrounding the nonpolar molecules, entropy increases relative to the unaggregated state, and an energetically more favorable state is reached. In a sense, then, water squeezes the nonpolar molecules into aggregates. Rather than constituting an attractive force, as in hydrogen bonds, the hydrophobic effect results from an avoidance of an unstable state — that is, extensive water cages around individual nonpolar molecules.

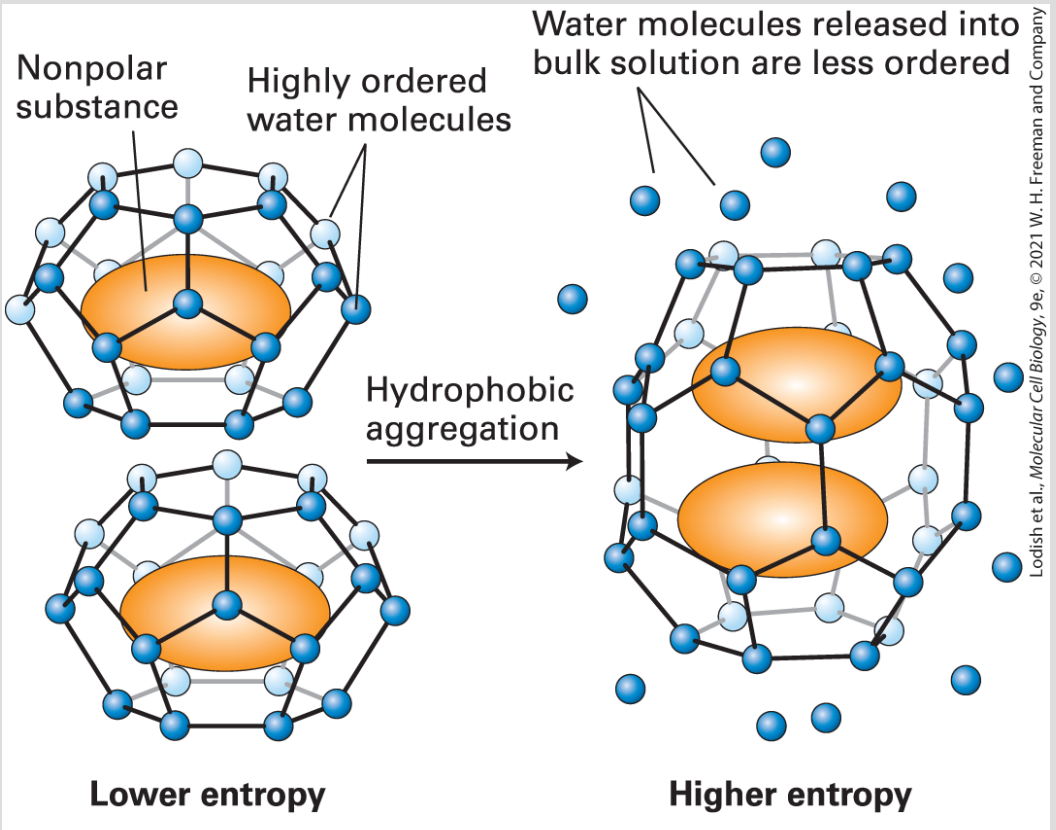


FIGURE 2-11 Schematic depiction of the hydrophobic effect. Cages of water molecules form around nonpolar molecules in solution. These water molecules in cages are more ordered than water molecules in the surrounding bulk liquid. Aggregation of nonpolar molecules reduces the number of water molecules involved in forming highly ordered cages, resulting in a higher entropy, more energetically favorable state (*right*) compared with the unaggregated state (*left*).

Nonpolar molecules can also associate, albeit weakly, through van der Waals interactions. The net result of the hydrophobic effect and van der Waals interactions is a very powerful tendency for hydrophobic molecules in water to interact with one another, not with water molecules. Simply put, *like dissolves like*. Polar molecules

dissolve in polar solvents such as water; nonpolar molecules dissolve in nonpolar solvents such as hexane.

One well-known hydrophobic molecule is cholesterol (see Section 2.2). Cholesterol, triglycerides, and other poorly water-soluble molecules are called *lipids*. Unlike hydrophilic molecules such as glucose or amino acids, lipids cannot readily dissolve in the liquid portion of the blood, the aqueous circulatory system that transports molecules and cells throughout the body. Instead, lipids, such as cholesterol, must be packaged into special hydrophilic carriers, called *lipoproteins*, that can themselves dissolve in the blood and be transported throughout the body. There can be hundreds to thousands of lipid molecules packed into the center, or core, of each lipoprotein. The hydrophobic core is surrounded by amphipathic molecules that have hydrophilic parts that interact with water and hydrophobic parts that interact with one another and the core. The packaging of lipids into lipoproteins (discussed in Chapter 14) permits their efficient transport in blood and is reminiscent of the containerization of cargo for efficient, long-distance transport via cargo ships, trains, and trucks.

High-density lipoprotein (HDL) and low-density lipoprotein (LDL) are two such lipoprotein carriers. HDL is associated with reduced heart disease, and LDL with increased heart disease, and they are therefore often referred to as *good* and *bad* cholesterol, respectively. Actually, the cholesterol molecules and their derivatives that are

carried by both HDL and LDL are essentially identical and in themselves are neither good nor bad. However, HDL and LDL have different effects on cells. The bad cholesterol, LDL, contributes to clogging of the arteries (known as *atherosclerosis*) and consequent heart disease and stroke. The mechanisms linking HDL to reduced risk for heart attacks are unclear. ■

Molecular Complementarity Due to Noncovalent Interactions Leads to a Lock-and-Key Fit Between Biomolecules

Both inside and outside cells, ions and molecules constantly collide. The higher the concentration of any two types of molecules, the more likely they are to encounter each other. When two molecules do encounter each other, they are most likely to simply bounce apart because the noncovalent interactions that would bind them together are weak and have a transient existence at physiological temperatures. However, two distinct molecules can exhibit **molecular complementarity**, a lock-and-key kind of fit between their shapes, charges, or other physical properties. Two such structurally complementary molecules can form multiple noncovalent interactions at close range. When two such molecules bump into each other, these multiple interactions can result in their sticking together, or binding.

Figure 2-12 illustrates how a variety of multiple weak interactions can cause two hypothetical proteins to bind together tightly. Numerous examples of such protein-to-protein molecular complementarity may be found throughout this book (see, e.g., Figures 16-9 and 16-18b). Almost any other arrangement of the same groups of atoms on the two surfaces would not allow the molecules to bind so tightly. Such molecular complementarity between regions within a protein molecule allow it to fold into a unique three-dimensional shape (see Chapter 3); it is also what holds the two chains of DNA together in a double helix (see Chapter 5). Similar interactions underlie the association of groups of molecules into multimolecular assemblies, or complexes, leading to the formation of muscle fibers, to the gluelike associations between cells in solid tissues, and to numerous other cellular structures. The antibodies that help neutralize pathogens (see Chapter 24) bind to them using similar principles of complementary molecular shapes.

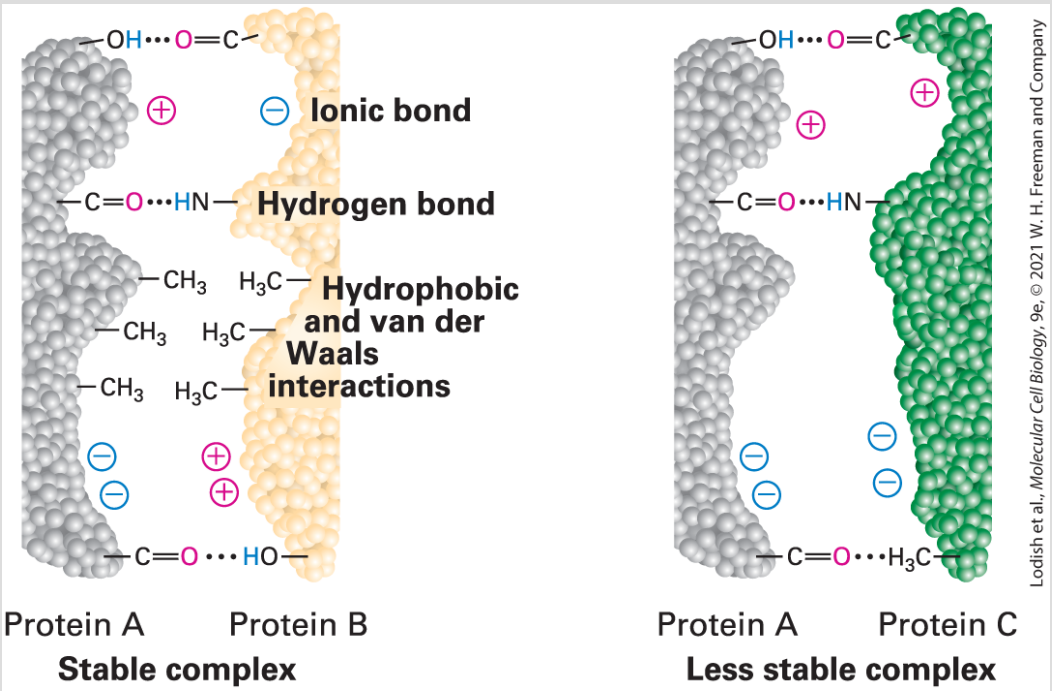


FIGURE 2-12 Molecular complementarity permits tight protein bonding via multiple noncovalent interactions. The complementary shapes, charges, polarity, and hydrophobicity of two protein surfaces permit multiple weak interactions, which in combination produce a strong interaction and tight binding. Because deviations from molecular complementarity substantially weaken binding, a particular surface region of any given biomolecule usually can bind tightly to only one or a very limited number of other molecules. The complementarity of the two protein molecules on the left permits them to bind much more tightly than the two noncomplementary proteins on the right.

Depending on the number and strength of the noncovalent interactions between the two molecules and on their environment, their binding may be tight or loose and, as a consequence, either lasting or transient. The higher the *affinity* of two molecules for each other, the better the molecular fit between them, the more noncovalent interactions can form, and the more tightly they can bind together. An important quantitative measure of affinity is the

binding dissociation constant K_d, described in Section 2.3. In many cases in cell biology, molecules must bind to one another but also be able to subsequently separate (dissociate). In other words, the binding needs to be reversible. Thus the binding cannot be too tight, or the molecules would not be able to dissociate when necessary. It is important to note that many large biological molecules are not hard, rigid structures, but rather can be somewhat malleable. Thus the binding of a molecule to another has the potential to induce a change in the shape of its binding partner. When the molecular complementarity increases after such interactions, the process is called *induced fit*.

As we discuss in Chapter 3, nearly all the chemical reactions that occur in cells depend on the binding properties of enzymes. These proteins not only speed up, or catalyze, reactions, but do so with a high degree of *specificity* because they are able to bind tightly to only one or a few related molecules. The specificity of intermolecular interactions and reactions, which depends on molecular complementarity, is essential for many processes critical to life.

KEY CONCEPTS OF SECTION 2.1

Covalent Bonds and Noncovalent Interactions

- The terms *hydrophilic, hydrophobic*, and *amphipathic/amphiphilic* refer to the tendency of molecules to be water-loving, incapable of interacting with water, and having features of or being tolerant of both, respectively. Hydrophilic molecules typically dissolve readily in water, whereas hydrophobic molecules are poorly soluble or insoluble in water.

- Covalent bonds consist of pairs of electrons shared by two atoms. Covalent bonds arrange the atoms of a molecule into a specific geometry.
- Many molecules in cells contain at least one asymmetric carbon atom, which is bonded to four dissimilar atoms. Such molecules can exist as stereoisomers (mirror images), designated D and L (see Figure 2-4), which have different biological activities. Nearly all amino acids are L isomers.
- Electrons may be shared equally or unequally in covalent bonds. Atoms that differ in electronegativity form polar covalent bonds, in which the bonding electrons are distributed unequally. One end of a polar bond has a partial positive charge and the other end has a partial negative charge (see Figure 2-5).
- Covalent bonds are stable in biological systems because the relatively high energies required to break them $(50-200\ \mathrm{kcal/mol})$ are much larger than the thermal kinetic energy available at room $(25\ ^{\circ}\mathrm{C})$ or body $(37\ ^{\circ}\mathrm{C}$ in mammals) temperatures.
- Noncovalent interactions between atoms are considerably weaker than covalent bonds, with energies ranging from about $1-5\ \mathrm{kcal/mol}$ (see Figure 2-6).
- Four main types of noncovalent interactions occur in biological systems: ionic bonds, hydrogen bonds, van der Waals interactions, and interactions due to the hydrophobic effect.
- Ionic bonds result from the electrostatic attraction between the positive and negative charges of ions. In aqueous solutions, all cations and anions are surrounded by a shell of bound water molecules (see Figure 2-7c). Increasing the salt (e.g., NaCl) concentration of a solution can weaken the relative strength of and even break the ionic bonds between biomolecules.
- In a hydrogen bond, a hydrogen atom covalently bonded to an electronegative atom associates with an acceptor atom whose nonbonding electrons attract the hydrogen (see Figure 2-8).
- Weak van der Waals interactions result from the attraction between transient dipoles associated with all molecules. They can form when two atoms approach each other closely (see Figure 2-10).
- In an aqueous environment, nonpolar molecules or nonpolar parts of larger molecules are driven together by the hydrophobic effect, thereby reducing the extent of their direct contact with water molecules (see Figure 2-11).
- Molecular complementarity is the lock-and-key fit between molecules whose shapes, charges, and other physical properties are complementary. Multiple noncovalent interactions can form between complementary molecules, causing them to bind tightly (see Figure 2-12), but cannot form between molecules that are not complementary.
- The high degree of binding specificity that results from molecular complementarity is one of the features that underlies intermolecular interactions in biology and thus is

essential for many processes critical to life.

2.2 Chemical Building Blocks of Cells

A common theme in biology is the construction of large **macromolecules** and macromolecular complexes (clusters of macromolecules) out of smaller molecular subunits, which can be thought of as building blocks. Often these subunits are similar or identical. The three main types of biological macromolecules — proteins, nucleic acids, and polysaccharides — are all **polymers** composed of multiple covalently linked small molecules, or **monomers** (Figure 2-13). **Proteins** are linear polymers containing up to several thousand amino acids linked by **peptide bonds**. **Nucleic acids** are linear polymers containing dozens to millions of nucleotides linked by **phosphodiester bonds**. **Polysaccharides** are linear or branched polymers of monosaccharides (sugars) such as glucose linked by **glycosidic bonds**.

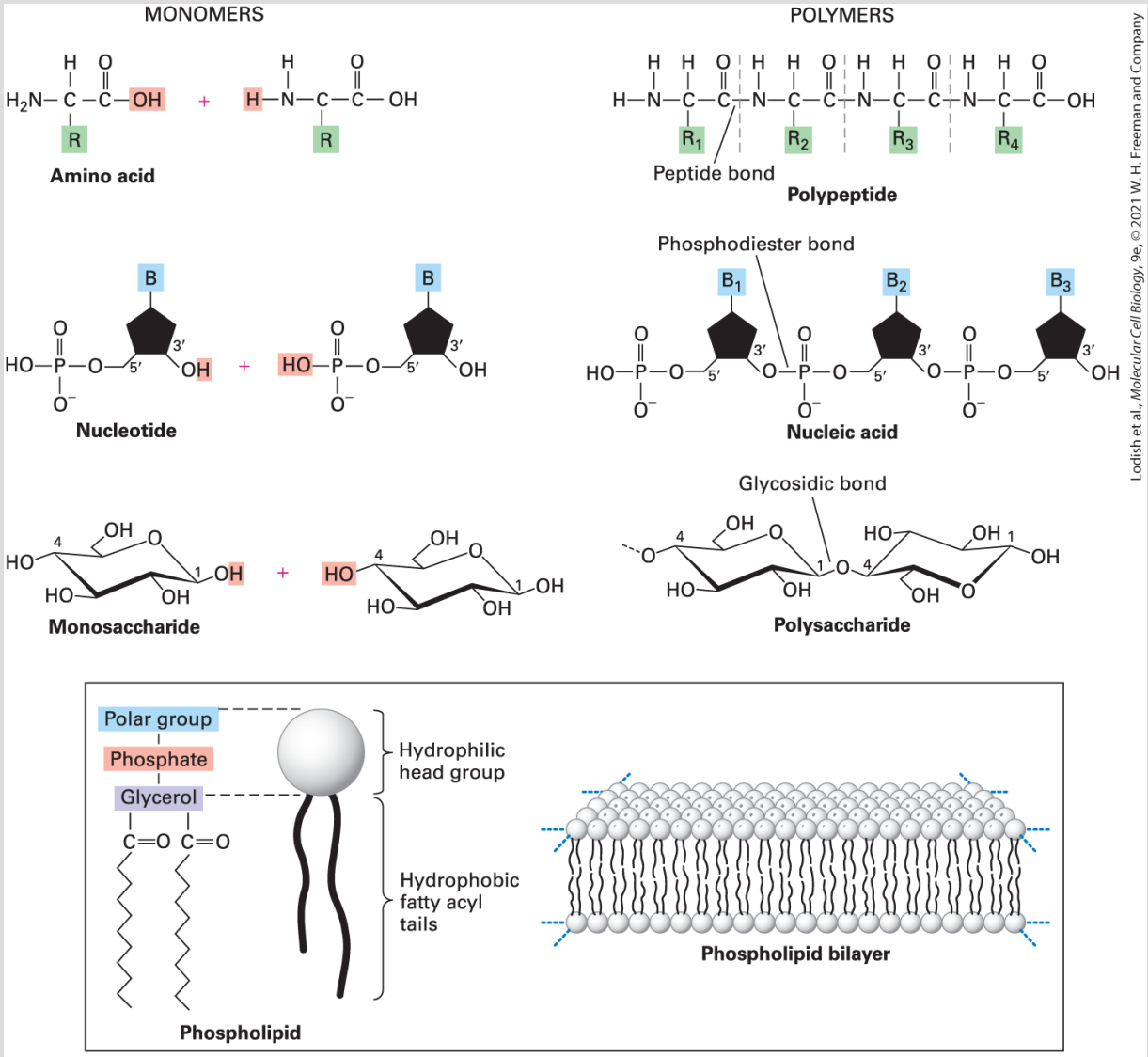


FIGURE 2-13 Overview of the cell's principal chemical building blocks. *Top*: The three major types of biological macromolecules are each assembled by the polymerization of multiple small molecules (monomers) of a particular type: proteins from amino acids (see Chapter 3), nucleic acids from nucleotides (see Chapter 5), and polysaccharides from monosaccharides (sugars). Each monomer is covalently linked into the polymer by a reaction whose net result is loss of a water molecule (dehydration). *Bottom*: In contrast, phospholipid monomers noncovalently assemble into a bilayer structure, which forms the basis of all cellular membranes (see Chapter 10).

Although the actual mechanisms of covalent bond formation between monomers are complex, as we will see, the formation of a covalent bond between two monomers usually results in the net loss of a hydrogen (H) from one monomer and a hydroxyl (OH) from the other monomer — or the net loss of one water molecule — and can

therefore be thought of as a *dehydration reaction*. The cleavage of a covalent bond in a polymer is the result of the reverse reaction, or the addition of water, called *hydrolysis*. Hydrolysis can release a single monomer from the end of the polymer or can break the polymer into two shorter polymers, usually of unequal length, when the break occurs somewhere within the polymer chain. The covalent bonds that link monomers together are usually stable under normal biological conditions (e.g., 37 °C, neutral pH), so these biopolymers are stable and can perform a wide variety of jobs in cells, such as storing information, catalyzing chemical reactions, serving as structural elements that define cellular shape and movement, and many others. Macromolecular complexes can also be assembled using noncovalent interactions. The two-ply, or *bilayer*, structure of cellular membranes is built up by the noncovalent assembly of many thousands of small molecules called *phospholipids* (see Figure 2-13).

In this chapter, we focus on the chemical building blocks making up cells — amino acids, nucleotides, sugars, and phospholipids. The structure, function, and assembly of proteins, nucleic acids, polysaccharides, and biomembranes are discussed in subsequent chapters.

Amino Acids Differing Only in Their Side Chains Compose Proteins

The monomeric building blocks of proteins are 20 **amino acids**, which — when incorporated into a protein polymer — are sometimes called **residues**. All amino acids have a characteristic structure consisting of a central **alpha carbon atom** (C_α) bonded to four different chemical groups (Figure 2-14): an amino (—NH_2) group, a carboxyl or carboxylic acid (—COOH) group (hence the name *amino acid*), a hydrogen (H) atom, and one variable group, called a *side chain* or **R group**. The α carbon in all amino acids is asymmetric with the exception of glycine, whose R group is a hydrogen. Because they are asymmetric, these molecules can exist in two mirror-image forms, called by convention the *D* (dextro) and the *L* (levo) *isomers* (see Figure 2-4). The two isomers cannot be interconverted (one made identical to the other)

without breaking and then re-forming a chemical bond in one of them. With rare exceptions, only the L forms of amino acids are found in proteins. However, D amino acids are prevalent in bacterial cell walls and other microbial products.

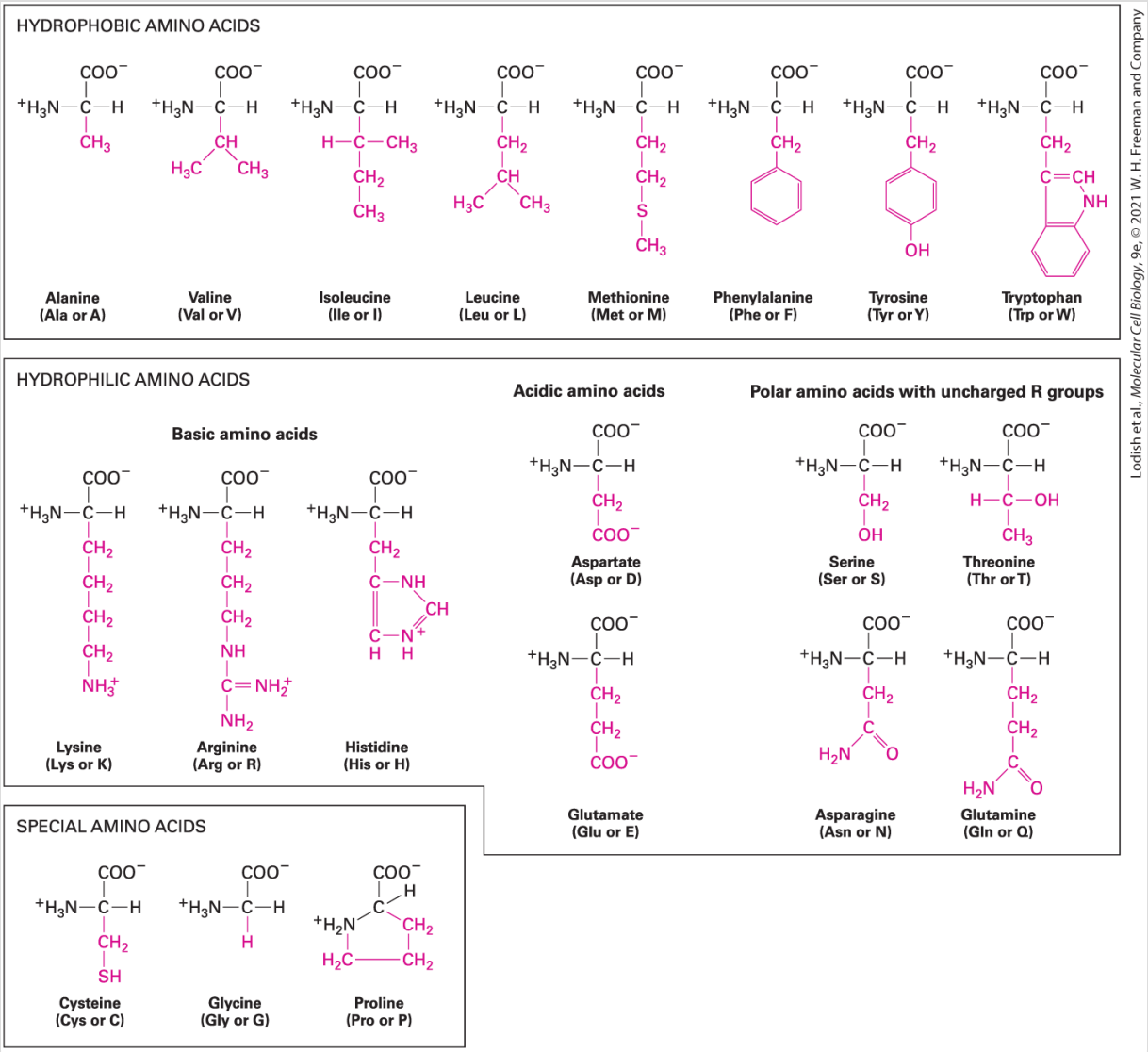


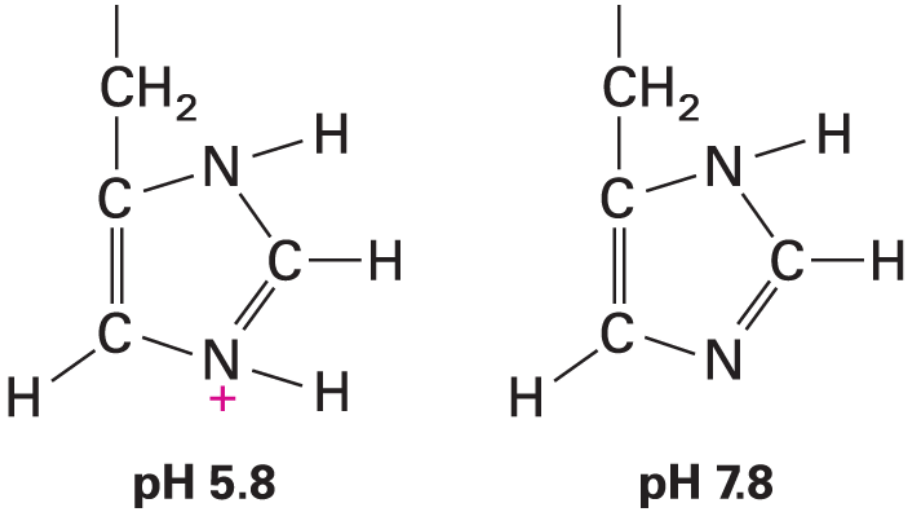
FIGURE 2-14 The 20 common amino acids used to build proteins. The side chain (R group; red) determines the distinctive properties of each amino acid and is the basis for grouping amino acids into three main categories: hydrophobic, hydrophilic, and special. Shown are the ionized forms that exist at the pH (~ 7) of the cytosol. In parentheses are the three-letter and one-letter abbreviations for each amino acid.

To understand the three-dimensional structures and functions of proteins, discussed in detail in Chapter 3, you must be familiar with some of the distinctive properties of

amino acids. These properties are determined in part by their side chains. You need not memorize the detailed structure of each type of side chain to understand how proteins work because amino acids can be classified into several broad categories based on the size, shape, charge, hydrophobicity (a measure of water solubility), and chemical reactivity of their side chains (Figure 2-14).

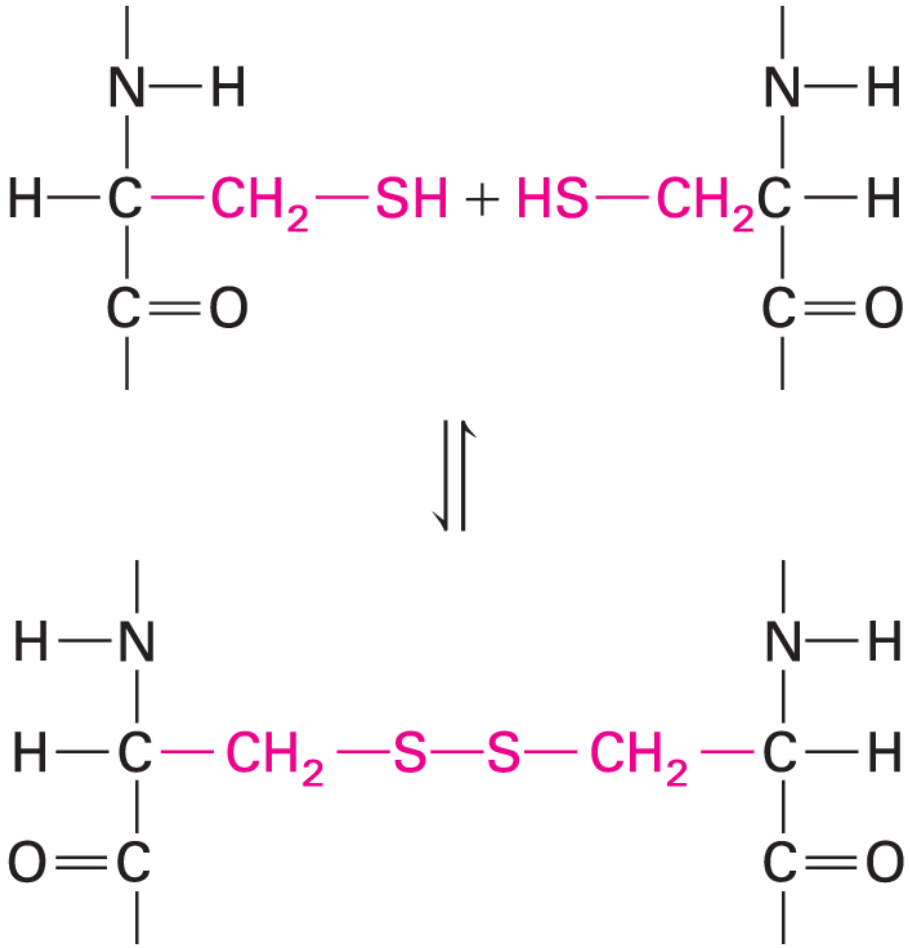
Amino acids with nonpolar side chains, called *hydrophobic amino acids*, are poorly soluble in water. The larger the nonpolar side chain, the more hydrophobic the amino acid. The side chains of *alanine, valine, leucine*, and *isoleucine* are linear or branched hydrocarbons that do not form a ring, and they are therefore called *aliphatic amino acids*. These amino acids are all nonpolar, as is *methionine*, which is similar to them except that it contains one sulfur atom. *Phenylalanine, tyrosine*, and *tryptophan* have large, hydrophobic, aromatic rings in their side chains. In later chapters, we will see in detail how hydrophobic side chains under the influence of the hydrophobic effect often pack into the interior of proteins or line the surfaces of proteins that are embedded within hydrophobic regions of biomembranes.

Amino acids with polar side chains are called *hydrophilic amino acids*. The most hydrophilic of these amino acids is the subset with side chains that are charged (ionized) at the pH typical of biological fluids ($\sim$7) both inside and outside the cell (see Section 2.3). *Arginine* and *lysine* have positively charged side chains and are called *basic amino acids; aspartic acid* and *glutamic acid* have negatively charged side chains due to the carboxylic acid groups in their side chains (their charged forms are called *aspartate* and *glutamate*) and are called *acidic amino acids*. A fifth amino acid, *histidine*, has a side chain containing a ring with two nitrogens, called *imidazole*. The imidazole ring can shift from being positively charged to uncharged in response to small changes in the acidity of its environment (defined as pH, see Section 2.3):



The activities of many proteins are modulated by shifts in environmental acidity (pH) through protonation or deprotonation of histidine side chains. *Asparagine* and *glutamine* are uncharged amino acids but have polar side chains containing amide groups with extensive hydrogen-bonding capacities. Similarly, *serine* and *threonine* are uncharged but have polar hydroxyl groups, which also participate in hydrogen bonds with other polar molecules.

Finally, *cysteine, glycine,* and *proline* play special roles in proteins because of the unique properties of their side chains. The side chain of cysteine contains a reactive *sulfhydryl group* (—*SH*). On release of a proton (H^+), a sulfhydryl group is converted into a thiolate anion (S^-). Thiolate anions in certain enzymes can play important roles in catalysis, notably in some proteases that destroy proteins. In many proteins, especially those secreted from cells or that extend outward from the plasma membrane into the extracellular fluid, each of two adjacent sulfhydryl groups can be oxidized, each releasing a proton and an electron, to form a covalent *disulfide bond* (—*S*—*S*—):



Disulfide bonds serve to cross-link regions within a single polypeptide chain (intramolecular cross-linking) or between two separate chains (intermolecular cross-linking). Disulfide bonds stabilize the folded structure of some proteins.

The smallest amino acid, *glycine*, has a single hydrogen atom as its R group. Its small size allows it to fit into tight spaces. We will see an example of this tight fit when we examine the structure of the most abundant protein in humans, *collagen* (see Chapter 20).

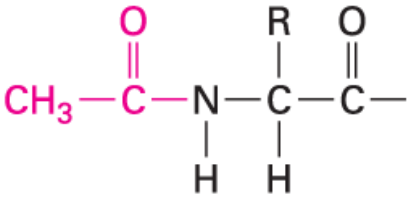
Unlike the side chains of the other common amino acids, the side chain of proline (pronounced *pro-leen*) bends around to form a ring by covalently bonding to the nitrogen atom in the amino group attached to the C_α. As a result, proline is very rigid, and its amino group is not available for typical hydrogen bonding. The presence of

proline in a protein creates a fixed kink in the polymer chain, limiting how it can fold in the vicinity of the proline residue.

Some amino acids are more abundant in proteins than others. Cysteine, tryptophan, and methionine are not common amino acids: together, they constitute approximately 5 percent of the amino acids in a typical protein. Four amino acids — leucine (hydrophobic), serine (polar), lysine (positively charged), and glutamic acid (negatively charged) — are the most abundant amino acids, constituting 32 percent of all the residues in a typical protein. However, the amino acid compositions of particular proteins may vary widely from these values.

Humans and other mammals can synthesize 11 of the 20 amino acids. The other nine are called *essential amino acids* and must be included in the diet to permit normal protein production. These essential amino acids are phenylalanine, valine, threonine, tryptophan, isoleucine, methionine, leucine, lysine, and histidine. Adequate provision of these essential amino acids in feed is key to the livestock industry. Indeed, a genetically engineered variety of corn with a high lysine content is now in use as an enhanced feed to promote the growth of animals. ■

Although cells use the 20 amino acids shown in Figure 2-14 in the *initial* synthesis of proteins, analysis of cellular proteins reveals that they contain over 100 different amino acids. The difference is due to the covalent modification by the addition of a variety of chemical groups to some amino acids after they are incorporated into proteins (Figure 2-15). One important modification is the addition of acetyl groups (CH_3CO) to amino acids, a process known as acetylation. Acetylation of the amino group of the N-terminal residue is the most common form of amino acid covalent modification, affecting an estimated 80 percent of all proteins:



Acetylated N-terminus

This modification may play an important role in controlling the life span of proteins within cells because many non-acetylated proteins are rapidly degraded.

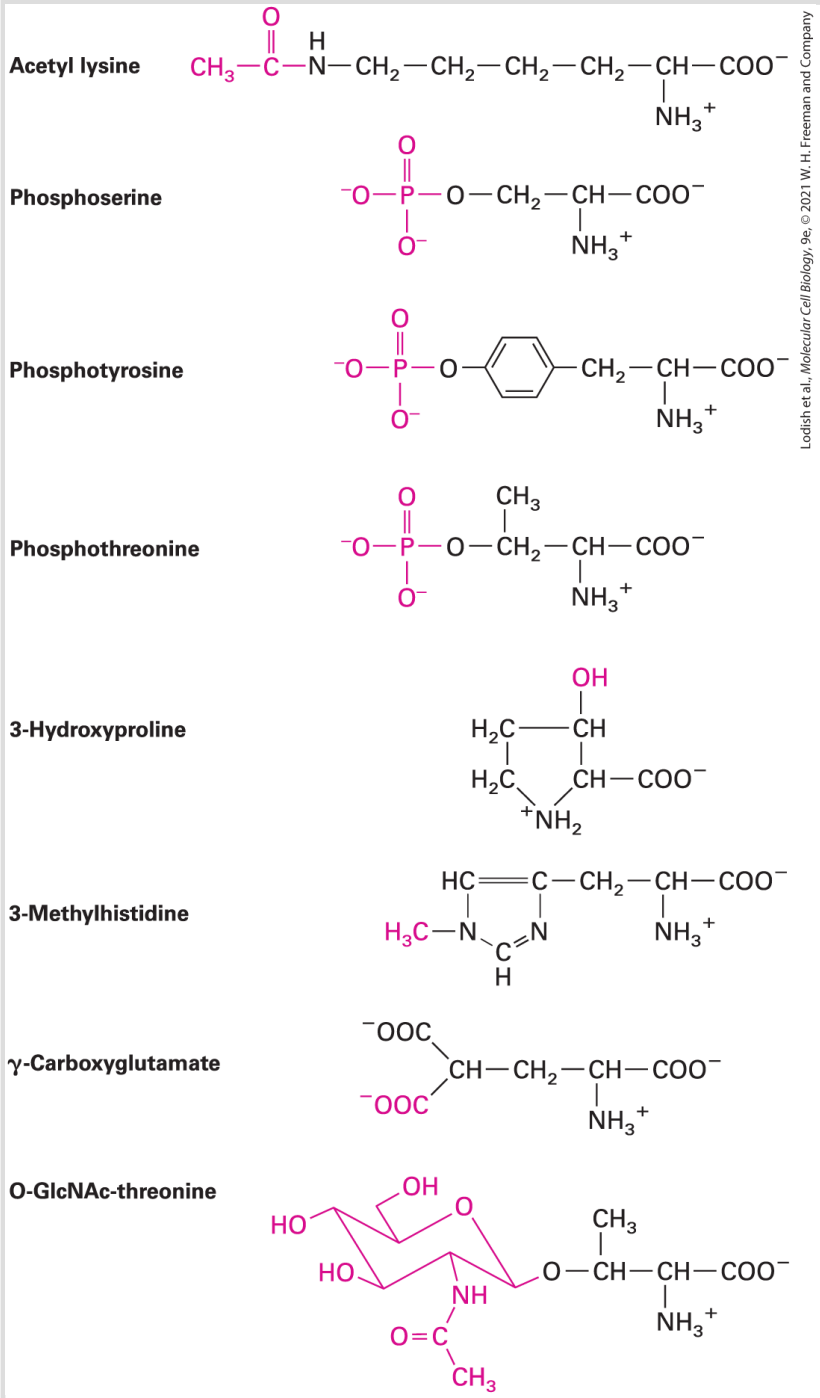


FIGURE 2-15 Common modifications of amino acid side chains in proteins. These modified residues and numerous others are formed by addition of various chemical groups (red) to the amino acid side chains during or after synthesis of a polypeptide chain.

Another important modification is the addition of a phosphate group (PO_4) to hydroxyl groups in serine, threonine, and tyrosine residues, a process known as

phosphorylation. We will encounter numerous examples of proteins whose activity is regulated by reversible phosphorylation and dephosphorylation. A well-known site of phosphorylation in bacteria, fungi, and plants is the nitrogen in the side chain of histidine, but the prevalence and functional significance of this modification in higher eukaryotes, such as mammals, has not been well-defined. Like phosphorylation and dephosphorylation, controlled methylation and demethylation of arginine and lysine side chains on proteins are important regulatory processes. For example, methylation of these side chains on proteins called *histones* is an important regulator of gene expression in eukaryotes (see Chapter 8). The side chains of asparagine, serine, and threonine are sites for glycosylation, the attachment of linear and branched carbohydrate chains. Many secreted proteins and membrane proteins contain glycosylated residues, and the reversible modification of hydroxyl groups on specific serines and threonines by a sugar called *N-acetylglucosamine* also appears to regulate protein activities inside cells. Other amino acid modifications found in selected proteins include the hydroxylation of proline and lysine residues in collagen (see Chapter 20), the methylation of histidine residues in membrane receptors, and the γ-carboxylation of glutamate in blood-clotting factors such as prothrombin. Deamidation of asparagine and glutamine to form the corresponding acidic amino acids, aspartate and glutamate, also occurs.

Five Different Nucleotides Are Used to Build Nucleic Acids

Two types of chemically similar nucleic acids, **DNA (deoxyribonucleic acid)** and **RNA (ribonucleic acid)**, are the cell's principal molecules that carry genetic information. RNAs can also help catalyze chemical reactions. DNA and RNA polymers are built from monomers called **nucleotides**. All nucleotides have a common three-part structure: a phosphate group is linked by a phosphoester bond to a pentose (five-carbon) sugar, and the pentose sugar in turn is linked to a nitrogen- and carbon-containing ring structure commonly referred to as a *base* (Figure 2-16a). In RNA, the pentose is *r*ibose; in DNA, it is *d*eoxyribose, which has a proton, rather than a hydroxyl group, at position 2′ (Figure 2-16b). (We describe the structures of sugars in

more detail below) The bases *adenine, guanine*, and *cytosine* (Figure 2-17) are found in both DNA and RNA; *thymine* is found only in DNA, and *uracil* is found only in RNA. The bases are often abbreviated *A, G, C, T*, and *U*, respectively; these same single-letter abbreviations are also commonly used to denote all the nucleotides in nucleic acid polymers.

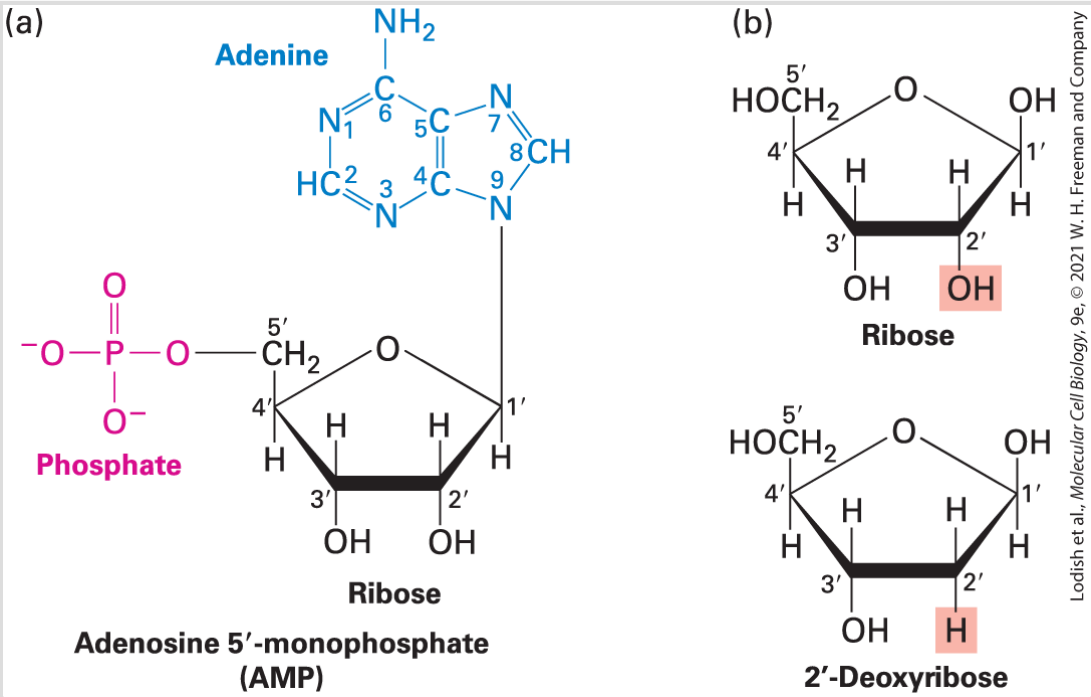


FIGURE 2-16 Common structure of nucleotides. (a) Adenosine 5′-monophosphate (AMP), a nucleotide present in RNA. By convention, the carbon atoms of the pentose sugar in nucleotides are numbered with primes. In natural nucleotides, the 1′ carbon is joined by a β linkage to the base (in this case, adenine); both the base (blue) and the phosphate on the 5′ hydroxyl (red) extend above the plane of the sugar ring. (b) Ribose and deoxyribose, the pentoses in RNA and DNA, respectively.

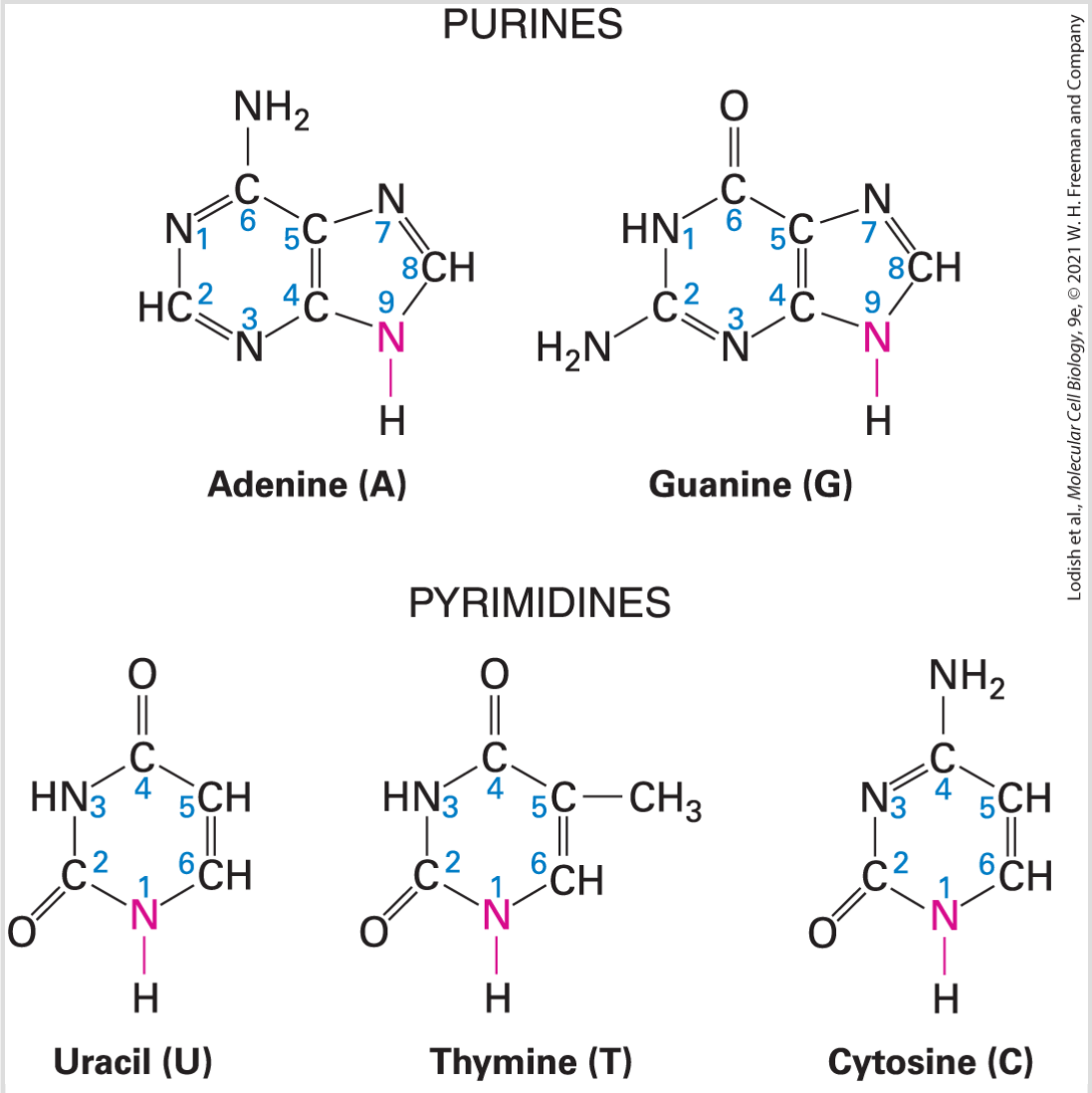
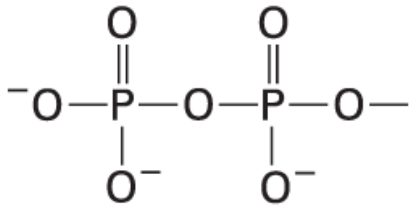


FIGURE 2-17 Chemical structures of the principal bases in nucleic acids. In nucleic acids and nucleotides, nitrogen 9 of purines and nitrogen 1 of pyrimidines (red) are bonded to the 1′ carbon of ribose or deoxyribose. U is found only in RNA, and T is found only in DNA. Both RNA and DNA contain A, G, and C.

Adenine and guanine are **purines**, bases that contain a pair of fused rings; cytosine, thymine, and uracil are **pyrimidines**, bases that contain a single ring (see Figure 2-17). In nucleotides, the 1′ carbon atom of the sugar (ribose or deoxyribose) is attached to the nitrogen atom at position 9 of a purine (N_9) or at position 1 of a pyrimidine (N_1). For example, the connection of deoxyribose to a purine base in the nucleotide adenosine 5′-monophosphate (AMP) is shown in Figure 2-16a. Due to the phosphate group, nucleotides are acidic: under normal intracellular conditions the phosphate

group releases hydrogen ions (H^+), becoming negatively charged (see Figure 2-16a). Most nucleic acids in cells are associated with proteins, which form both ionic interactions with the negatively charged phosphates and other noncovalent interactions with other portions of the nucleic acid.

Cells and extracellular fluids in organisms contain small concentrations of **nucleosides**, combinations of a base and a sugar without a phosphate group. Nucleotides are nucleosides that have one, two, or three phosphate groups covalently attached to the ribose via an ester bond. An ester bond is a covalent bond linking an acid, such as a carboxylic acid or a phosphoric acid, with an alcohol. The formation of such a bond, called *esterification*, is accompanied by the release of a hydroxyl (—OH) group from the acid and a proton (—H) from the hydroxyl group on the alcohol, which together form a water molecule. In a nucleotide, a phosphoric acid is esterified with the 5′ hydroxyl group of the ribose (see Figure 2-16a). Nucleoside monophosphates have a single esterified phosphate (see Figure 2-16a); nucleoside diphosphates contain a pyrophosphate group:



Pyrophosphate

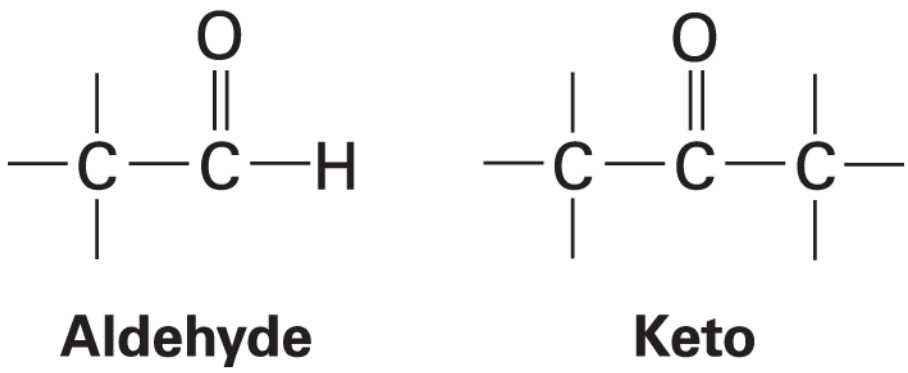
Nucleoside triphosphates have a third phosphate. Table 2-3 lists the names of the nucleosides and nucleotides in nucleic acids and the various forms of nucleoside phosphates. Later in this chapter (Section 2.4), we will consider in detail the structure of the nucleoside triphosphate ATP (see Figures 2-1d and 2-31). The nucleoside triphosphates are used in the synthesis of nucleic acids, as we explain in Chapter 5. Among their other functions in the cell, GTP participates in intracellular signaling and acts as an energy reservoir, particularly in protein synthesis, and ATP is the most widely used biological energy carrier.

TABLE 2-3 • Terminology of Nucleosides and Nucleotides

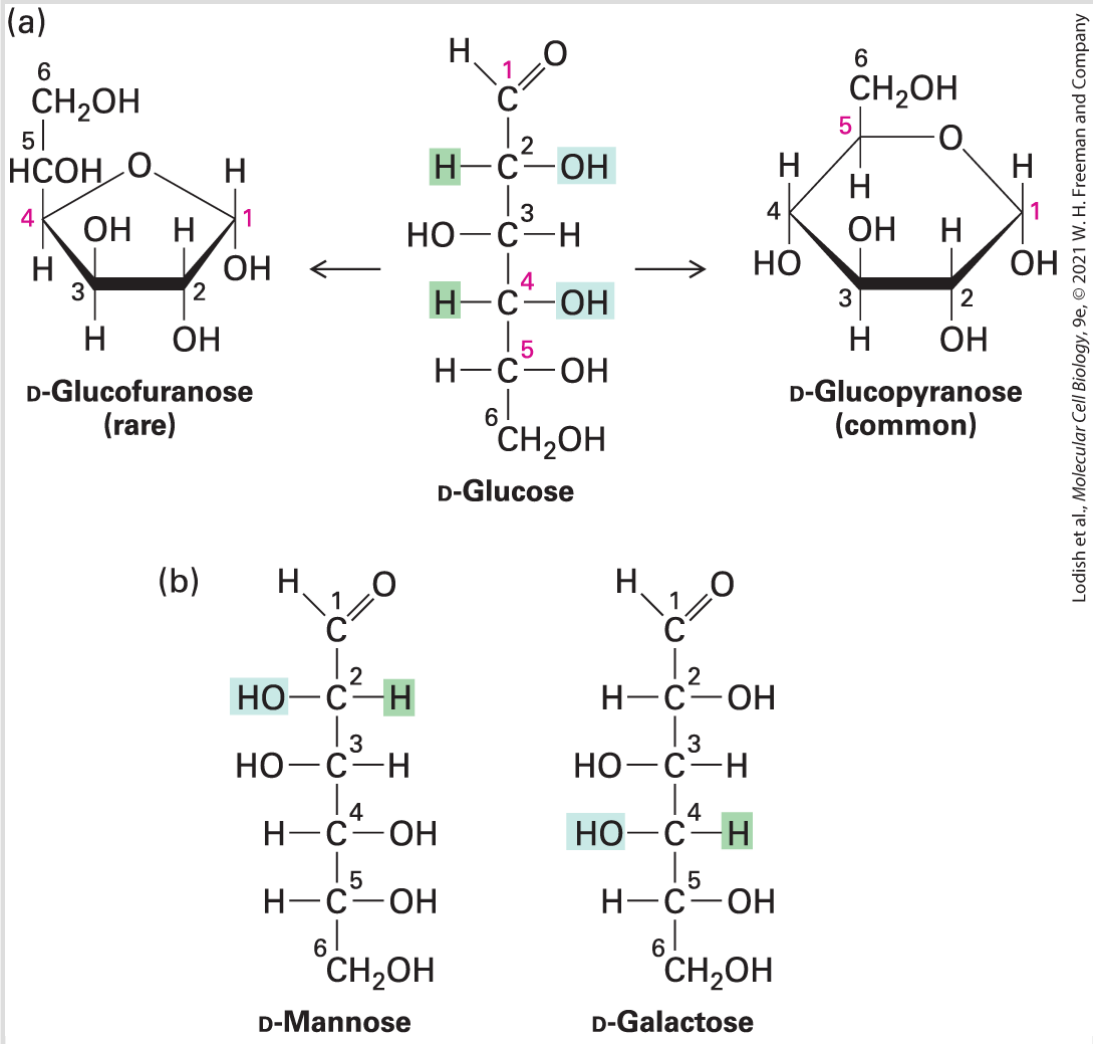
Bases		Purines		Pyrimidines	
		Adenine (A)	**Guanine (G)**	**Cytosine (C)**	**Uracil (U)** **Thymine (T)**
Nucleosides	in RNA	Adenosine	Guanosine	Cytidine	Uridine
	in DNA	Deoxyadenosine	Deoxyguanosine	Deoxycytidine	Deoxythymidine
Nucleotides	in RNA	Adenylate	Guanylate	Cytidylate	Uridylate
	in DNA	Deoxyadenylate	Deoxyguanylate	Deoxycytidylate	Deoxythymidylate
Nucleoside monophosphates		AMP	GMP	CMP	UMP
Nucleoside diphosphates		ADP	GDP	CDP	UDP
Nucleoside triphosphates		ATP	GTP	CTP	UTP
Deoxynucleoside mono-, di-, and triphosphates		dAMP, etc.	dGMP, etc.	dCMP, etc.	dTMP, etc.

Monosaccharides Covalently Assemble into Linear and Branched Polysaccharides

The building blocks of the polysaccharides are the simple sugars, or **monosaccharides**. Monosaccharides are **carbohydrates**, which are literally covalently bonded combinations of carbon atoms ("carbo") and water molecules ("hydrate") in a one-to-one ratio $(CH_2O)_n$, where n equals 3, 4, 5, 6, or 7. **Hexoses** ($n = 6$) and **pentoses** ($n = 5$) are the most common monosaccharides. All monosaccharides contain hydroxyl (—OH) groups and either an aldehyde or a keto group:



Many biologically important sugars are hexoses, including glucose, mannose, and galactose (Figure 2-18). Mannose is identical to glucose except that the orientation of the groups bonded to carbon 2 is reversed — they are stereoisomers (also called *epimers*). Similarly, galactose, another hexose, differs from glucose only in the orientation of the groups attached to carbon 4. Interconversion of glucose and mannose or galactose requires the breaking and making of covalent bonds; such reactions are carried out by enzymes called *epimerases*.



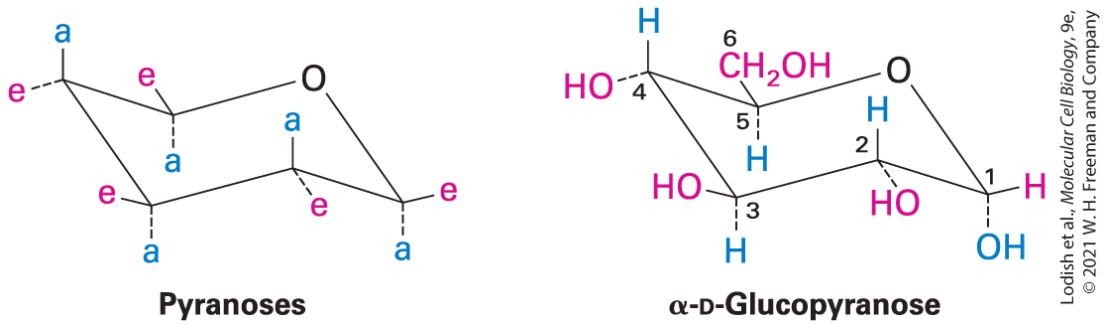
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 2-18 Chemical structures of hexoses. All hexoses have the same chemical formula ($C_6H_{12}O_6$) and contain an aldehyde or a keto group. (a) The ring forms of D-glucose are generated from the linear molecule by reaction of the aldehyde at carbon 1 with the hydroxyl on carbon 5 or carbon 4. The three forms are readily interconvertible, although the pyranose form (*right*) predominates in biological systems. (b) In D-mannose and D-galactose, the configuration of the H (green) and OH (blue) bound to one carbon atom differs from that in glucose. These sugars, like glucose, exist primarily as pyranoses (six-member rings).

D-Glucose ($C_6H_{12}O_6$) is one of the principal external sources of energy for most cells in complex multicellular organisms. It can exist in three different forms: a linear structure and two different hemiacetal ring structures (Figure 2-18a). If the aldehyde group on carbon 1 combines with the hydroxyl group on carbon 5, the result is the formation of D-glucopyranose, a hemiacetal containing a six-atom (usually called a

six-membered) ring. In the α anomer of D-glucopyranose, the hydroxyl group attached to carbon 1 points *downward* from the ring, as shown in Figure 2-18a; in the β anomer, this hydroxyl points *upward*. In aqueous solution, the α and β anomers readily interconvert spontaneously; at equilibrium there is about one-third α anomer and two-thirds β, with very little of the open-chain form. Because enzymes can distinguish between the α and β anomers of D-glucose, these forms have distinct biological roles. Condensation of the hydroxyl group on carbon 4 of the linear glucose with its aldehyde group results in the formation of D-glucofuranose, a hemiacetal containing a five-membered ring. Although all three forms of D-glucose exist in biological systems, the pyranose (six-membered ring) form is by far the most abundant.

The pyranose ring in Figure 2-18a is depicted as planar. In fact, because of the tetrahedral geometry of each of the five carbon atoms, the most stable conformation of a pyranose ring has a nonplanar, chairlike shape. In this conformation, each bond from a ring carbon to a nonring atom (e.g., H or O) is either nearly perpendicular to the ring — referred to as *axial (a)* — or nearly in the plane of the ring — referred to as *equatorial (e)*:



Disaccharides, formed from two monosaccharides, are the simplest polysaccharides. The disaccharide lactose, composed of galactose and glucose, is the major sugar in milk; the disaccharide sucrose, composed of glucose and fructose, is a principal product of plant photosynthesis and is refined into common table sugar (Figure 2-19).

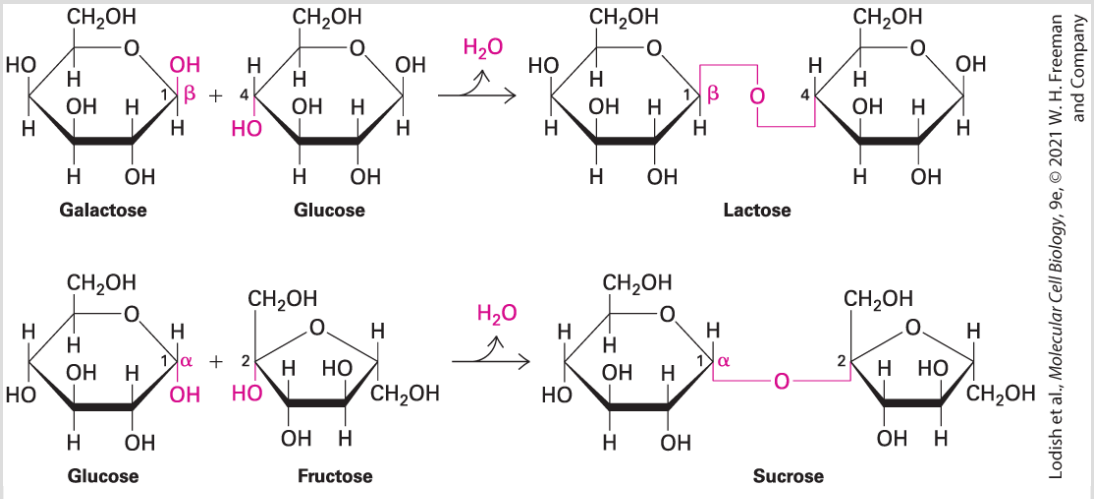
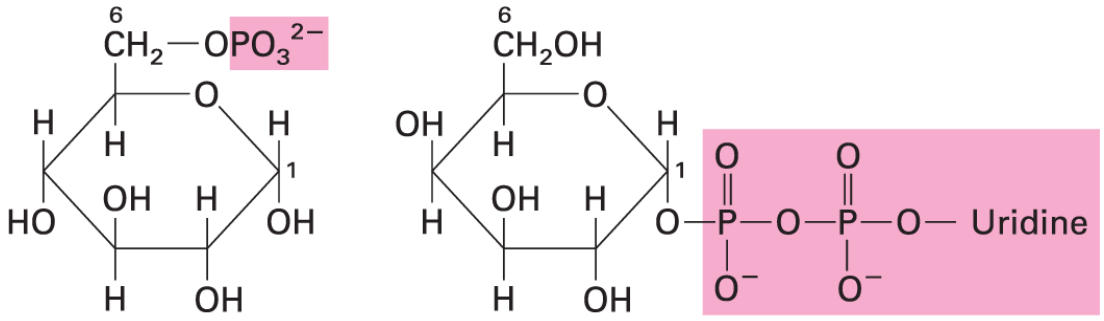


FIGURE 2-19 Formation of the disaccharides lactose and sucrose. In any glycosidic linkage, the anomeric carbon of one sugar molecule (in either the α or β conformation) is linked to a hydroxyl oxygen on another sugar molecule. The linkages are named accordingly; thus lactose contains a $\beta(1 \rightarrow 4)$ bond, and sucrose contains an $\alpha(1 \rightarrow 2)$ bond.

Larger polysaccharides contain dozens to hundreds of monosaccharide units. They can function as reservoirs for glucose, as structural components, or as adhesives that help hold cells together in tissues. The most common storage carbohydrate in animal cells is **glycogen**, a very long, highly branched polymer of glucose. As much as 10 percent of the liver by weight can be glycogen. The primary storage carbohydrate in plant cells, **starch**, is also a glucose polymer. It occurs in an unbranched form (amylose) and a lightly branched form (amylopectin). Both glycogen and starch are composed of the α anomer of glucose. In contrast, **cellulose**, the major constituent of plant cell walls, is an unbranched polymer of the β anomer of glucose. Cellulose confers stiffness to many plant structures (see Chapter 20). Human digestive enzymes can hydrolyze the α glycosidic bonds in starch but not the β glycosidic bonds in cellulose. Many species of plants, bacteria, and molds produce cellulose-degrading enzymes. Cows and termites can break down cellulose because they harbor cellulose-degrading bacteria in their gut. Bacterial cell walls consist of **peptidoglycan** rather than cellulose. A peptidoglycan has a polysaccharide chain cross-linked by peptide cross-bridges, which confer rigidity and cell shape. Human tears and gastrointestinal fluids contain lysozyme, an enzyme capable of hydrolyzing peptidoglycan in the bacterial cell wall.

The enzymes that make the glycosidic bonds, linking monosaccharides into polysaccharides, are specific for the α or β anomer of one sugar and a particular hydroxyl group on the other. In principle, any two sugar molecules can be linked in a variety of ways because each monosaccharide has multiple hydroxyl groups that can participate in the formation of glycosidic bonds. Furthermore, any one monosaccharide has the potential to be linked to more than one other monosaccharide, thus generating a branch point and nonlinear polymers. Glycosidic bonds are usually formed between the growing polysaccharide chain and a covalently modified form of a monosaccharide. Such modifications include the addition of a phosphate (e.g., glucose 6-phosphate) or a nucleotide (e.g., UDP-galactose):



Glucose 6-phosphate **UDP-galactose**

The epimerase enzymes that interconvert different monosaccharides often do so using the nucleotide sugars rather than the unmodified, or *free*, sugars.

Many complex polysaccharides contain modified sugars that are covalently linked to various small groups, particularly amino, sulfate, and acetyl groups. Such modifications are abundant in **glycosaminoglycans**, major polysaccharide components of the extracellular matrix that we describe in Chapter 20.

As you can see here and in many examples throughout this book, a relatively simple set of monosaccharides can be assembled into a wide variety of polysaccharides with dramatically different biological functions.

Phospholipids Associate Noncovalently to Form the Basic Bilayer Structure of Biomembranes

Biomembranes are large, flexible sheets with a two-ply, or bilayer, structure. They serve as the boundaries of cells and their intracellular organelles and form the outer surfaces of some viruses. Membranes literally define what is a cell (the outer membrane and the contents within the membrane) and what is not (the extracellular space outside the membrane). Unlike proteins, nucleic acids, and polysaccharides, membranes are assembled by the *noncovalent* association of their component building blocks. The primary building blocks of all biomembranes are **phospholipids**, whose physical properties are responsible for the formation of the sheet-like bilayer structure of membranes. In addition to phospholipids, biomembranes can contain a variety of other molecules, including cholesterol, glycolipids, and proteins. The structure and function of biomembranes will be described in detail in Chapter 10. Here we will focus on the phospholipids in biomembranes.

To understand the structure of a phospholipid molecule, we have to understand each of its component parts and how they are assembled. As we will see shortly, a phospholipid molecule consists of two long-chain, nonpolar fatty acid groups linked (usually by an ester bond) to small, highly polar groups. These polar groups typically include a short organic molecule such as glycerol (trihydroxy propane), a phosphate, and typically a small organic molecule (Figure 2-20).

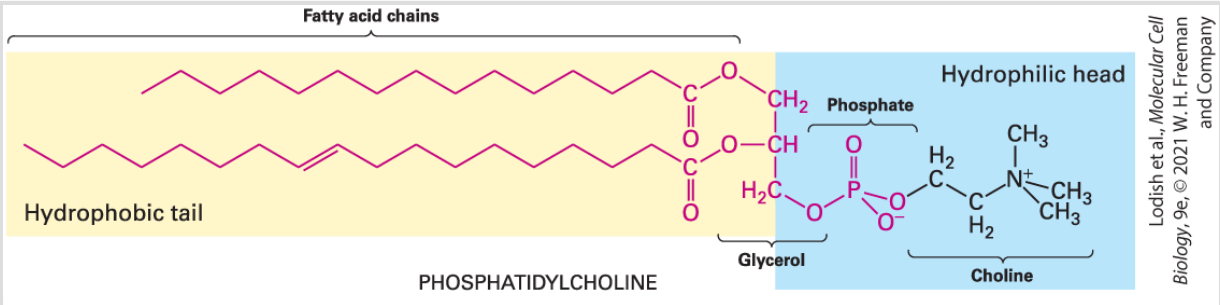


FIGURE 2-20 Phosphatidylcholine, a typical phosphoglyceride. All phosphoglycerides are amphipathic phospholipids, having a hydrophobic tail (yellow) and a hydrophilic head (blue) in which glycerol is linked via a

phosphate group to an alcohol. Either or both of the fatty acyl side chains in a phosphoglyceride may be saturated or unsaturated. In phosphatidic acid (red), the simplest phospholipid, the phosphate is not linked to an alcohol.

Fatty acids consist of a hydrocarbon chain attached to a carboxyl group (—COOH). Like glucose, fatty acids are an important energy source for many cells (see Chapter 12). They differ in length, although the predominant fatty acids in cells have an even number of carbon atoms, usually 14, 16, 18, or 20. The major fatty acids in phospholipids are listed in Table 2-4. Fatty acids are often designated by the abbreviation *Cx:y*, where *x* is the number of carbons in the chain and *y* is the number of double bonds. Fatty acids containing 12 or more carbon atoms are nearly insoluble in aqueous solutions because of their long hydrophobic hydrocarbon chains.

TABLE 2-4 • Fatty Acids That Predominate in Phospholipids

Common Name of Acid (ionized form in parentheses)	Abbreviation	Chemical Formula
Saturated Fatty Acids		
Myristic (myristate)	C14:0	$CH_3(CH_2)_{12}COOH$
Palmitic (palmitate)	C16:0	$CH_3(CH_2)_{14}COOH$
Stearic (stearate)	C18:0	$CH_3(CH_2)_{16}COOH$
Unsaturated Fatty Acids		
Oleic (oleate)	C18:1	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$
Linoleic (linoleate)	C18:2	$CH_3(CH_2)_4CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$
Arachidonic (arachidonate)	C20:4	$CH_3(CH_2)_4(CH{=}CHCH_2)_3CH{=}CH(CH_2)_3COOH$

Fatty acids in which all the carbon-carbon bonds are single bonds — that is, the fatty acids have no carbon-carbon double bonds — are said to be **saturated**; those with at least one carbon-carbon double bond are called **unsaturated**. Unsaturated fatty acids with more than one carbon-carbon double bond are referred to as **polyunsaturated**. Two *essential* polyunsaturated fatty acids, linoleic acid (C18:2) and linolenic acid

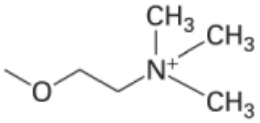
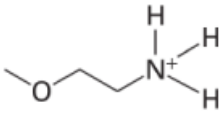
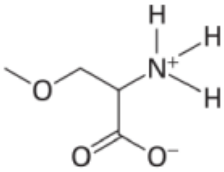
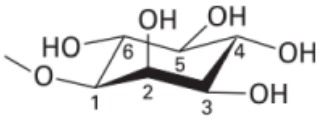
(C18:3), cannot be synthesized by mammals and must be supplied in their diet. Mammals can synthesize other common fatty acids.

In phospholipids, fatty acids are covalently attached to another molecule by esterification. In the combined molecule formed by this reaction, the part derived from the fatty acid is called an *acyl group*, or *fatty acyl group*. The most common phospholipids are the **phosphoglycerides**, which contain two acyl groups attached to two of the three hydroxyl groups of glycerol (see Figure 2-20). Other types of phospholipids are described in Chapter 10 (also see Figure 10-8).

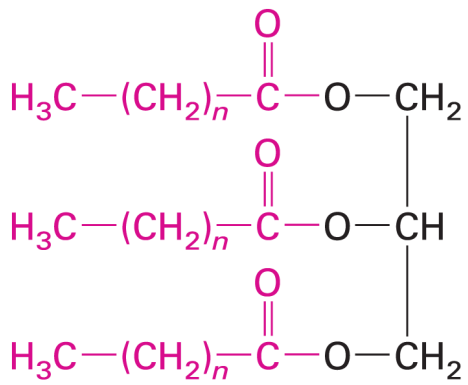
In phosphoglycerides, one hydroxyl group of the glycerol is esterified to phosphate while the other two are normally esterified to fatty acids. The simplest phospholipid, phosphatidic acid, contains only these components. Phospholipids such as phosphatidic acids are not only membrane building blocks but also important signaling molecules. Lysophosphatidic acid, for example, can be a potent inducer of cell division (called a **mitogen**). In this phospholipid, the acyl chain at the 2 position (attached to the hydroxyl group on the central carbon of the glycerol) has been removed, so the molecule is relatively water soluble. In most phospholipids found in membranes, the phosphate group is also esterified to a hydroxyl group on another hydrophilic compound. In phosphatidylcholine, for example, the phosphate is covalently attached to the small molecule called *choline* (see Figure 2-20).

The negatively charged phosphate, as well as the charged or polar groups esterified to it, can interact strongly with water. The phosphate and its associated esterified group constitute the *head* group of a phospholipid, which is hydrophilic, whereas the fatty acyl chains, the *tails*, are hydrophobic. Other common phosphoglycerides and associated head groups are shown in Table 2-5. Molecules such as phospholipids that have both hydrophobic and hydrophilic regions are called *amphipathic*. In Chapter 10, we will see how the amphipathic properties of phospholipids allow their assembly into sheet-like bilayers in which the fatty acyl tails point into the center of the sheet and the head groups point outward toward the aqueous environment (see Figure 2-13 bottom).

TABLE 2-5 • Common Phosphoglycerides and Head Groups

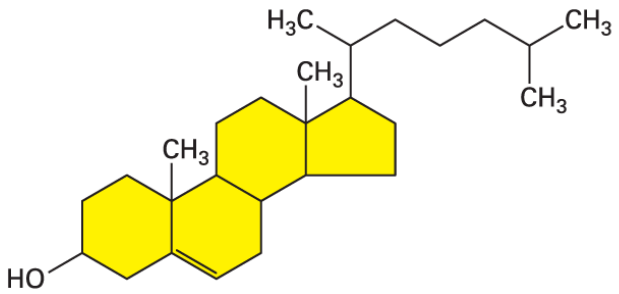
Common Phosphoglycerides	Head Group
Phosphatidylcholine	 Choline
Phosphatidylethanolamine	 Ethanolamine
Phosphatidylserine	 Serine
Phosphatidylinositol	 Inositol

Fatty acyl groups also can be covalently linked in other fatty molecules, including **triacylglycerols** (or **triglycerides**), which contain three acyl groups esterified to glycerol:

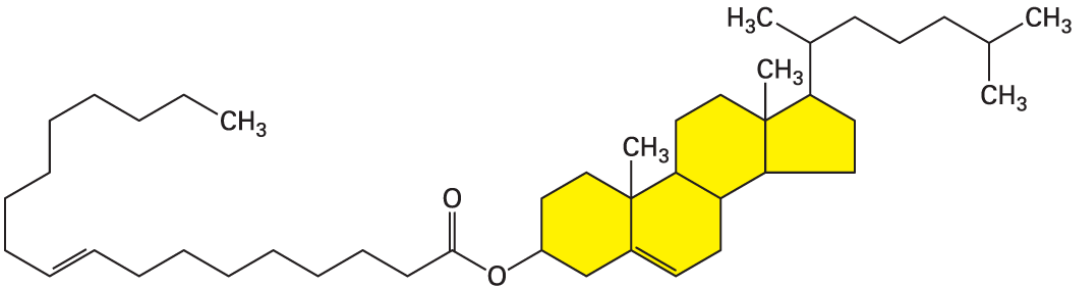


Triacylglycerol

They can also be covalently attached to the very hydrophobic molecule cholesterol, an alcohol, to form cholesteryl esters:



Cholesterol

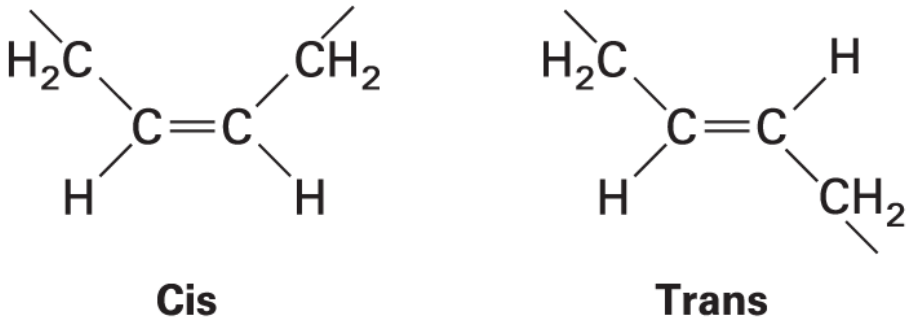


Cholesteryl ester

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Triglycerides and cholesteryl esters are extremely water-insoluble molecules in which fatty acids and cholesterol are either stored or transported. Triglycerides are the storage form of fatty acids in the fat cells of adipose tissue and are the principal components of dietary fats. Cholesteryl esters and triglycerides are transported between tissues through the bloodstream in specialized carriers called *lipoproteins* (see Chapter 14).

We saw previously that the fatty acids, which are key components of both phospholipids and triglycerides, can be either saturated or unsaturated. An important consequence of the carbon-carbon double bond (C=C) in an unsaturated fatty acid is that two stereoisomeric configurations, cis and trans, are possible around each of these bonds:



A cis double bond introduces a rigid kink in the otherwise flexible straight acyl chain of a saturated fatty acid (Figure 2-21). In general, the unsaturated fatty acids in biological systems contain only cis double bonds. Saturated fatty acids without the kink can pack together tightly and so have higher melting points than unsaturated fatty acids. The main fatty molecules in butter are triglycerides with saturated fatty acyl chains, which is why butter is usually solid at room temperature. Unsaturated fatty acids or fatty acyl chains with the cis double bond kink cannot pack as closely together as saturated fatty acyl chains. Thus vegetable oils, composed of triglycerides with unsaturated fatty acyl groups, usually are liquid at room temperature. Vegetable and similar oils may be partially hydrogenated to convert some of their unsaturated fatty acyl chains to saturated fatty acyl chains. As a consequence, the hydrogenated vegetable oil can be molded into solid sticks of margarine. A by-product of the hydrogenation reaction is the conversion of some of the fatty acyl chains into trans fatty acids, popularly called *trans fats*. These trans fats, found in partially hydrogenated margarine and other food products, are not natural. Saturated and trans fatty acids have similar physical properties; for example, they tend to be solids at room temperature. Their consumption, relative to the consumption of unsaturated fats, is associated with increased plasma cholesterol levels and is discouraged by some nutritionists. ■

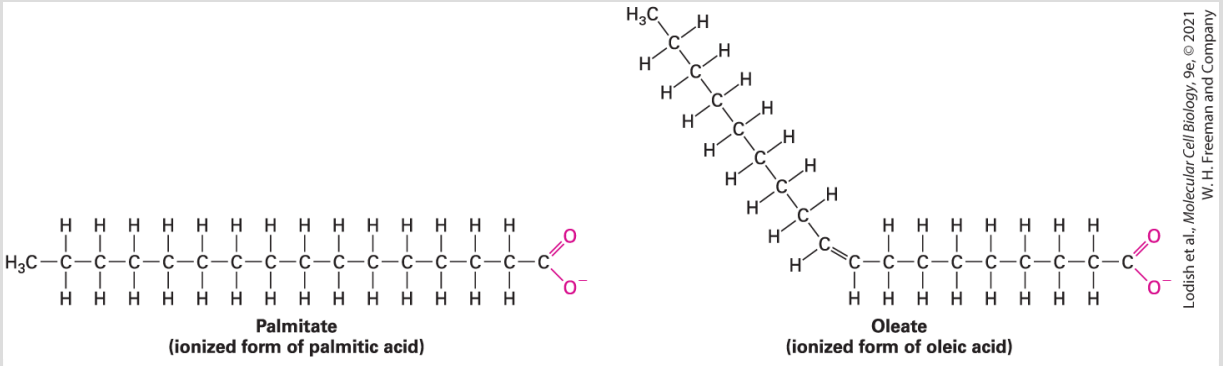


FIGURE 2-21 The effect of a double bond on the shape of fatty acids. Shown are chemical structures of the ionized form of palmitic acid, a saturated fatty acid with 16 C atoms, and oleic acid, an unsaturated one with 18 C atoms. In saturated fatty acids, the hydrocarbon chain is often linear; the cis double bond in oleate creates a rigid kink in the hydrocarbon chain.

KEY CONCEPTS OF SECTION 2.2

Chemical Building Blocks of Cells

- Macromolecules are polymers of monomer subunits linked together by covalent bonds via dehydration reactions. Three major types of macromolecules are found in cells: proteins, composed of amino acids linked by peptide bonds; nucleic acids, composed of nucleotides linked by phosphodiester bonds; and polysaccharides, composed of monosaccharides (sugars) linked by glycosidic bonds (see Figure 2-13). Phospholipids, the fourth major chemical building block, assemble noncovalently into biomembranes.
- Differences in the size, shape, charge, hydrophobicity, and reactivity of the side chains of the 20 common amino acids determine the chemical and structural properties of proteins (see Figure 2-14). The three general categories into which the side chains fall are hydrophobic, hydrophilic (basic, acidic, polar), and special (see Figure 2-14). It is helpful to remember which amino acids fall into each of these categories.
- The bases in the nucleotides composing DNA and RNA are carbon- and nitrogen-containing rings attached to a pentose sugar. They form two groups: the purines, with two rings — adenine (A) and guanine (G) — and the pyrimidines, with one ring — cytosine (C), thymine (T), and uracil (U) (see Figure 2-17). A, G, T, and C are found in DNA, and A, G, U, and C are found in RNA.
- Glucose and other hexoses can exist in three forms: an open-chain linear structure, a six-member (pyranose) ring, and a five-member (furanose) ring (see Figure 2-18). In biological systems, the pyranose form of D-glucose predominates.
- Glycosidic bonds are formed between either the α or the β anomer of one sugar and a hydroxyl group on another sugar, leading to formation of disaccharides and other polysaccharides (see Figure 2-19).
- Phospholipids are amphipathic molecules with a hydrophobic tail (often two fatty acyl chains) connected by a small organic molecule (often glycerol) to a hydrophilic head (see Figure 2-20).
- The long hydrocarbon chain of a fatty acid may be saturated (containing no carbon-carbon double bonds) or unsaturated (containing one or more double bonds). Fatty substances such as butter that have primarily saturated fatty acyl chains tend to be solid at room temperature, whereas unsaturated fats with cis double bonds have kinked chains that cannot pack closely together and so tend to be liquids at room temperature.

2.3 Chemical Reactions and Chemical Equilibrium

We now shift our discussion to chemical reactions in which bonds, primarily covalent bonds in *reactant* chemicals, are broken and new bonds are formed to generate reaction *products*. At any one time, several hundred different kinds of chemical reactions are occurring simultaneously in every cell, and many chemicals can, in principle, undergo multiple chemical reactions. Both the *extent* to which reactions can proceed and the *rate* at which they take place determine the chemical composition of cells. In this section, we discuss the concepts of equilibrium and steady state as well as dissociation constants and pH. These concepts will arise again and again throughout this textbook, so it is important for you to be familiar with them. In Section 2.4, we discuss how energy influences the extents and rates of chemical reactions.

A Chemical Reaction Is in Equilibrium When the Rates of the Forward and Reverse Reactions Are Equal

When reactants first mix together — before any products have been formed — the rate of the forward reaction to form products is determined in part by the reactants' initial concentrations, which

determine the likelihood of reactants bumping into one another and reacting (Figure 2-22). As the reaction products accumulate, the concentration of each reactant decreases, and so does the forward reaction rate. Meanwhile, some of the product molecules begin to participate in the reverse reaction, which re-forms the reactants. The ability of a reaction to go backward is called *microscopic reversibility*. The reverse reaction is slow at first but speeds up as the concentration of product increases. Eventually, the rates of the forward and reverse reactions become equal, so that the concentrations of reactants and products stop changing. The system is then said to be in **chemical equilibrium** (plural, *equilibria*).

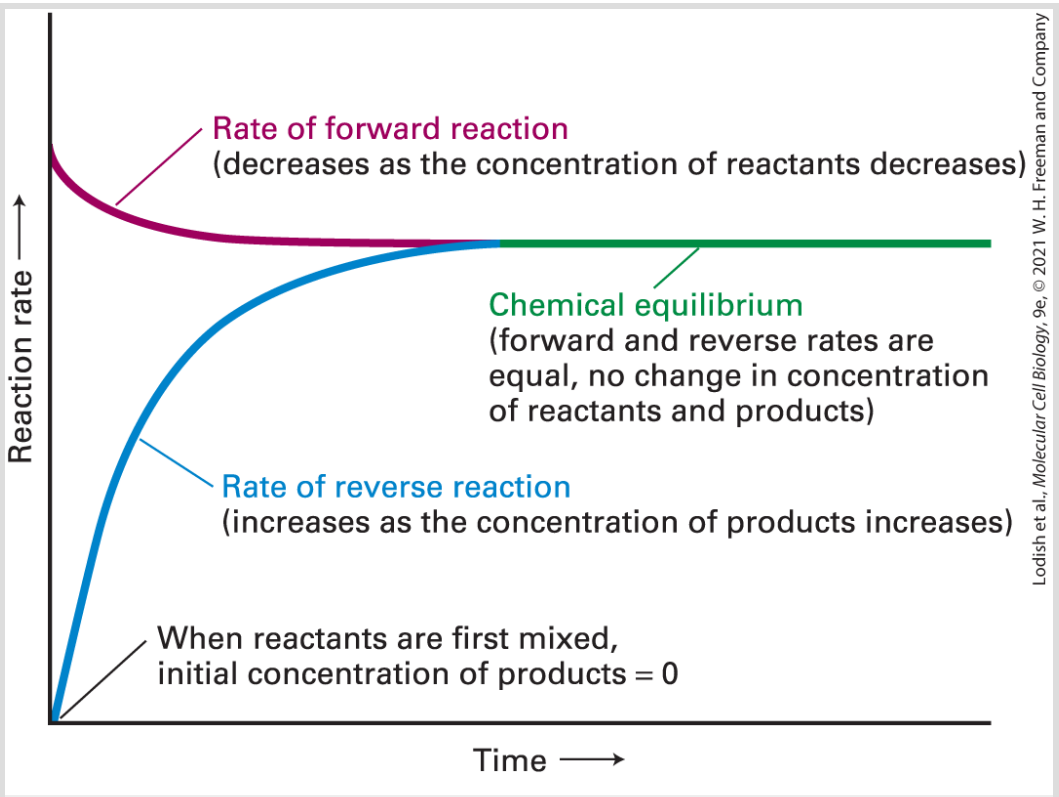


FIGURE 2-22 Time dependence of the rates of a chemical reaction. The forward and reverse rates of a reaction depend in part on the initial

concentrations of reactants and products. The net forward reaction rate slows as the concentration of reactants decreases, whereas the net reverse reaction rate increases as the concentration of products increases. At equilibrium, the rates of the forward and reverse reactions are equal, and the concentrations of reactants and products remain constant.

The ratio of the concentrations of the products to the concentrations of the reactants when they reach equilibrium, called the **equilibrium constant (K_{eq}),** is a fixed value. Thus K_{eq} provides a measure of the extent to which a reaction occurs by the time it reaches equilibrium. The rate of a chemical reaction can be increased by a **catalyst**, a chemical that accelerates the making and breaking of covalent bonds in other molecules but itself is not permanently changed during a reaction. Although a catalyst increases the reaction rate, it does not change the equilibrium constant (see Section 2.4).

The Equilibrium Constant Reflects the Extent of a Chemical Reaction

For any chemical reaction, K_{eq} depends on the chemical nature of the reactants and products, the temperature, and the pressure (particularly in reactions involving gases). Under standard physical conditions (25 $^{\circ}$C and 1 atm pressure for biological systems), K_{eq} is always the same for a given reaction, whether or not a catalyst is present.

Consider the general reaction with three reactants and three products,

$$\mathrm{aA} + \mathrm{bB} + \mathrm{cC} \rightleftharpoons \mathrm{zZ} + \mathrm{yY} + \mathrm{xX} \tag{2-1}$$

where capital letters represent particular molecules or atoms and lowercase letters represent the number of each in the reaction. For this reaction, the formula for the equilibrium constant is given by

$$K_{\mathrm{eq}} = \frac{[\mathrm{X}]^{\mathrm{x}}[\mathrm{Y}]^{\mathrm{y}}[\mathrm{Z}]^{\mathrm{z}}}{[\mathrm{A}]^{\mathrm{a}}[\mathrm{B}]^{\mathrm{b}}[\mathrm{C}]^{\mathrm{c}}} \tag{2-2}$$

where brackets denote the concentrations of the molecules. In Equation 2-2, the concentrations of reactants and products are those present at equilibrium. The rate of the forward reaction (left to right in Equation 2-1) is

$$\mathrm{Rate}_{\mathrm{forward}} = k_{\mathrm{f}}[\mathrm{A}]^{\mathrm{a}}[\mathrm{B}]^{\mathrm{b}}[\mathrm{C}]^{\mathrm{c}}$$

where k_{f} is the **rate constant** for the forward reaction. Similarly, the rate of the reverse reaction (right to left in Equation 2-1) is

$$\mathrm{Rate}_{\mathrm{reverse}} = k_{\mathrm{r}}[\mathrm{X}]^{\mathrm{x}}[\mathrm{Y}]^{\mathrm{y}}[\mathrm{Z}]^{\mathrm{z}}$$

where k_r is the rate constant for the reverse reaction. These reaction rate equations apply whether or not the reaction has reached equilibrium. It is important to remember that the forward and reverse rates of a reaction can change because of changes in reactant or product concentrations, yet at the same time the forward and reverse rate constants do not change; hence the name *constant*. Confusing rates and rate constants is a common error. At equilibrium the forward and reverse rates are equal, so $\text{Rate}_{\text{forward}}/\text{Rate}_{\text{reverse}} = 1$. By rearranging these equations, we can express the equilibrium constant as the ratio of the rate constants:

$$K_{\text{eq}} = \frac{k_{\text{f}}}{k_{\text{r}}} \tag{2-3}$$

The concept of K_{eq} is particularly helpful when we want to think about the energy that is released or absorbed when a chemical reaction occurs. We will discuss this concept in considerable detail in Section 2.4.

Chemical Reactions in Cells Are at Steady State

Under appropriate conditions and given sufficient time, a single biochemical reaction carried out in a test tube eventually reaches

equilibrium, at which the concentrations of reactants and products do not change with time because the rates of the forward and reverse reactions are equal. Within cells, however, many reactions are linked in pathways in which the fate of the product of one reaction is not simply to remain unchanged or to be reconverted via a reverse reaction to the reactants. For example, the product of one reaction might serve as a reactant in another reaction, or it might be pumped out of the cell. In this more complex situation, the original reaction can never reach equilibrium because some of the product molecules do not have a chance to be converted back to reactants. Nevertheless, in such non-equilibrium conditions, the rate of formation of a substance can be equal to the rate of its consumption, and as a consequence, the concentration of that substance remains constant over time. In such circumstances, the system of linked reactions for producing and consuming that substance is said to be in a **steady state** (Figure 2-23). One consequence of such linked reactions is that they prevent the accumulation of excess intermediates, protecting cells from the harmful effects of intermediates that are toxic at high concentrations. When the concentration of a product of an ongoing reaction is not changing over time, it might be a consequence of a state of equilibrium, or it might be a consequence of a steady state. In biological systems, when metabolite concentrations, such as blood glucose levels, are not changing with time — a condition called *homeostasis* (from the Greek *homoios* meaning “similar” and

stasis meaning “standing still or stability”) — it is a consequence of a steady state rather than equilibrium.

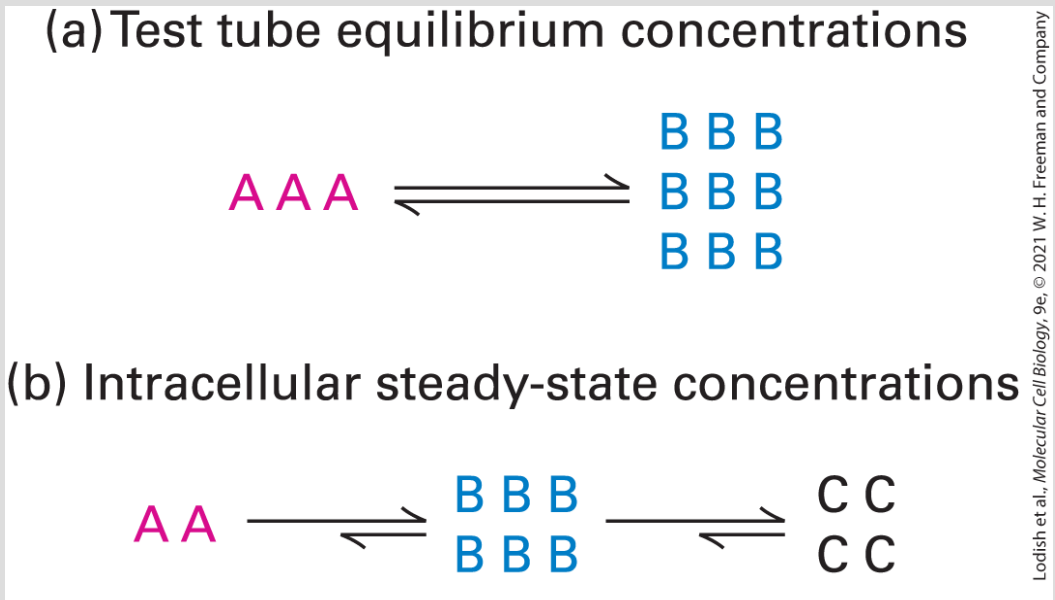


FIGURE 2-23 Comparison of reactions at equilibrium and at steady state. (a) In the test tube, a biochemical reaction ($A \rightarrow B$) eventually reaches equilibrium, at which the rates of the forward and reverse reactions are equal (as indicated by the reaction arrows of equal length), and the relative concentrations of the reactants (A) and products (B) do not change (in this example, the B:A ratio at equilibrium is 9:3). (b) In metabolic pathways within cells, the product B is commonly consumed — in this example, by conversion to C. A pathway of linked reactions is at steady state when the rate of formation of the intermediates (e.g., B) equals its rate of consumption. As indicated by the unequal length of the arrows, the individual reversible reactions constituting a metabolic pathway do not reach equilibrium. Moreover, the concentrations of the intermediates at steady state can differ from what they would be at equilibrium.

Dissociation Constants of Binding Reactions Reflect the Affinity of Interacting Molecules

The concept of equilibrium also applies to the binding of one molecule to another even when there are no covalent changes to either molecule. Many important cellular processes depend on such binding reactions, which involve the making and breaking of various noncovalent interactions rather than covalent bonds, as discussed above. In some cases, the noncovalent binding can be relatively weak, or loose; in other cases the binding can be very strong, or tight. A common example is the binding of a **ligand** (e.g., the hormone insulin or adrenaline) to its **receptor** on the surface of a cell, which triggers an intracellular signaling pathway (see Chapter 15). Another example is the binding of a protein to a specific sequence of bases in a molecule of DNA, which frequently causes the expression of a nearby gene to increase or decrease (see Chapter 8). Here we will review the most common approach used to think about how tightly or loosely two molecules bind to one another.

To illustrate the general approach, let's calculate the extent to which a protein (P) binds to a particular, short sequence of As, Ts, Gs, and Cs present in a relatively long DNA molecule (D) at equilibrium, forming a protein-DNA complex (PD):

$$\mathrm{P} + \mathrm{D} \rightleftharpoons \mathrm{PD}$$

Most commonly, binding reactions are described in terms of the **dissociation constant** ($K_\mathbf{d}$), which is the reciprocal of the equilibrium constant ($K_\mathrm{d} = 1/K_\mathrm{eq}$). For this binding reaction, the dissociation constant is calculated from the concentrations of the three components when they are at equilibrium by

$$K_\mathrm{d} = \frac{[\mathrm{P}][\mathrm{D}]}{[\mathrm{PD}]} \tag{2-4}$$

It is worth noting that in such a binding reaction, when half of the DNA is bound to the protein (PD) and the other half is unbound (D), then the concentration of PD is equal to that of D, $[\mathrm{PD}] = [\mathrm{D}]$, and the concentration of P ([P]) is equal to K_d. The lower the K_d, the tighter that P binds to D — often referred to as a higher binding affinity of P for D — and thus the lower the concentration of P needed to bind to half of D. This can sometimes lead to a bit of confusion. Remember, the *lower* the K_d, the *higher* the binding affinity, and the tighter the binding.

Typically, a protein's binding to a specific DNA sequence exhibits a K_d of 10^{-10} M, where M symbolizes molarity, or moles per liter ($\mathrm{mol/L}$). To relate the magnitude of this dissociation constant to the intracellular ratio of bound to unbound DNA, let's consider the simple example of a bacterial cell having a volume of 1.5×10^{-15} L and containing 1 molecule of DNA and 10 molecules of the DNA-

binding protein P. In this case, given a K_d of 10^{-10} M and the total concentration of the P in the cell (~111×10^{-10} M, about a hundredfold higher than the K_d), 99 percent of the time the specific DNA sequence to which P binds will have a molecule of protein P bound to it and 1 percent of the time it will not, even though the cell contains only 10 molecules of the protein! Clearly P and D have a high affinity for each other and bind tightly, as reflected by the low value of the dissociation constant for their binding. For protein-protein and protein-DNA binding, K_d values of ~10^{-9} M (nanomolar) or less are considered to be tight, ~10^{-6} M (micromolar) modestly tight, and ~10^{-3} M (millimolar) relatively weak.

A large biological macromolecule, such as a protein, can have multiple binding surfaces for binding several molecules simultaneously (Figure 2-24). In some cases, these binding reactions are independent, each with its own distinct K_d value that is independent of the others. In other cases, binding of a molecule at one site on a macromolecule can change the three-dimensional shape, or conformation, of a distant site (a phenomenon called **allostery**, see Chapter 3). That change in conformation alters the site's binding interactions with some other molecule (altering the K_d). The covalent modifications of amino acid side chains — mentioned previously — often influence the three-dimensional shapes of binding sites required for such binding interactions. These covalent and noncovalent binding reactions are important

mechanisms by which one molecule can alter, and thus regulate, the structure and binding activity of another. We examine this regulatory mechanism in more detail in Chapter 3.

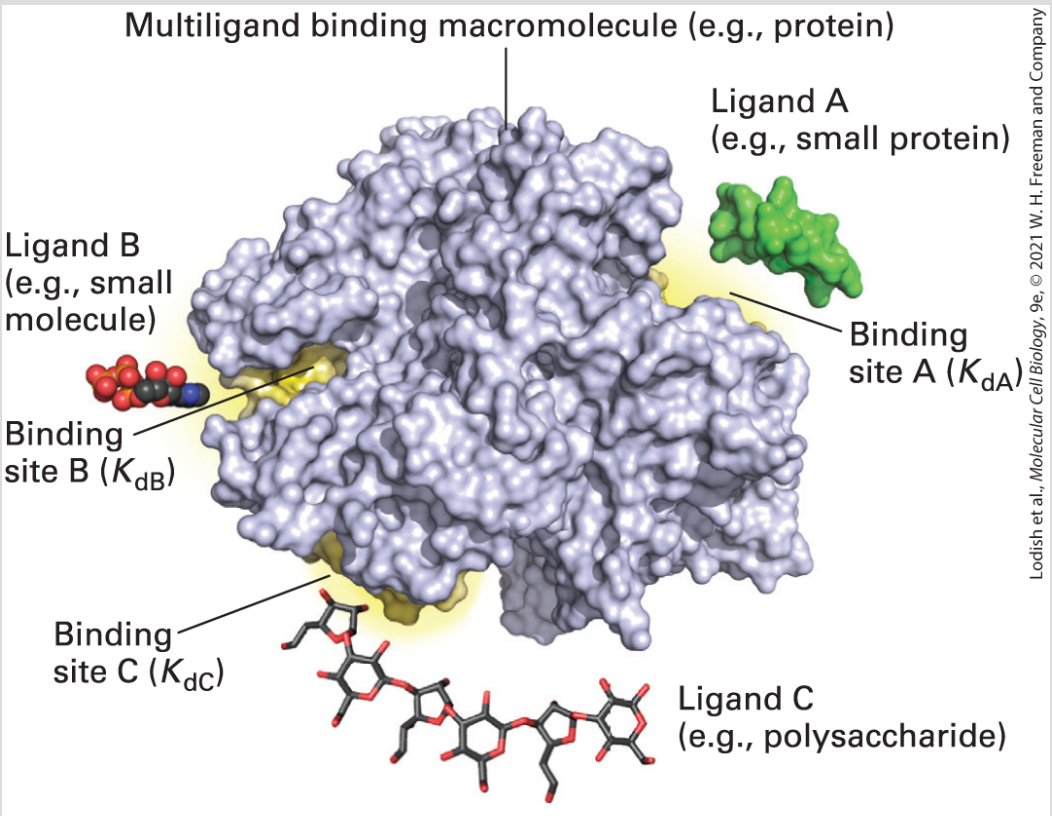


FIGURE 2-24 Macromolecules can have distinct binding sites for multiple ligands. A large macromolecule (e.g., a protein, purple) with three distinct binding sites (yellow, A–C) is shown; each of the three binding sites exhibit molecular complementarity to three different binding partners (ligands A–C) with distinct dissociation constants ($K_{dA–C}$).

Biological Fluids Have Characteristic pH Values

The solvent inside cells and in all extracellular fluids is water. An important characteristic of any aqueous solution is the concentration of positively charged hydrogen ions (H^+) and negatively charged hydroxyl ions (OH^-). Because these ions are the dissociation products of H_2O, they are constituents of all living systems, and they are liberated by many reactions that take place between molecules within cells. These ions can also be transported into or out of cells, as when highly acidic gastric juice is secreted by cells lining the walls of the stomach.

When a water molecule dissociates, one of its polar H—O bonds breaks. The resulting hydrogen ion, referred to as a *proton*, has a short lifetime as a free ion and quickly combines with a water molecule to form a hydronium ion (H_3O^+). For convenience, we refer to the concentration of hydrogen ions in a solution, $[H^+]$, even though this quantity really represents the concentration of hydronium ions, $[H_3O^+]$. Dissociation of H_2O generates one OH^- ion along with each H^+. The dissociation of water is a reversible reaction:

$$H_2O \rightleftharpoons H^+ + OH^-$$

At 25 °C, $[H^+][OH^-] = 10^{-14}\ M^2$, so that in pure water, $[H^+] = [OH^-] = 10^{-7}$ M.

The concentration of hydrogen ions in a solution is expressed conventionally as its **pH**, defined as the negative log of the hydrogen ion concentration. The pH of pure water at 25 $^\circ$C is 7:

$$\mathrm{pH} = -\log[\mathrm{H}^+] = \log \frac{1}{[\mathrm{H}^+]} = \log \frac{1}{10^{-7}} = 7$$

It is important to keep in mind that a one-unit difference in pH represents a tenfold difference in the concentration of protons. On the pH scale, 7.0 is considered neutral: pH values below 7.0 indicate acidic solutions (higher $[\mathrm{H}^+]$), and values above 7.0 indicate basic, or alkaline, solutions (Figure 2-25). For instance, gastric juice, which is rich in hydrochloric acid (HCl), has a pH of about 1. Its $[\mathrm{H}^+]$ is roughly 1-million-fold greater than that of cytoplasm, which has a pH of about 7.2–7.4.

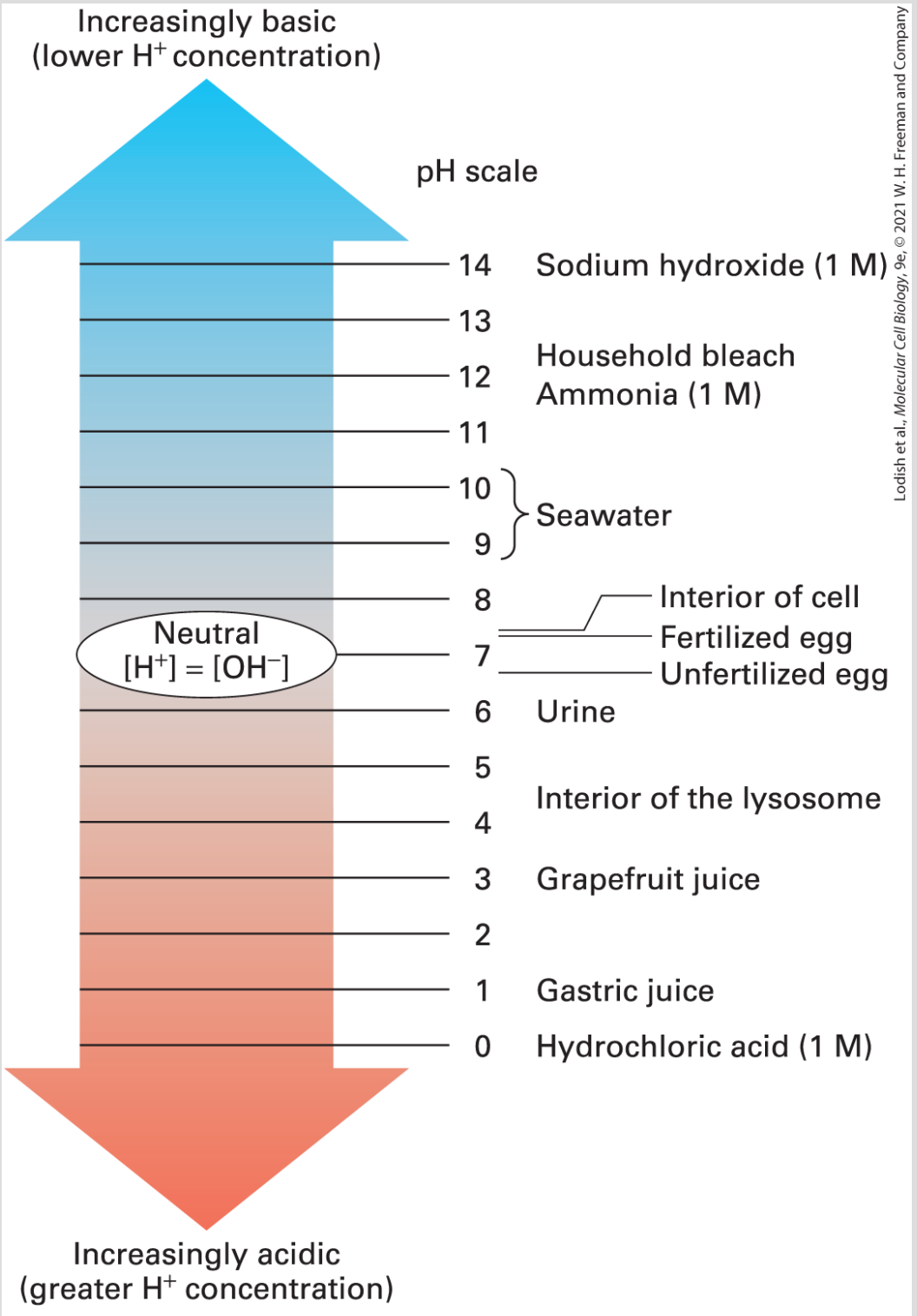


FIGURE 2-25 Some pH values for common solutions. The pH of an aqueous solution is the negative log of the hydrogen ion concentration. The pH values for most intracellular and extracellular biological fluids are near 7 and are

carefully regulated to permit the proper functioning of cells, organelles, and cellular secretions. The pH values for solutions of ammonia and hydrochloric acid are for one molar (1 M) solutions.

Although the cytosol of cells normally has a pH of about 7.2, the interior of certain organelles in eukaryotic cells (see Chapter 1) can have a much lower pH. The internal (luminal) fluid in lysosomes, for example, has a pH of about 4.5. The many degradative enzymes within lysosomes function optimally in an acidic environment, whereas their action is inhibited in the near neutral pH environment of the cytoplasm. As this example illustrates, maintenance of a particular pH is essential for the proper functioning of some cellular structures. On the other hand, dramatic shifts in cellular pH may play an important role in controlling cellular activity. For example, the pH of the cytoplasm of an unfertilized egg of the sea urchin, an aquatic animal, is 6.6. Within 1 minute of fertilization, however, the pH rises to 7.2; that is, the $[H^+]$ decreases to about one-fourth its original value, a change that is necessary for subsequent growth and division of the egg.

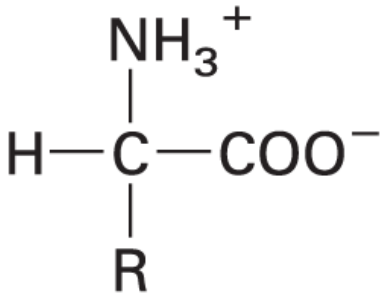
Hydrogen Ions Are Released by Acids and Taken Up by Bases

In general, an **acid** is any molecule, ion, or chemical group that tends to release a hydrogen ion (H^+). Examples are the carboxyl

group (—COOH), which tends to dissociate to form the negatively charged carboxylate ion ($—COO^-$) and a hydrogen ion, and hydrochloric acid (HCl). Conversely, a **base** is any molecule, ion, or chemical group that readily combines with an H^+. Examples are the hydroxyl ion (OH^-); ammonia (NH_3), which forms an ammonium ion ($NH_4{}^+$); and the amino group ($—NH_2$).

When an acid is added to an aqueous solution, the $[H^+]$ increases, and the pH goes down. Conversely, when a base is added to an aqueous solution, the $[H^+]$ decreases, and the pH goes up. Because $[H^+][OH^-] = 10^{-14}\ M^2$, any increase in $[H^+]$ is coupled with a commensurate decrease in $[OH^-]$, and vice versa.

Many biological molecules contain both acidic and basic groups. For example, in neutral solutions ($pH = 7.0$), many amino acids exist predominantly in the doubly ionized form, in which the carboxyl group has lost a proton and the amino group has accepted one:



where R represents the uncharged side chain. Such a molecule, containing an equal number of positive and negative ions, is called a

zwitterion. Zwitterions, having no net charge, are neutral. At extreme pH values, only one of these two ionizable groups of an amino acid is charged: the $—NH_2{}^+$ at low pH or the $—COO^-$ at high pH.

The dissociation reaction for an acid (or acid group in a larger molecule), HA, can be written as $HA \rightleftharpoons H^+ + A^-$. The equilibrium constant for this reaction, denoted K_a (the subscript *a* stands for "acid"), is defined as $K_a = [H^+][A^-]/[HA]$. Taking the logarithm of both sides and rearranging the result yields a very useful relation between the equilibrium constant K_a and pH:

$$pH = pK_a + \log \frac{[A^-]}{[HA]} \tag{2-5}$$

where pK_a equals $-\log\ K_a$.

From this expression, commonly known as the *Henderson–Hasselbalch equation*, it can be seen that the pK_a of any acid is equal to the pH at which half the molecules are dissociated and half are neutral (undissociated). This is because when $[A^-] = [HA]$, then $\log\ ([A^-]/[HA]) = 0$, and thus $pK_a\ = PH$. The Henderson–Hasselbalch equation allows us to calculate the degree of dissociation of an acid — that is, the ratio of dissociated and undissociated forms — if both the pH of the solution and the pK_a of

the acid are known. Experimentally, by measuring the $[A^-]$ and $[HA]$ as a function of the solution's pH, one can calculate the pK_a of the acid and thus the equilibrium constant K_a for the dissociation reaction (Figure 2-26). Knowing the pK_a of a molecule not only provides an important description of its properties, but also allows us to exploit these properties to manipulate the acidity of an aqueous solution and to understand how biological systems control this critical characteristic of their aqueous fluids.

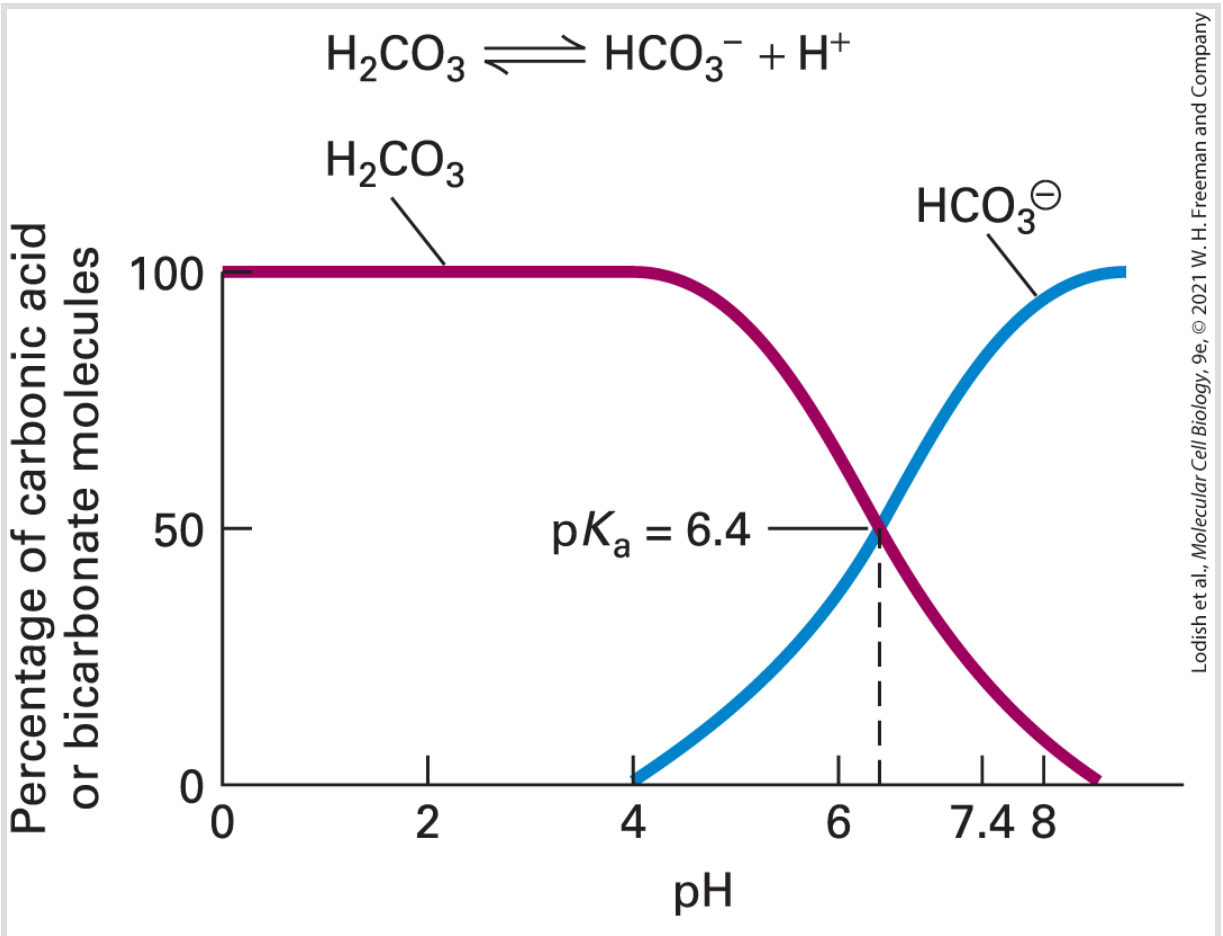


FIGURE 2-26 The relationship between pH, pK_a, and the dissociation of an acid. As the pH of a solution of carbonic acid rises from 0 to 8.5, the percentage of the compound in the undissociated, or un-ionized, form (H_2CO_3) decreases from 100 percent and that of the

ionized form increases from 0 percent. When the pH (6.4) is equal to the acid's $\mathrm{p}K_a$, half of the carbonic acid has ionized. When the pH rises to above 8, virtually all of the acid has ionized to the bicarbonate form $(\mathrm{HCO_3}^-)$.

Buffers Maintain the pH of Intracellular and Extracellular Fluids

A living, actively metabolizing cell must maintain a constant pH in the cytoplasm of about 7.2–7.4, and it must do so even as its metabolism is producing many acids. Cells have a reservoir of weak bases and weak acids, called **buffers**, which ensure that the cell's cytoplasmic pH remains relatively constant despite small fluctuations in the amounts of H^+ or OH^- being generated by metabolism or by the uptake or secretion of molecules and ions by the cell. Buffers maintain a constant pH by soaking up excess H^+ or OH^- when these ions are added to the cell or are produced by metabolism. As we shall see, buffers are most effective at preventing changes in pH when the pH of the solution is similar to the $\mathrm{p}K_\mathrm{a}$ of the buffer.

Consider a buffered solution whose pH is equal to the $\mathrm{p}K_\mathrm{a}$ of the buffer $([\mathrm{HA}] = [\mathrm{A}^-])$. If additional acid (or base) is added to this solution, the pH of the solution changes, but it changes less than it would if the buffer had not been present. This is because protons released by the added acid are taken up by the ionized form of the

buffer (A^-); likewise, hydroxyl ions released by the added base are neutralized by protons released by the undissociated buffer (HA). The capacity of a buffer or any other substance to release hydrogen ions or take them up depends partly on the extent to which the substance has already taken up or released protons, which in turn depends on the pH of the solution relative to the pK_a of the substance. The ability of a buffer to minimize changes in pH, its *buffering capacity*, depends on the concentration of the buffer and the relationship between its pK_a value and the pH, which is expressed by the Henderson–Hasselbalch equation.

The titration curve for acetic acid shown in Figure 2-27 illustrates the effect of pH on the fraction of molecules in the un-ionized (HA) and ionized (A^-) forms. When the pH is equal to the pK_a, half of the acetic acid is dissociated (dashed lines). At one pH unit below the pK_a of an acid, 91 percent of the molecules are in the HA form; at one pH unit above the pK_a, 91 percent are in the A^- form. At pH values more than one unit above or below the pK_a (unshaded regions in Figure 2-27), the buffering capacity of weak acids and bases declines rapidly. In other words, the addition of the same number of moles of base — for example, hydroxyl ions added as sodium hydroxide (NaOH) — to a solution containing a mixture of HA and A^- that is at a pH near the pK_a will cause less of a pH change than it would if the HA and A^- were not present or if the pH were far from the pK_a value.

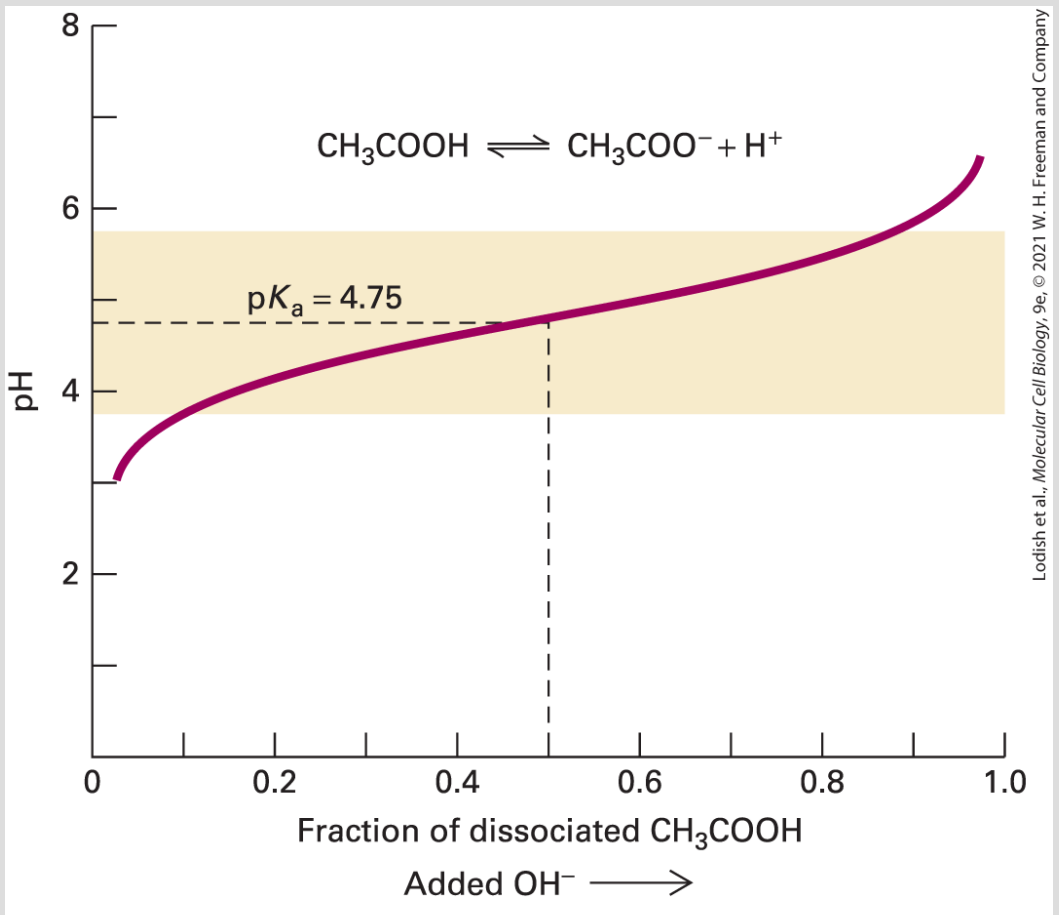


FIGURE 2-27 The titration curve of the buffer acetic acid (CH_3COOH). The pK_a for the dissociation of acetic acid to hydrogen and acetate ions is 4.75. At this pH, half the acetic acid molecules are dissociated. Because pH is measured on a logarithmic scale, the solution changes from 91 percent CH_3COOH at pH 3.75 to 9 percent CH_3COOH at pH 5.75. The acid has maximum buffering capacity in this pH range.

All biological systems contain one or more buffers. Phosphate ions, the ionized forms of phosphoric acid, are present in considerable quantities in cells and are an important buffer in maintaining, or buffering, the pH of the cytoplasm. Phosphoric acid (H_3PO_4) has three protons that are capable of dissociating, but they do not dissociate simultaneously. Loss of each proton can be described by a

discrete dissociation reaction and pK_a, as shown in Figure 2-28. When hydroxyl ions are added to a solution of phosphoric acid, the pH change is much less steep at pH values near the three pK_a values (shaded regions) than when the pH of the solution is not similar to any of the pK_as. The titration curve for phosphoric acid shows that the pK_a for the dissociation of the second proton is 7.2. Thus at pH 7.2, about 50 percent of cellular phosphate is $H_2PO_4^-$ and about 50 percent is HPO_4^{2-} according to the Henderson–Hasselbalch equation. For this reason, phosphate is an excellent buffer at pH values around 7.2, the approximate pH of the cytoplasm of cells, and at pH 7.4, the pH of human blood. The amino (lysine), guanidinium (arginine), and carboxylate (aspartate, glutamate) portions of amino acid side chains of proteins as well as the amino and carboxylate groups at the N- and C-termini of proteins can also bind and release protons. Thus proteins that are present in high concentrations inside cells and in many extracellular fluids can themselves serve as buffers.

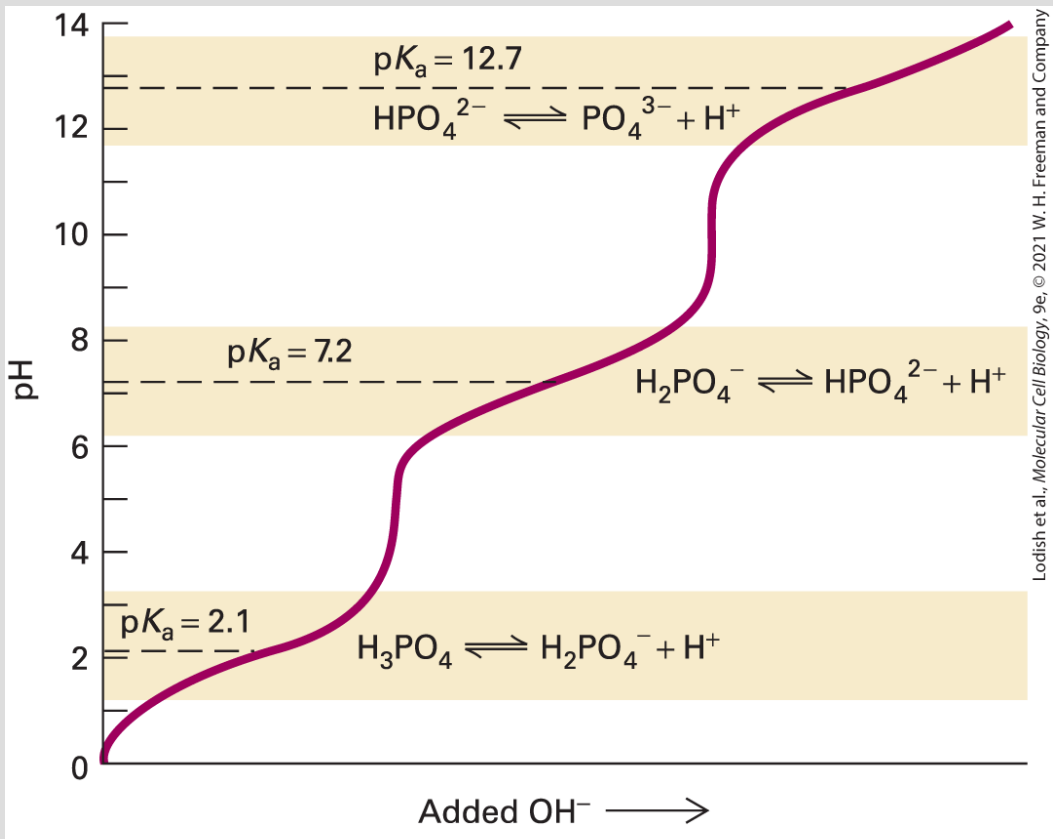


FIGURE 2-28 The titration curve of phosphoric acid (H_3PO_4), a common buffer in biological systems. This biologically ubiquitous molecule has three hydrogen atoms that dissociate at different pH values; thus phosphoric acid has three pK_a values, as noted on the graph. The shaded areas denote the pH ranges — within one pH unit of the three pK_a values — where the buffering capacity of phosphoric acid is high. In these regions, the addition of an acid (or base) will cause relatively small changes in the pH.

KEY CONCEPTS OF SECTION 2.3

Chemical Reactions and Chemical Equilibrium

- A chemical reaction is at equilibrium when the rate of the forward reaction is equal to the rate of the reverse reaction, and thus there is no net change over time in the concentration of the reactants or products.

- The equilibrium constant K_{eq} of a reaction reflects the ratio of products to reactants at equilibrium and thus is a measure of the extent of the reaction and the relative stabilities of the reactants and products.
- The K_{eq} depends on the temperature, pressure, and chemical properties of the reactants and products but is independent of the reaction rate and of the initial concentrations of reactants and products.
- For any reaction, the equilibrium constant K_{eq} equals the ratio of the forward rate constant to the reverse rate constant (k_{f}/k_r). The rates of conversion of reactants to products and vice versa depend on the rate constants and the concentrations of the reactants or products.
- Within cells, the linked reactions in metabolic pathways generally are not at equilibrium, but rather at steady state, at which the rate of formation of the intermediates equals their rate of consumption (see Figure 2-23), and thus the concentrations of the intermediates are not changing.
- The dissociation constant K_{d} for the noncovalent binding of two molecules is a measure of the tightness of binding or the stability of the complex formed between the molecules (e.g., ligand-receptor or protein-DNA complexes). The *lower* the K_d, the *higher* the binding affinity, and the tighter the binding. K_{d} values of ~$10^{-9}\ \mathrm{M}$ (nanomolar) are considered to be tight, ~$10^{-6}\ \mathrm{M}$ (micromolar) modestly tight, and ~$10^{-3}\ \mathrm{M}$ (millimolar) relatively weak.
- The pH is the negative logarithm of the concentration of hydrogen ions $(-\log\ [\mathrm{H}^+])$. The pH of the cytoplasm is normally about 7.2–7.4, whereas the interior of lysosomes has a pH of about 4.5.
- Acids release protons (H^+), and bases bind them.
- Buffers are mixtures of a weak acid (HA) and its corresponding base form (A^-), which minimize the change in pH of a solution when an acid or base is added. Biological systems use various buffers to maintain their pH within a very narrow range to permit the optimal functioning of biomolecules.

2.4 Biochemical Energetics

Obtaining, storing, and using energy are central to cellular life. Energy may be defined as the ability to do work, a concept that is as applicable to cells as to automobile engines and electric power plants. The energy stored within chemical bonds can be harnessed to support chemical work and the physical movements of cells. In this section, we review how energy influences the extents of chemical reactions (chemical thermodynamics) and the rates of chemical reactions (chemical kinetics).

Several Forms of Energy Are Important in Biological Systems

There are two principal forms of energy: kinetic and potential. **Kinetic energy** is the energy of movement — the motion of molecules, for example. **Potential energy** is stored energy — the energy stored in covalent bonds, for example. Potential energy plays a particularly important role in the energy economy of cells.

Thermal energy, or heat, is a form of kinetic energy — the energy of the motion of molecules. For heat to do work, it must flow from a region of higher temperature, where the average speed of molecular motion is greater, to one of lower temperature. Although differences in temperature can exist between the internal and external

environments of cells, these thermal gradients do not usually serve as the source of energy for cellular activities. The thermal energy in warm-blooded animals, which have evolved a mechanism for thermoregulation, is used chiefly to maintain constant body temperature. This is an important homeostatic function because the rates of many cellular activities are temperature dependent. For example, cooling mammalian cells from their normal body temperature of 37 °C to 4 °C can almost completely freeze or stop many cellular processes (e.g., intracellular membrane movements). Overheating can be lethal because heat disrupts the noncovalent interactions that are required to maintain the structures and functions of biological macromolecules and macromolecular complexes.

Radiant energy is the kinetic energy of photons, or waves of light. This form of kinetic energy is critical to biology. Radiant energy can be converted to thermal energy, for instance, when light is absorbed by molecules and the energy is converted to molecular motion. Radiant energy absorbed by molecules can also change the electronic structure of the molecules by moving electrons into higher energy orbitals, from whence the energy can later be recovered to perform work. For example, during photosynthesis, light energy absorbed by pigment molecules such as chlorophyll is subsequently converted into the energy of chemical bonds (see Chapter 12).

Mechanical energy, a major form of kinetic energy in biology, usually results from the conversion of stored chemical energy. For example, changes in the lengths of cytoskeletal filaments powered by chemical energy generate mechanical forces that push or pull on membranes and organelles (see Chapters 17 and 18).

Electric energy — the energy of moving electrons or other charged particles — is yet another major form of kinetic energy. Electric energy has particular importance to membrane function, as in electrically active neurons (see Chapters 11, 12, and 23).

Several forms of potential energy are biologically significant. Central to biology is **chemical potential energy**, the energy stored in the bonds connecting atoms in molecules. Indeed, most of the biochemical reactions described in this book involve the making or breaking of at least one covalent chemical bond. In general, energy must be expended to make covalent bonds in typical biomolecules, and energy is released when those bonds are broken. For example, the high potential energy in the covalent bonds of glucose can be released by controlled enzymatic combustion in cells (see Chapter 12). This energy is harnessed by the cell to do many kinds of work.

A second biologically important form of potential energy is the energy in a **concentration gradient**. When the concentration of a substance on one side of a barrier, such as a membrane, is different from that on the other side, a concentration gradient exists. All cells

form concentration gradients between their interior and the external fluids by selectively exchanging nutrients, waste products, and ions with their surroundings. Furthermore, the fluids within organelles in cells (e.g., mitochondria, lysosomes) frequently contain different concentrations of ions and other molecules than those in the cytosol; the concentration of protons within a lysosome (pH $\sim$4.5), as we saw in the last section, is about 500 times that in the cytosol (pH 7.2–7.4). Concentration gradients of protons across membranes are an important driver of energy production in mitochondria (see Chapter 12).

A third form of potential energy in cells is an **electric potential** — the energy of charge separation. For instance, there is a gradient of electric charge of about 200,000 volts per centimeter across the plasma membranes of virtually all cells. In Chapter 11, we discuss how concentration gradients and electric potential gradients are generated and maintained. In Chapter 12, we cover how they are converted to chemical potential energy.

Cells Can Transform One Type of Energy into Another

According to the first law of thermodynamics, energy is neither created nor destroyed, but can be converted from one form to another. (In nuclear reactions, mass is converted to energy, but this

is irrelevant in biological systems.) Energy conversions are very important in biology. In photosynthesis, for example, the radiant energy of light is transformed into the chemical potential energy of the covalent bonds between the atoms in a sucrose or starch molecule. In muscles and nerves, chemical potential energy stored in covalent bonds is transformed, respectively, into the kinetic energy of muscle contraction and the electric energy of neural transmission. In all cells, potential energy — released by breaking certain chemical bonds — is used to generate potential energy in the form of concentration and electric potential gradients. Similarly, energy stored in chemical concentration gradients or electric potential gradients is used to synthesize chemical bonds or to transport molecules from one side of a membrane to another to generate a concentration gradient. The latter process occurs during the transport of nutrients such as glucose into certain cells and the transport of many waste products out of cells.

Because all forms of energy are interconvertible, they can be expressed in the same units of measurement. Although the standard unit of energy is the joule, biochemists have traditionally used an alternative unit, the *calorie* ($1 \text{ joule} = 0.239 \text{ calorie}$). A calorie is the amount of energy required to raise the temperature of one gram of water by $1\ ^\circ\text{C}$. Throughout this book, we use the *kilocalorie* to measure energy changes ($1 \text{ kcal} = 1000 \text{ cal}$). When you read or hear about the "Calories" in food (note the capital *C*), the reference is almost always to kilocalories as defined here.

The Change in Free Energy Determines If a Chemical Reaction Will Occur Spontaneously

Chemical reactions can be divided into two types, depending on whether energy is absorbed or released in the process. In an **exergonic** ("energy-releasing") reaction, the products contain less energy than the reactants. Exergonic reactions take place spontaneously. The liberated energy is usually released as heat (the energy of molecular motion) and generally results in a rise in temperature, as in the oxidation (burning) of wood. In an **endergonic** ("energy-absorbing") reaction, the products contain more energy than the reactants, and energy is absorbed from the surroundings during the reaction. If there is no external source of energy to drive an endergonic reaction, it cannot take place. Endergonic reactions are responsible for the ability of the instant cold packs often used to treat injuries to rapidly cool below room temperature. Crushing the pack mixes the reactants, initiating the reaction.

A fundamentally important concept in understanding if a reaction is exergonic or endergonic, and therefore if it occurs spontaneously or not, is free energy (G), or *Gibbs free energy*, named after J. W. Gibbs. Gibbs, who received the first PhD in engineering in America in 1863, showed that "all systems change in such a way that free energy [*G*] is minimized." In other words, a chemical reaction occurs

spontaneously when the free energy of the products is lower than the free energy of the reactants. In the case of a chemical reaction, $\text{reactants} \rightleftharpoons \text{products}$, the **free-energy change, ΔG**, is given by

$$\Delta G = G_{\text{products}} - G_{\text{reactants}}$$

The relation of ΔG to the direction of any chemical reaction can be summarized in three statements:

- If ΔG is negative, the forward reaction (generation of products) will tend to occur spontaneously, and energy usually will be released as the reaction takes place (exergonic reaction) (Figure 2-29a). A reaction with a negative ΔG is referred to as thermodynamically favorable.
- If ΔG is positive, the forward reaction will not occur spontaneously; energy will have to be added to the system in order to force the reactants to become products (endergonic reaction, Figure 2-29b).
- If ΔG is zero, both forward and reverse reactions will occur without a gain or loss of energy, and there will be no spontaneous net conversion of reactants to products, or vice versa; the system is at equilibrium.

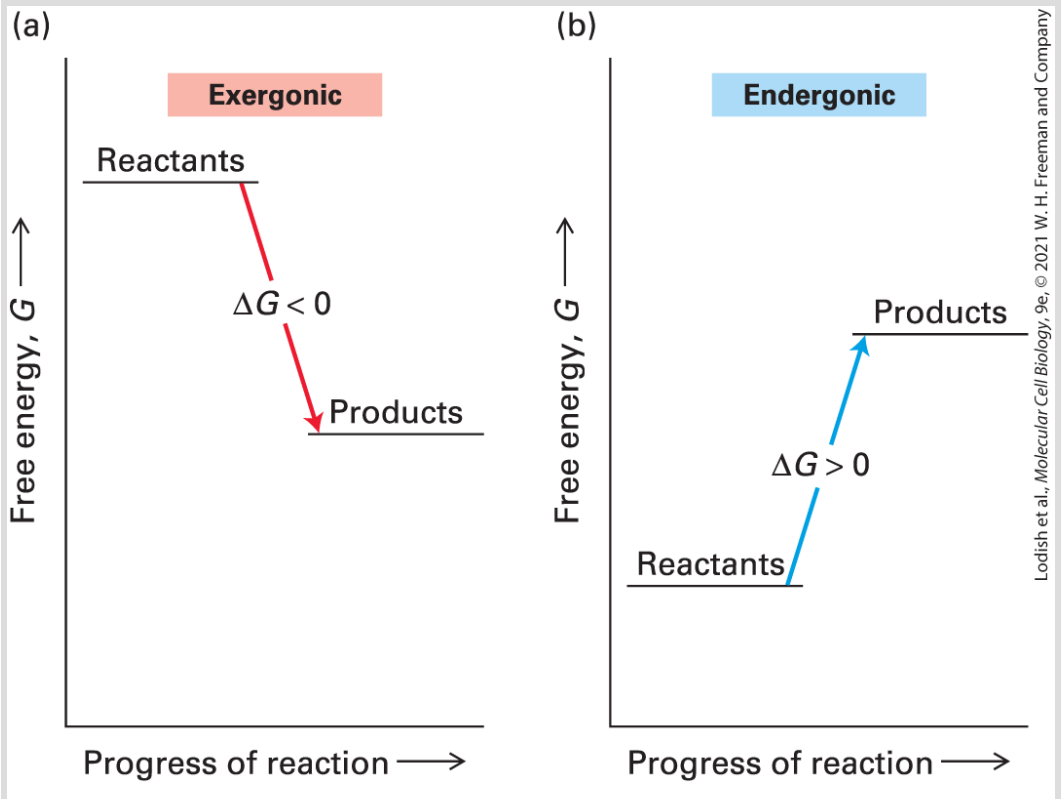


FIGURE 2-29 Changes in the free energy (ΔG) of exergonic and endergonic reactions. (a) In exergonic reactions, the free energy of the products is less than that of the reactants. Consequently, these reactions occur spontaneously, and energy is released as the reactions proceed. (b) In endergonic reactions, the free energy of the products is greater than that of the reactants, and these reactions do not occur spontaneously. An external source of energy must be supplied if the reactants are to be converted into products.

By convention, the *standard free-energy change* of a reaction ($\Delta G^{\circ\prime}$) is the value of the change in free energy at 298 K (25 °C), 1 atm pressure, pH 7.0 (as in pure water), and initial concentrations of 1 M for all reactants and products except protons, which are kept at 10^{-7}M (pH 7.0). The conditions of most biological reactions differ from these standard conditions, particularly in the concentrations

of reactants, which are normally less than 1 M, and in the temperature of mammals (37 °C).

The free energy of a chemical system can be defined as $G = H - TS$, where H is the bond energy, or **enthalpy (*H*)**, of the system; T is its temperature in degrees Kelvin (K); and S is the **entropy (*S*)**, a measure of its randomness or disorder. According to the second law of thermodynamics, the natural tendency of any isolated system is to become more disordered — that is, for entropy to increase. For example, a drop of ink placed in a large container of water will spontaneously spread out (become less ordered) rather than remain in a concentrated drop in the container. A reaction can occur spontaneously only if the combined effects of changes in enthalpy and entropy lead to a lower ΔG. That is, if temperature remains constant, a reaction proceeds spontaneously only if the free-energy change, ΔG, in the following equation is negative:

$$\Delta G = \Delta H - T\Delta S \tag{2-6}$$

In an **exothermic** ("heat-releasing") chemical reaction, ΔH is negative. In an **endothermic** ("heat-absorbing") reaction, ΔH is positive. The combined effects of the changes in the enthalpy and entropy determine if the ΔG for a reaction is positive or negative and thus if the reaction occurs spontaneously. An exothermic reaction ($\Delta H < 0$) in which entropy increases ($\Delta S > 0$) occurs

spontaneously ($\Delta G < 0$). An endothermic reaction ($\Delta H > 0$) will occur spontaneously if ΔS increases enough so that the $T\Delta S$ term can overcome the positive ΔH.

Many biological reactions lead to an increase in order and thus a decrease in entropy ($\Delta S < 0$). An obvious example is the reaction that links amino acids to form a protein. A solution of protein molecules has a lower entropy than does a solution of the same amino acids unlinked because the free movement of any amino acid is more restricted (greater order) when it is bound into a long chain than when it is not. Thus when cells synthesize polymers such as proteins from their constituent monomers, the polymerizing reaction will be spontaneous only if the cells can efficiently transfer energy to both generate the bonds that hold the monomers together and overcome the loss in entropy that accompanies polymerization. Often cells accomplish this feat by *coupling* such synthetic, entropy-lowering reactions with independent reactions that have a very highly negative ΔG, such as the hydrolysis of nucleoside triphosphates (discussed later in this section). In this way, cells can convert sources of energy in their environment into the highly organized structures and metabolic pathways that are essential for life.

The actual change in free energy during a reaction is influenced by temperature, pressure, and the initial concentrations of reactants and products, so it usually differs from the standard free-energy

change $\Delta G^{\circ\prime}$. Most biological reactions — like others that take place in aqueous solutions — are also affected by the pH of the solution. We can estimate free-energy changes for temperatures and initial concentrations that differ from the standard conditions by using the equation

$$\Delta G = \Delta G^{\circ\prime} + RT \ln Q = \Delta G^{\circ\prime} + RT \ln \frac{[\text{products}]}{[\text{reactants}]} \tag{2-7}$$

where R is the gas constant of $1.987\ \text{cal}/(\text{degree}\cdot\text{mol})$, T is the temperature (in degrees Kelvin), and Q is the *initial* ratio of products to reactants. For a reaction $\text{A} + \text{B} \rightleftharpoons \text{C}$, in which two molecules combine to form a third, Q in Equation 2-7 equals $[\text{C}]/[\text{A}][\text{B}]$. In this case, an increase in the initial concentration of either $[\text{A}]$ or $[\text{B}]$ will result in a larger negative value for ΔG and thus drive the reaction toward spontaneous formation of C.

Regardless of the $\Delta G^{\circ\prime}$ of a particular biochemical reaction, it will proceed spontaneously within cells only if ΔG is negative given the intracellular concentrations of reactants and products. For example, the conversion of glyceraldehyde 3-phosphate (G3P) to dihydroxyacetone phosphate (DHAP), two intermediates in the cellular breakdown of glucose (see Chapter 12),

$$\text{G3P} \rightleftharpoons \text{DHAP}$$

has a $\Delta G^{\circ\prime}$ of -1840 cal/mol. If the initial concentrations of G3P and DHAP are equal, then $\Delta G = \Delta G^{\circ\prime}$ because $RT\ \ln[\text{products}]/[\text{reactants}] = 0$; in this situation, the reversible reaction G3P $\rightleftharpoons$ DHAP will proceed spontaneously in the direction of DHAP formation until equilibrium is reached. However, if the initial [DHAP] is 0.1 M and the initial [G3P] is 0.001 M, with other conditions standard, then Q in Equation 2-7 equals $0.1/0.001 = 100$, giving a ΔG of $+887$ cal/mol. Under these conditions, the reaction will proceed in the direction of formation of G3P.

The ΔG of a reaction is independent of the reaction rate. Indeed, under normal physiological conditions, few, if any, of the biochemical reactions needed to sustain life would occur quickly enough without some mechanism for increasing reaction rates. As we describe below and in more detail in Chapter 3, the rates of reactions in biological systems are usually determined by the activity of **enzymes**, the protein catalysts that accelerate the formation of products from reactants without altering the value of ΔG.

The $\Delta G^{\circ\prime}$ of a Reaction Can Be Calculated from Its K_{eq}

A chemical mixture at equilibrium is in a stable state of minimal free energy. For a system at equilibrium ($\Delta G = 0, Q = K_{\text{eq}}$) under standard conditions, we can write

$$\Delta G^{\circ\prime} = -2.3RT \ \log K_{\text{eq}} = -1362 \ \log K_{\text{eq}} \tag{2-8}$$

(note the change to base 10 logarithms). Thus if we determine the concentrations of reactants and products at equilibrium (i.e., the K_{eq}), we can calculate the value of $\Delta G^{\circ\prime}$. For example, the K_{eq} for the interconversion of glyceraldehyde 3-phosphate to dihydroxyacetone phosphate (G3P $\rightleftharpoons$ DHAP) is 22.2 under standard conditions. Substituting this value into Equation 2-8, we can easily calculate the $\Delta G^{\circ\prime}$ for this reaction as -1840 cal/mol.

By rearranging Equation 2-8 and taking the antilogarithm, we obtain

$$K_{\text{eq}} = 10^{-(\Delta G^{\circ\prime}/2.3RT)} \tag{2-9}$$

From this expression, it is clear that if $\Delta G^{\circ\prime}$ is negative, the exponent will be positive, and hence K_{eq} will be greater than 1. Therefore, at equilibrium there will be more products than reactants; in other words, the formation of products from reactants is favored. Conversely, if $\Delta G^{\circ\prime}$ is positive, the exponent will be negative, and K_{eq} will be less than 1. The relationship between K_{eq}

and $\Delta G^{\circ\prime}$ further emphasizes the influence of the relative free energies of reactants and products on the extent to which a reaction will occur spontaneously.

The Rate of a Reaction Depends on the Activation Energy Necessary to Energize the Reactants into a Transition State

As a chemical reaction proceeds, reactants approach each other; some bonds begin to form while others begin to break. One way to think of the state of the molecules during this transition is that there are strains in the electronic configurations of the atoms and their bonds. The collection of atoms moves from the relatively stable state of the reactants to a transient, intermediate, and higher energy state during the course of the reaction (Figure 2-30). The state during a chemical reaction at which the system is at its highest energy level is called the **transition state**, and the collection of reactants in that state is called the **transition-state intermediate**. The energy needed to excite the reactants to this higher energy state is called the **activation energy** of the reaction. The activation energy is the difference in free energy between the reactants and the transition-state intermediate. The activation energy is usually represented by $\Delta G^{\ddagger}$, which is analogous to the representation of the change in Gibbs free energy (ΔG) already discussed. From the transition state,

the collection of atoms can either release energy as the reaction products are formed (reaction proceeds to the right in Figure 2-30) or release energy as the atoms go backward (leftward in Figure 2-30) and re-form the original reactants.

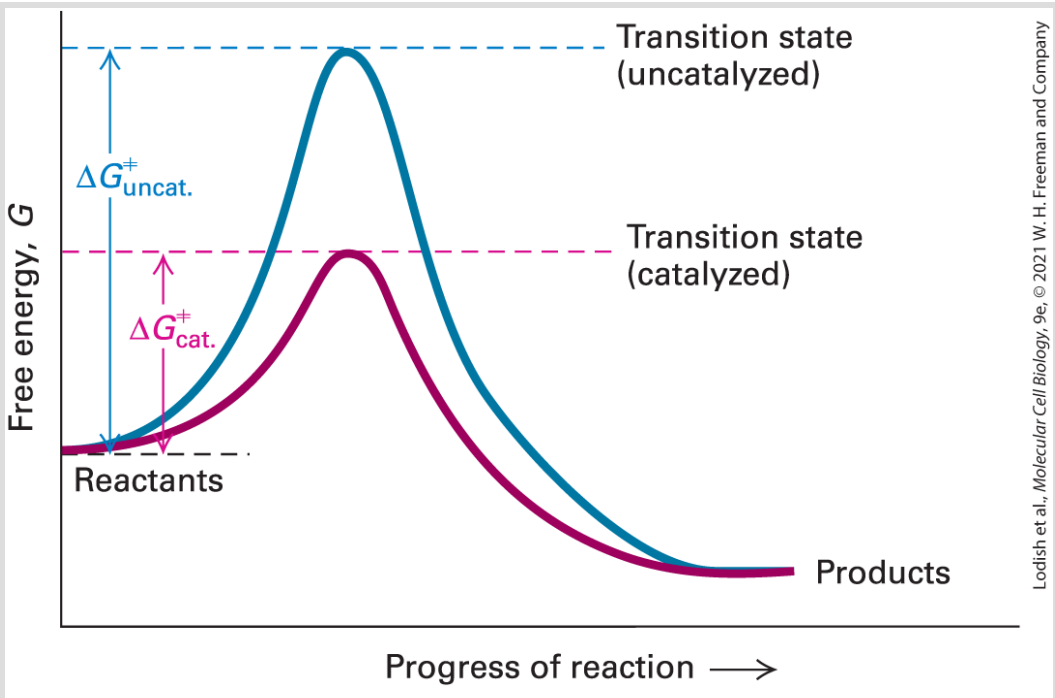


FIGURE 2-30 Activation energy of uncatalyzed and catalyzed chemical reactions. This hypothetical reaction pathway (blue) depicts the changes in free energy, *G*, as a reaction proceeds. A reaction will take place spontaneously if the free energy (*G*) of the products is less than that of the reactants ($\Delta G < 0$). However, all chemical reactions proceed through one (shown here) or more high-energy transition states, and the rate of a reaction is inversely proportional to the activation energy ($\Delta G^{\ddagger}$), which is the difference in free energy between the reactants and the transition state. In a catalyzed reaction (red), the free energies of the reactants and products are unchanged, but the free energy of the transition state is lowered, thus increasing the velocity of the reaction.

The velocity (V) at which products are generated from reactants during the reaction under a given set of conditions (temperature, pressure, reactant concentrations) will depend on the concentration of material in the transition state. That concentration will depend on (1) the rate at which reactants are converted into the transition state, which depends on the activation energy, and (2) the characteristic rate constant (v) at which the material in the transition state is converted to products. The higher the activation energy, the lower the fraction of reactants that reach the transition state, and the slower the overall rate of the reaction. The relationship between the concentration of reactants, v, and V is

$$V = v\,[\text{reactants}] \times 10^{-(\Delta G^{\ddagger}/2.3\text{RT})}$$

From this equation, we can see that lowering the activation energy $\Delta G^{\ddagger}$ — that is, decreasing the difference in free energy between the reactants and the transition state — leads to an acceleration of the overall reaction rate V. A reduction in $\Delta G^{\ddagger}$ of $1.36\ \text{kcal/mol}$ leads to a tenfold increase in the rate of the reaction, whereas a $2.72\ \text{kcal/mol}$ reduction increases the rate a hundredfold. Thus relatively small changes in $\Delta G^{\ddagger}$ can lead to large changes in the overall rate of the reaction.

Catalysts such as enzymes (discussed further in Chapter 3) accelerate reaction rates by lowering the relative energy of the transition state and thus the activation energy required to reach it

(see Figure 2-30). The relative energies of reactants and products determine if a reaction is thermodynamically favorable (negative ΔG), whereas the activation energy determines how rapidly products form — that is, the reaction kinetics. Thermodynamically favorable reactions will not occur at appreciable rates if the activation energies are too high.

Life Depends on the Coupling of Energetically Unfavorable Chemical Reactions with Energetically Favorable Ones

Many processes in cells are energetically unfavorable ($\Delta G > 0$) and will not proceed spontaneously. Examples include the synthesis of proteins from amino acids and the transport of a substance across the plasma membrane from a lower to a higher concentration. Cells can carry out an energy-requiring, or endergonic, reaction ($\Delta G_1 > 0$) by coupling it to an energy-releasing, or exergonic, reaction ($\Delta G_2 < 0$) if the sum of the two reactions has an overall net negative ΔG.

Suppose, for example, that the reaction $\mathrm{A} \rightleftharpoons \mathrm{B} + \mathrm{X}$ has a ΔG of $+5\ \mathrm{kcal/mol}$ and that the reaction $\mathrm{X} \rightleftharpoons \mathrm{Y} + \mathrm{Z}$ has a ΔG of $-10\ \mathrm{kcal/mol}$:

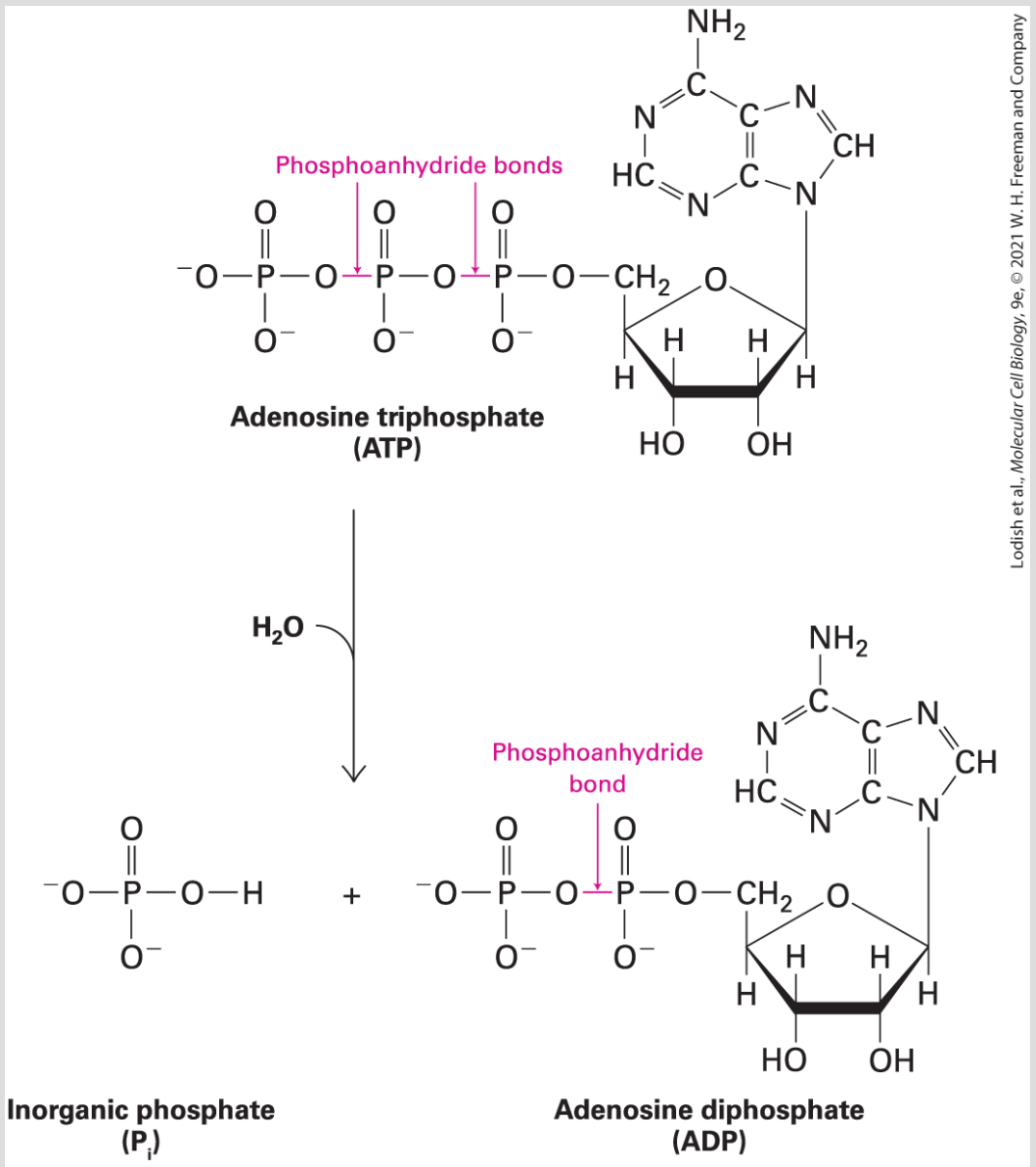
$$
\begin{array}{lcccccl}
(1) & \mathrm{A} & \rightleftharpoons & \mathrm{B} + \mathrm{X} & \Delta G & = +5\ \mathrm{kcal/mol} \\
(2) & \mathrm{X} & \rightleftharpoons & \mathrm{Y} + \mathrm{Z} & \Delta G & = -10\ \mathrm{kcal/mol} \\
\hline
Sum: & \mathrm{A} & \rightleftharpoons & \mathrm{B} + \mathrm{Y} + \mathrm{Z} & \Delta G^{\circ\prime} & = -5\ \mathrm{kcal/mol}
\end{array}
$$

In the absence of the second reaction, there would be much more A than B at equilibrium. However, because the conversion of X to $\mathrm{Y} + \mathrm{Z}$ is such a favorable reaction, it will pull the first process toward the formation of B and the consumption of A. Energetically unfavorable reactions in cells are often coupled to the energy-releasing hydrolysis of ATP.

Hydrolysis of ATP Releases Substantial Free Energy and Drives Many Cellular Processes

In almost all organisms, the nucleoside triphosphate **adenosine triphosphate**, or **ATP** (Figures 2-1d and 2-31), is the most important molecule for capturing, transiently storing, and subsequently transferring energy to perform work such as biosynthesis or mechanical motion. Commonly referred to as a cell's energy *currency*, ATP is a type of usable potential energy that cells can *spend* in order to power their activities. The storied history of ATP begins with its discovery in 1929, apparently simultaneously by Kurt Lohmann, who was working with the famous biochemist Otto

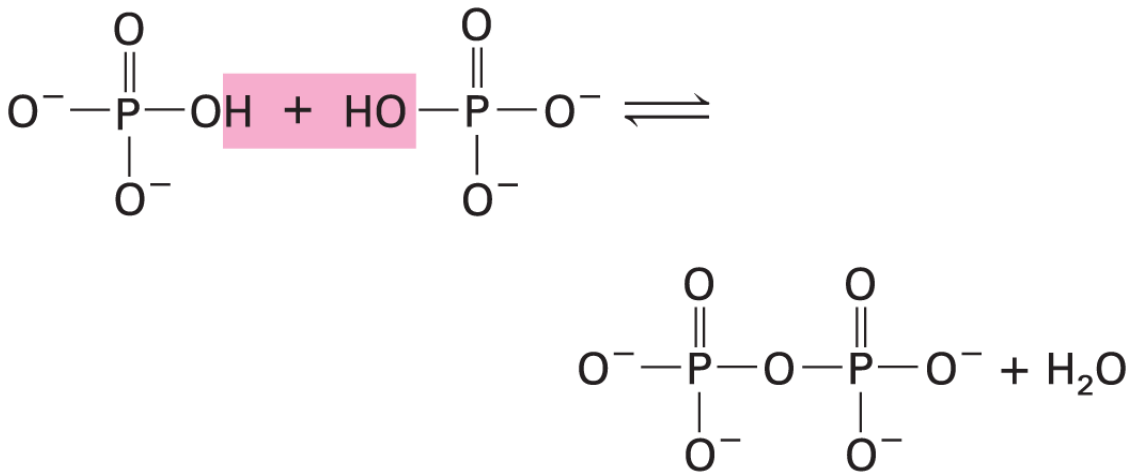
Meyerhof in Germany and who published first, and by the less well-known Cyrus Fiske and Yellapragada SubbaRow in the United States. [It is noteworthy that SubbaRow, who died in 1948 at the age of 53, also made key contributions to treatments for major diseases, including anemia (folic acid), cancer and rheumatoid arthritis (methotrexate), and bacterial infections (first tetracycline antibiotic).] Muscle contractions were shown to depend on ATP in the 1930s. The proposal that ATP is the main intermediary for the transfer of energy in cells is credited to Fritz Lipmann around 1941. Many Nobel Prizes have been awarded for the study of ATP and its role in cellular-energy metabolism, and its importance in understanding molecular cell biology cannot be overstated.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 2-31 Hydrolysis of adenosine triphosphate (ATP). The two phosphoanhydride bonds (red) in ATP (*top*), which link the three phosphate groups, each have a $\Delta G^{\circ\prime}$ of about $-7.3\ \mathrm{kcal/mol}$ for hydrolysis. Hydrolysis of the terminal phosphoanhydride bond by the addition of water results in the release of phosphate and generation of ADP. Hydrolysis of the phosphoanhydride bonds of ATP is exergonic (Figure 2-29a), especially the terminal one, and the hydrolysis of these bonds is the source of energy that drives many energy-requiring reactions in biological systems.

The useful energy in an ATP molecule is contained in **phosphoanhydride bonds**. These are covalent bonds formed from the condensation of two molecules of phosphate by the loss of water:



As shown in Figure 2-31, an ATP molecule has two key phosphoanhydride (also called *phosphodiester*) bonds. Forming these bonds (represented here by the symbol ~) in ATP requires an input of energy. When these bonds are hydrolyzed, or broken by the addition of water, that energy is released. Hydrolysis of a phosphoanhydride bond in each of the following reactions has a highly negative $\Delta G^{\circ\prime}$ of about -7.3 kcal/mol:

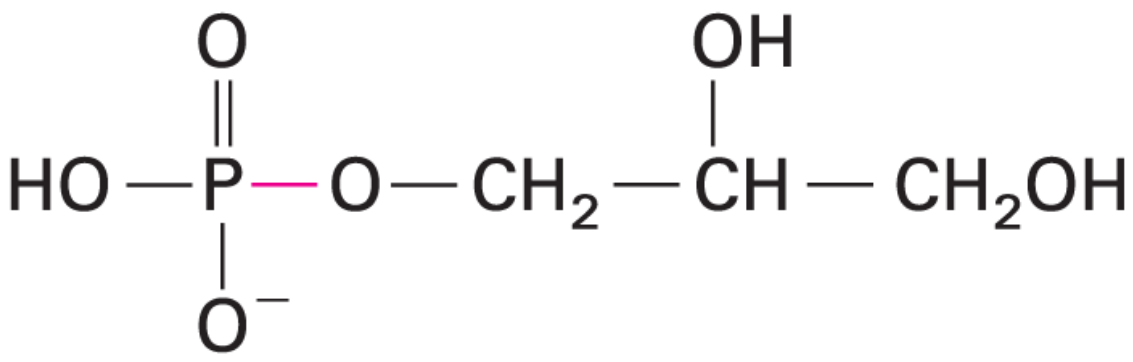
$$\underset{\text{(ATP)}}{\text{Ap}\sim\text{p}\sim\text{p}} + \text{H}_2\text{O} \rightarrow \underset{\text{(ADP)}}{\text{Ap}\sim\text{p}} + \text{P}_\text{i} + \text{H}^+$$

$$\underset{\text{(ATP)}}{\text{Ap}\sim\text{p}\sim\text{p}} + \text{H}_2\text{O} \rightarrow \underset{\text{(AMP)}}{\text{Ap}} + \text{PP}_\text{i} + \text{H}^+$$

$$\underset{\text{(ADP)}}{\text{Ap}\sim\text{p}} + \text{H}_2\text{O} \rightarrow \underset{\text{(AMP)}}{\text{Ap}} + \text{P}_\text{i} + \text{H}^+$$

P_i stands for inorganic phosphate (PO_4^{3-}) and PP_i for inorganic pyrophosphate, two phosphate groups linked by a phosphoanhydride bond. As the top two reactions show, the removal of a phosphate group from ATP leaves adenosine diphosphate (ADP), and the removal of a pyrophosphate group from ATP leaves adenosine monophosphate (AMP).

A phosphoanhydride bond or other *high-energy bond* (commonly denoted by ~) is not intrinsically different from other covalent bonds. High-energy bonds simply release substantial amounts of energy when hydrolyzed. For instance, the $\Delta G^{\circ\prime}$ for hydrolysis of a phosphoanhydride bond in ATP (–7.3 kcal/mol) is more than three times the $\Delta G^{\circ\prime}$ for hydrolysis of the phosphoester bond (red) in glycerol 3-phosphate (–2.2 kcal/mol):



Glycerol 3-phosphate

A principal reason for this difference is that ATP and its hydrolysis products, ADP and P_i, are charged at neutral pH. During synthesis of ATP, a large amount of energy must be used to force the negative charges in ADP and P_i together. Conversely, this energy is released when ATP is hydrolyzed to ADP and P_i. In comparison, formation of the phosphoester bond between an uncharged hydroxyl in glycerol and P_i requires less energy, and less energy is released when this bond is hydrolyzed.

Cells have evolved mechanisms by which enzymes transfer the free energy released by hydrolysis of phosphoanhydride bonds to other molecules, thereby driving reactions that would otherwise be energetically unfavorable. For example, if the ΔG for the reaction $B + C \rightarrow D$ is positive but less than the absolute value of the ΔG for hydrolysis of ATP, the reaction can be driven to the right by coupling it to hydrolysis of the terminal phosphoanhydride bond in ATP. In one common mechanism of such *energy coupling*, some of the

energy stored in this phosphoanhydride bond is transferred to one of the reactants (here, B) by the breaking of the bond in ATP and the formation of a covalent bond between the released phosphate group and that reactant. The phosphorylated intermediate generated in this way can then react with reactant C to form product $\mathrm{D} + \mathrm{P_i}$ in a reaction that has an overall negative ΔG:

$$\mathrm{B} + \mathrm{Ap\sim p\sim p} \rightarrow \mathrm{B\sim p} + \mathrm{Ap\sim p}$$
$$\mathrm{B\sim p} + \mathrm{C} \rightarrow \mathrm{D} + \mathrm{P_i}$$

The overall reaction

$$\mathrm{B} + \mathrm{C} + \mathrm{ATP} \rightleftharpoons \mathrm{D} + \mathrm{ADP} + \mathrm{P_i}$$

is energetically favorable ($\Delta G < 0$). Similarly, hydrolysis of GTP to GDP can provide energy to perform work, including the synthesis of ATP (see Chapter 12), but most often GTP hydrolysis is used to control cellular systems (e.g., protein synthesis, hormonal signaling) rather than as a source of energy (see Chapter 15).

An alternative mechanism of energy coupling is to use the energy released by ATP hydrolysis to change the conformation of a molecule to an *energy-rich* stressed state. In turn, the energy stored as conformational stress can be released as the molecule relaxes back into its unstressed conformation. If this relaxation process can

be coupled to another reaction, the released energy can be harnessed to drive cellular processes.

As with many biosynthetic reactions, transport of molecules into or out of the cell often has a positive ΔG and thus requires an input of energy to proceed. Such simple transport reactions do not *directly* involve the making or breaking of covalent bonds; thus their $\Delta G^{\circ\prime}$ is 0. In the case of a substance moving into a cell, Equation 2-7 becomes

$$\Delta G = RT \ln \frac{[\mathrm{C_{in}}]}{[\mathrm{C_{out}}]} \tag{2-10}$$

where $[\mathrm{C_{in}}]$ is the initial concentration of the substance inside the cell and $[\mathrm{C_{out}}]$ is its concentration outside the cell. We can see from Equation 2-10 that ΔG is positive for transport of a substance into a cell against its concentration gradient (when $[\mathrm{C_{in}}] > [\mathrm{C_{out}}]$); the energy to drive such uphill transport is often supplied by the hydrolysis of ATP. Conversely, when a substance moves down its concentration gradient ($[\mathrm{C_{out}}] > [\mathrm{C_{in}}]$), ΔG is negative. Such downhill transport releases energy that can be coupled to an energy-requiring reaction, such as the movement of another substance uphill across a membrane or the synthesis of ATP itself (see Chapters 11 and 12).

ATP Is Generated During Photosynthesis and Respiration

ATP is continuously being hydrolyzed to provide energy for many cellular activities. Some estimates suggest that humans daily hydrolyze a mass of ATP equal to their entire body weight. Clearly to continue functioning, cells must constantly replenish their ATP supply. Therefore, cells must obtain energy from their environment that can be used to resynthesize the ATP. For nearly all cells, the ultimate source of energy used to make ATP is sunlight.

Some organisms can use sunlight directly. Through the process of **photosynthesis**, plants, algae, and certain photosynthetic bacteria trap the energy of sunlight and use it to synthesize ATP from ADP and P_i. Much of the ATP produced in photosynthesis is hydrolyzed to provide energy for the conversion of carbon dioxide to six-carbon sugars, a process called *carbon fixation*:

$$6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2 + \text{energy}$$

The sugars made during photosynthesis are a source of food, and thus energy, for the photosynthetic organisms making them and for the non-photosynthetic organisms, such as animals, that consume the plants either directly or indirectly by eating other animals that have eaten the plants. In this way, sunlight is the direct or indirect source of energy for most organisms (see Chapter 12).

In plants, animals, and nearly all other organisms, the free energy in sugars and other molecules derived from food is released in the processes of **glycolysis** and **cellular respiration**. During cellular respiration, energy-rich molecules in food (e.g., glucose) are oxidized to carbon dioxide and water. The complete oxidation of glucose,

$$\mathrm{C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O}$$

has a $\Delta G^{\circ\prime}$ of -686 kcal/mol and is the reverse of photosynthetic carbon fixation. Cells employ an elaborate set of protein-mediated reactions to couple the oxidation of 1 molecule of glucose to the synthesis of as many as 30 molecules of ATP from 30 molecules of ADP. This oxygen-dependent (**aerobic**) degradation (**catabolism**) of glucose is the major pathway for generating ATP in all animal cells, all non-photosynthetic plant cells, and many bacterial cells. Catabolism of fatty acids can also be an important source of ATP. We discuss the mechanisms of photosynthesis and cellular respiration in Chapter 12.

Although light energy captured in photosynthesis is the primary source of chemical energy for cells, it is not the only source. Certain microorganisms that live in or around deep-sea hydrothermal vents, where adequate sunlight is unavailable, derive the energy for converting ADP and $\mathrm{P_i}$ into ATP from the oxidation of reduced

inorganic compounds. These reduced compounds originate deep in the earth and are released at the vents.

NAD^+ and FAD Couple Many Biological Oxidation and Reduction Reactions

In many chemical reactions, electrons are transferred from one atom or molecule to another; this transfer may or may not accompany the formation of new chemical bonds or the release of energy that can be coupled to other reactions. The loss of electrons from an atom or a molecule is called **oxidation**, and the gain of electrons by an atom or a molecule is called **reduction**. An example of oxidation is the removal of electrons from the sulfhydryl group–containing side chains of two cysteine amino acids to form a disulfide bond, described in Section 2.2. Electrons are neither created nor destroyed in a chemical reaction, so if one atom or molecule is oxidized, another must be reduced. For example, oxygen draws electrons from Fe^{2+} (ferrous) ions to form Fe^{3+} (ferric) ions, a reaction that occurs as part of the process by which carbohydrates are degraded in mitochondria. Each oxygen atom receives two electrons, one from each of two Fe^{2+} ions:

$$2\,Fe^{2+} + \frac{1}{2}\,O_2 \rightarrow 2\,Fe^{3+} + O^{2-}$$

Thus Fe^{2+} is oxidized and O_2 is reduced. Such reactions in which one molecule is reduced and another is oxidized are often referred to as **redox reactions**. Oxygen is an electron acceptor in many redox reactions in cells under aerobic conditions.

Many biologically important oxidation and reduction reactions involve the removal or addition of hydrogen atoms (protons plus electrons) rather than the transfer of isolated electrons on their own. The oxidation of succinate to fumarate, which occurs in mitochondria, is an example (Figure 2-32). Protons are soluble in aqueous solutions (as H_3O^+), but electrons are not, so they must be transferred directly from one atom or molecule to another without a water-dissolved intermediate. In this type of oxidation reaction, electrons are often transferred to small, electron-carrying molecules, sometimes referred to as *coenzymes*. The most common of these electron carriers are **NAD^+ (nicotinamide adenine dinucleotide**), which is reduced to NADH, and **FAD (flavin adenine dinucleotide)**, which is reduced to $FADH_2$ (Figure 2-33). The reduced forms of these coenzymes can transfer protons and electrons to other molecules, thereby reducing them.

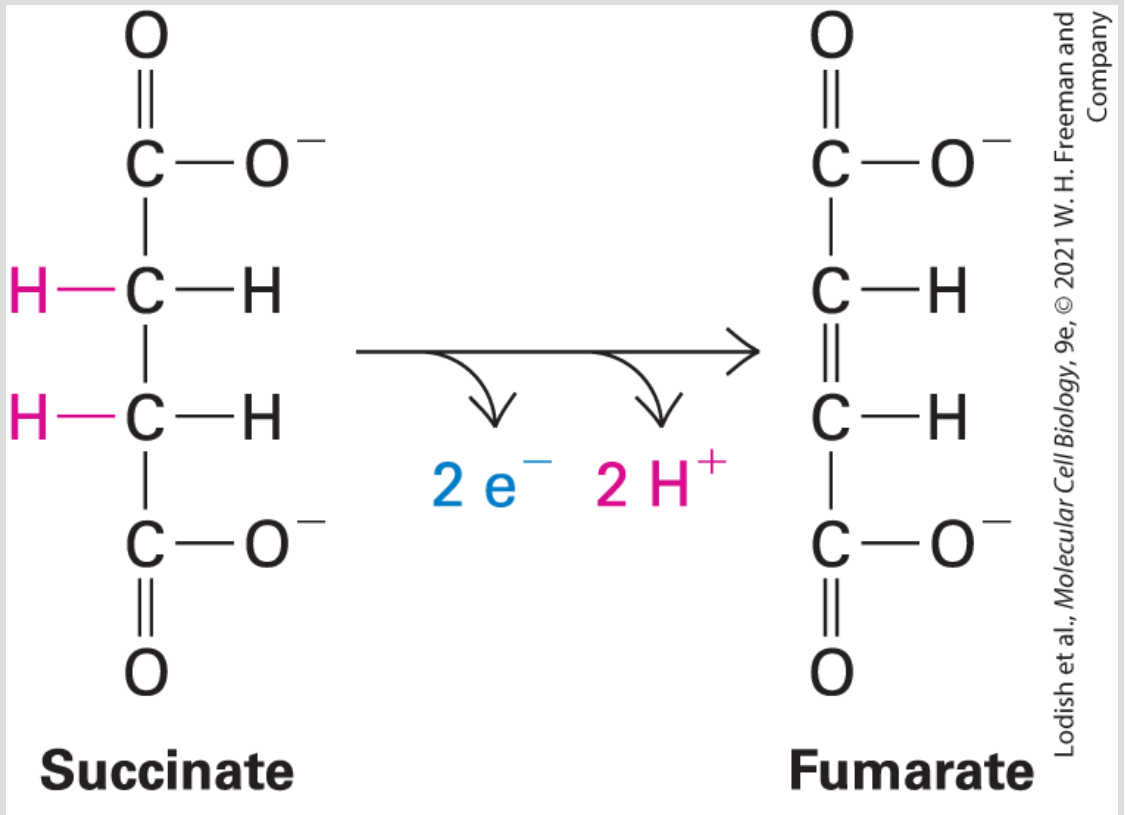


FIGURE 2-32 Conversion of succinate to fumarate. In this oxidation reaction, which occurs in mitochondria as part of the citric acid cycle, succinate loses two electrons and two protons. These protons and electrons are transferred to FAD, reducing it to $FADH_2$.

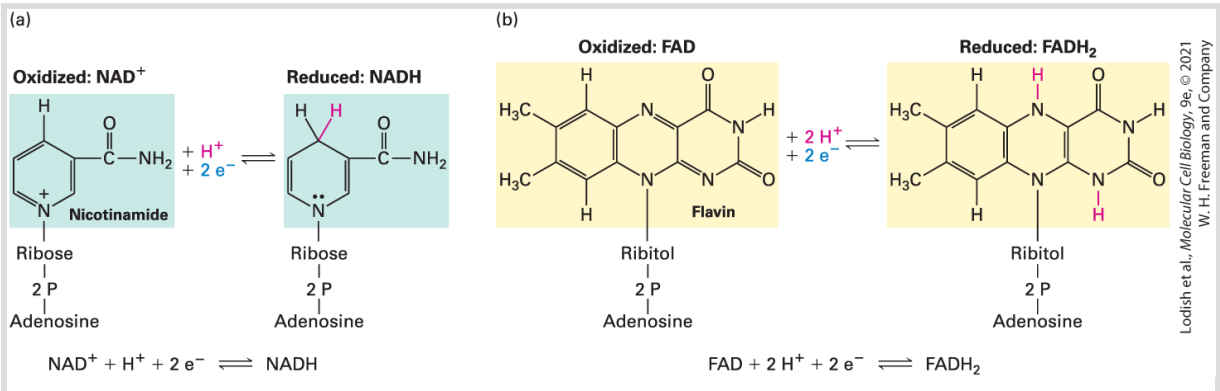


FIGURE 2-33 The electron-carrying coenzymes NAD^+ and FAD. (a) NAD^+ (nicotinamide adenine dinucleotide) is reduced to NADH by the addition of two electrons and one proton simultaneously. In many biological redox reactions, a pair of hydrogen atoms (two protons

and two electrons) is removed from a molecule. In some cases, one of the protons and both electrons are transferred to NAD^+; the other proton is released into solution. (b) FAD (flavin adenine dinucleotide) is reduced to $FADH_2$ by the addition of two electrons and two protons, as occurs when succinate is converted to fumarate (see Figure 2-32). In this two-step reaction, addition of one electron together with one proton first generates a short-lived semiquinone intermediate (not shown), which then accepts a second electron and proton.

To describe redox reactions, such as the reaction of ferrous ion (Fe^{2+}) and oxygen (O_2), it is easiest to divide them into two half-reactions:

$$\begin{aligned} \textit{Oxidation of } Fe^{2+}: &\quad 2\,Fe^{2+} \rightarrow 2\,Fe^{3+} + 2\,e^- \\ \textit{Reduction of } O_2: &\quad 2\,e^- + \frac{1}{2}\,O_2 \rightarrow O^{2-} \end{aligned}$$

In this case, the reduced oxygen (O^{2-}) readily reacts with two protons to form one water molecule (H_2O). The readiness with which an atom or a molecule *gains* an electron is its **reduction potential (*E*)**. The tendency to *lose* electrons, the **oxidation potential**, has the same magnitude as the reduction potential for the reverse reaction, but has the opposite sign.

Reduction potentials are measured in volts (V) from an arbitrary zero point set at the reduction potential of the following half-reaction under standard conditions (25 °C, 1 atm, and reactants at 1 M):

$$\mathrm{H^+ + e^-} \underset{\text{oxidation}}{\overset{\text{reduction}}{\rightleftharpoons}} \frac{1}{2}\,\mathrm{H_2}$$

The value of the reduction potential (E) for a molecule or an atom under standard conditions is its standard reduction potential, E'_0. A molecule or an ion with a positive E'_0 has a higher affinity for electrons than the $\mathrm{H^+}$ ion does under standard conditions. Conversely, a molecule or ion with a negative E'_0 has a lower affinity for electrons than the $\mathrm{H^+}$ ion does under standard conditions. Like the values of $\Delta G^{\circ\prime}$, standard reduction potentials may differ somewhat from those found under the conditions in a cell because the concentrations of reactants in a cell are not 1 M.

In a redox reaction, electrons move spontaneously toward atoms or molecules having *more positive* reduction potentials. In other words, a molecule having a more negative reduction potential can transfer electrons spontaneously to, or reduce, a molecule with a more positive reduction potential. In this type of reaction, the change in electric potential ΔE is the sum of the reduction and oxidation potentials for the two half-reactions. The ΔE for a redox reaction is related to the change in free energy ΔG by the following expression:

$$\Delta G\ (\mathrm{cal/mol}) = -n\ (23{,}064)\ \Delta E\ (\mathrm{volts}) \tag{2-11}$$

where n is the number of electrons transferred. Note that a redox reaction with a positive ΔE value will have a negative ΔG and thus will tend to proceed spontaneously from left to right.

KEY CONCEPTS OF SECTION 2.4

Biochemical Energetics

- The change in free energy, ΔG, is the most useful measure for predicting the potential of chemical reactions to occur spontaneously in biological systems. Chemical reactions tend to proceed spontaneously in the direction for which ΔG is negative. The magnitude of ΔG is independent of the reaction rate. A reaction with a negative ΔG is referred to as thermodynamically favorable.
- The chemical free-energy change, $\Delta G^{\circ\prime}$, equals $-2.3\ RT \log K_{\text{eq}}$. Thus the value of $\Delta G^{\circ\prime}$ can be calculated from the experimentally determined concentrations of reactants and products at equilibrium.
- The rate of a reaction depends on the activation energy needed to energize reactants to a transition state. Catalysts such as enzymes speed up reactions by lowering the activation energy of the transition state.
- A chemical reaction having a positive ΔG can proceed spontaneously if it is coupled with a reaction having a negative ΔG of larger magnitude.
- Many otherwise energetically unfavorable cellular processes are driven by the hydrolysis of phosphoanhydride bonds in ATP (see Figure 2-31).
- Directly or indirectly, light energy captured by photosynthesis in plants, algae, and photosynthetic bacteria is the ultimate source of chemical energy for nearly all cells on Earth.
- An oxidation reaction (loss of electrons) is always coupled with a reduction reaction (gain of electrons).
- Biological oxidation and reduction reactions are often coupled by electron-carrying coenzymes such as NAD^+ and FAD (see Figure 2-33).
- Oxidation-reduction reactions with a positive ΔE have a negative ΔG and thus tend to proceed spontaneously.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

acid
activation energy
adenosine triphosphate (ATP)
aerobic
allostery
alpha carbon atom
amino acid
amphipathic (or amphiphilic)
anion
base
buffer
carbohydrate
catabolism

catalyst
cation
cellular respiration
cellulose
chemical equilibrium
chemical potential energy
concentration gradient
covalent bond
dipole
dipole moment
disaccharide
dissociation constant (K_d)
DNA (deoxyribonucleic acid)
electric potential
endergonic
endothermic
enthalpy (H)
entropy (S)
enzyme
equilibrium constant (K_{eg})
exergonic
exothermic
FAD (flavin adenine dinucleotide)
fatty acids
free-energy change (ΔG)
glycogen
glycolysis
glycosaminoglycan

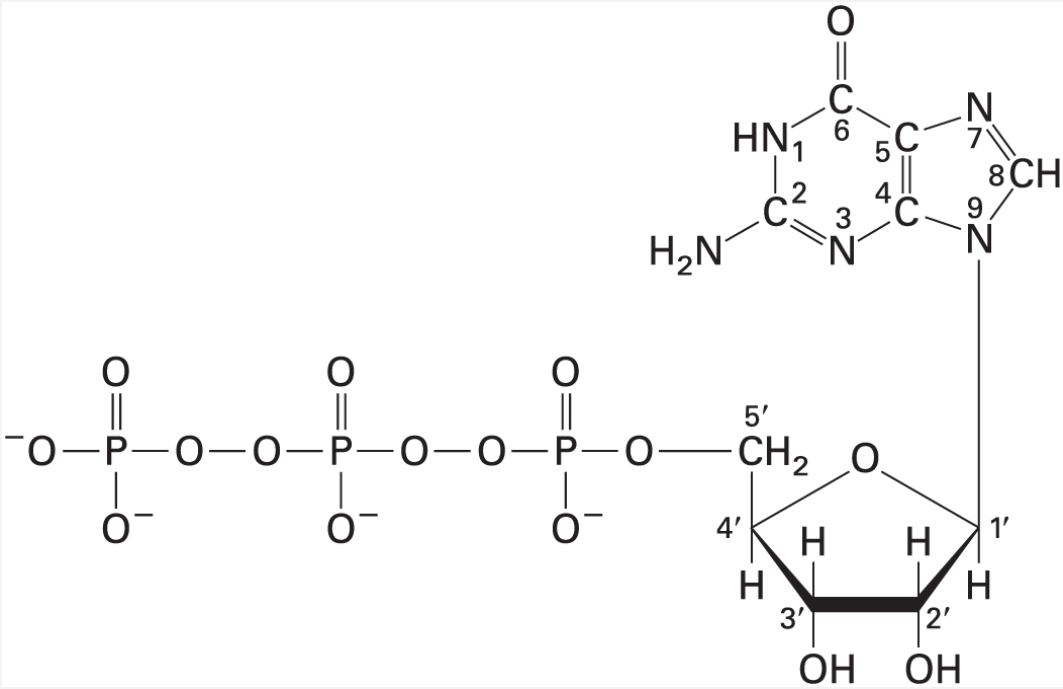
glycosidic bond
hexose
hydrocarbon
hydrogen bond
hydrophilic
hydrophobic
hydrophobic effect
ionic bond (ionic interaction)
kinetic energy
ligand
macromolecule
mitogen
molecular complementarity
monomer
monosaccharide
NAD^+ (nicotinamide adenine dinucleotide)
noncovalent interactions
nonpolar
nucleic acid
nucleoside
nucleotide
oxidation
oxidation potential
pentose
peptide bond
peptidoglycan
pH
phosphoanhydride bond

phosphodiester bond
phosphoglyceride
phospholipid
photosynthesis
polar
polymer
polysaccharide
polyunsaturated
potential energy
protein
purine
pyrimidine
rate constant
receptor
redox reaction
reduction
reduction potential (*E*)
residue
RNA (ribonucleic acid)
saturated
starch
steady state
stereoisomer
transition state
transition-state intermediate
triacylglycerol (triglyceride)
unsaturated
van der Waals interaction

Review the Concepts

1. The gecko is a reptile with an amazing ability to climb smooth surfaces, including glass. Geckos appear to stick to smooth surfaces via van der Waals interactions between septa on their feet and the smooth surface. How is this method of stickiness advantageous over covalent interactions? Given that van der Waals forces are among the weakest molecular interactions, how can the gecko's feet stick so effectively?
2. The K^+ channel is an example of a transmembrane protein (a protein that spans the phospholipid bilayer of the plasma membrane). What types of amino acids are likely to be found (a) lining the channel through which K^+ passes, (b) in contact with the hydrophobic core of the phospholipid bilayer containing fatty acyl groups, (c) in the cytosolic domain of the protein, and (d) in the extracellular domain of the protein?
3. V-M-Y-F-E-N: This is the single-letter amino acid abbreviation for a peptide. What is the net charge of this peptide at pH 7.0? An enzyme called a *protein tyrosine kinase* can attach a phosphate to the hydroxyl group of a tyrosine (Y). What is the net charge of the peptide at pH 7.0 after it has been phosphorylated by a tyrosine kinase? What is the likely source of phosphate used by the kinase for this reaction?

4. Disulfide bonds help to stabilize the three-dimensional structure of proteins. What amino acids are involved in the formation of disulfide bonds? Does the formation of a disulfide bond increase or decrease entropy (ΔS)?
5. In the 1960s, the drug *thalidomide* was prescribed to pregnant women to treat morning sickness. However, thalidomide caused severe limb defects in the children of some women who took the drug, and its use for morning sickness was discontinued. It is now known that thalidomide was administered as a mixture of two stereoisomeric compounds, one of which relieved morning sickness and the other of which was responsible for the birth defects. What are stereoisomers? Why might two such closely related compounds have such different physiological effects?
6. Name the compound shown below.



Is this nucleotide a component of DNA, RNA, or both? Name one other function of this compound.

7. The chemical basis of blood-group specificity resides in the carbohydrates displayed on the surfaces of red blood cells. Carbohydrates have the potential for great structural diversity. Indeed, the structural complexity of the oligosaccharides that can be formed from four sugars is greater than that of the oligopeptides that can be formed from four amino acids. What properties of carbohydrates make this great structural diversity possible?
8. Calculate the pH of 1 L of pure water at equilibrium. How will the pH change after 0.008 moles of the strong base sodium hydroxide (NaOH) are dissolved in the water? Now, calculate the pH of a 50 mM aqueous solution of the weak acid 3-(*N*-morpholino) propane-1-sulfonic acid (MOPS) in

which 61 percent of the solute is in its weak acid form and 39 percent is in the form of MOPS's corresponding base (the pK_a for MOPS is 7.20). What is the final pH after 0.008 moles of NaOH are added to 1 L of this MOPS buffer?

9. Ammonia (NH_3) is a weak base that under acidic conditions becomes protonated to the ammonium ion in the following reaction:

$$NH_3 + H^+ \rightarrow NH_4{}^+$$

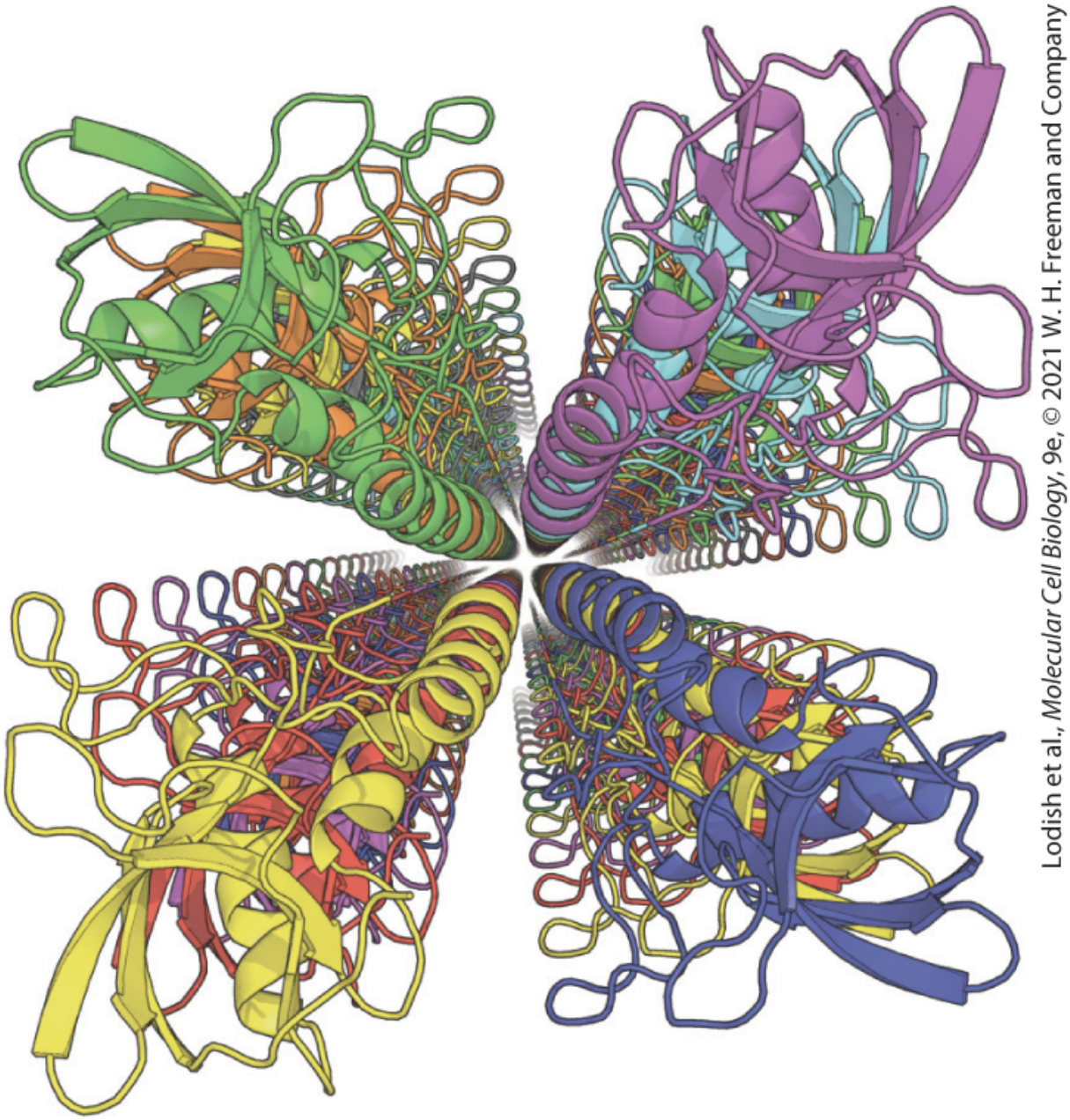
NH_3 freely permeates biological membranes, including those of lysosomes. The lysosome is a subcellular organelle with a pH of about 4.5–5.0; the pH of cytoplasm is about 7.0. What is the effect on the pH of the fluid content of lysosomes when cells are exposed to ammonia? *Note:* Ammonium ($NH_4{}^+$) does not diffuse freely across membranes.

10. Consider the binding reaction $L + R \rightarrow LR$, where L is a ligand and R is its receptor. When 1×10^{-3} M of L is added to a solution containing 5×10^{-2} M of R, 90 percent of the L binds to form LR. What is the K_{eq} of this reaction? How will the K_{eq} be affected by the addition of a protein that facilitates (catalyzes) this binding reaction? What is the dissociation equilibrium constant K_d?
11. What is the ionization state of phosphoric acid in the cytoplasm? Why is phosphoric acid such a physiologically important compound?

12. The $\Delta G^{\circ\prime}$ for the reaction $\mathrm{X} + \mathrm{Y} \rightarrow \mathrm{XY}$ is -1000 cal/mol. What is the ΔG at 25 $^{\circ}$C (298 $^{\circ}$Kelvin) starting with 0.01 M each of X, Y, and XY? Suggest two ways one could make this reaction energetically favorable.
13. According to health experts, saturated fatty acids, which come from animal fats, are a major factor contributing to coronary heart disease. What distinguishes a saturated fatty acid from an unsaturated fatty acid, and to what does the term *saturated* refer? Recently, trans unsaturated fatty acids, or trans fats, which raise total cholesterol levels in the body, have also been implicated in heart disease. How does the cis stereoisomer differ from the trans configuration, and what effect does the cis configuration have on the structure of the fatty acid chain?
14. Chemical modifications of amino acids contribute to the diversity and function of proteins. For instance, γ-carboxylation of specific amino acids is required to make some proteins biologically active. What particular amino acid undergoes this modification, and what is its biological relevance? Warfarin, a derivative of coumarin, which is present in many plants, inhibits γ-carboxylation of this amino acid and was used in the past as a rat poison. At present, it is also used clinically in humans. What patients might be prescribed warfarin and why?

CHAPTER 3

Protein Structure and Function



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Computationally designed, hypothetical four-stranded fiber protein viewed down the fiber axis. Using advanced methods for protein design, the amino acid sequence of the staphylococcal nuclease protein was modified so that the individual, folded protein chains (shown in different colors) would stack into long strands. The design predicts that four long strands would assemble into a four-stranded fiber held together by the binding of hydrophobic helices located in the center of the fiber.

[Data from H. Shen et al., 2018, *Science* **362**:705–709.]

OUTLINE

3.1 Hierarchical Structure of Proteins
3.2 Protein Folding
3.3 Protein Binding and Enzyme Catalysis
3.4 Regulating Protein Function
3.5 Purifying, Detecting, and Characterizing Proteins
3.6 Proteomics

Proteins, which are polymers of amino acids called *polypeptides*, fold into three-dimensional structures of many sizes and shapes. Their three-dimensional diversity principally reflects variations in their amino acid sequences and in the lengths of the polypeptides. In general, a polypeptide will fold into only one or a few closely related three-dimensional shapes — called **conformations**. A key concept in understanding how a protein works is that *its function is often derived from its three-dimensional structure, and its three-dimensional structure is determined by its amino acid sequence and the noncovalent interactions that stabilize its structure*.

In many cases, the conformation, and thus the function, of a protein can change when that protein associates noncovalently or covalently with other molecules or ions. Through these associations, the function — also called the *activity* — of a protein can be *regulated* (e.g., turned "on" or "off," or "up" or "down") to permit cells to adapt to changing conditions. Altered conditions include changes in nutrient availability (Chapter 21), hormonal signaling (Chapters 15 and 16), communication with other cells (Chapters 20 and 23), the

developmental state of the organism, and the presence of pathogens (Chapter 24), along with many other factors.

How many proteins are there in a typical eukaryotic cell? We can calculate that there are about 7.9×10^9 protein molecules in a mammalian hepatocyte (a liver cell). (This calculation is worked out in Section 3.1.) It is estimated that a hepatocyte contains about 10,000 different proteins; thus each cell, *on average*, contains close to a million molecules of each type of protein. In reality though, the abundances of different proteins vary widely, from the quite rare insulin-binding receptor protein (2×10^4 molecules per cell) to the structural protein actin (5×10^8 molecules per cell). Every cell closely regulates the abundance of each of its proteins so that they are present in the appropriate quantities to support the cell's needs at any given time. We will learn more about the mechanisms used by cells to regulate protein levels later in this chapter and in Chapters 8 and 9.

Because of their many different shapes and chemical properties, proteins can perform a dazzling array of functions inside and outside cells that either are essential for life or provide a selective evolutionary advantage to the cell or organism that contains them. It is, therefore, not surprising that characterizing the structures and activities of proteins — and how these change in response to regulation — is a fundamental prerequisite for understanding how cells work. Much of this textbook is devoted to examining how

proteins act with one another or with other types of molecules (e.g., DNA) to allow cells to live and function properly. In this chapter we emphasize the basic principles underlying the structures and functions of proteins and the regulation of their activity. We also look at some of the many methods used to study them.

Although their structures are diverse, most proteins can be grouped into a few broad functional classes. *Structural proteins*, for example, determine the shapes of cells and their extracellular environments and serve as guide wires or rails to direct the intracellular movement of molecules and organelles. Structural proteins are usually formed by the assembly of multiple protein subunits into very large and often very long structures. *Scaffold proteins* bring other proteins together into ordered arrays to perform specific functions more efficiently than those proteins would if they were not assembled together. *Enzymes* are proteins that catalyze chemical reactions (sometimes called *molecular transformations*). These molecular transformations are the fundamental activities in metabolic pathways (Chapter 12). They can also result in the modification of proteins to alter their activities (e.g., phosphorylation and dephosphorylation to activate and inactivate other proteins). *Membrane transport proteins* are embedded in cellular membranes and permit the flow of ions and molecules across the membranes. *Regulatory proteins* act as signals, sensors, and switches to control activities of cells by altering the functions of other proteins and genes. Regulatory proteins include *signaling*

proteins, such as some hormones and cell-surface receptors that transmit extracellular signals to the cell interior. *Motor proteins* are responsible for the *movement* of other proteins, organelles, and cells — even whole organisms. There are additional proteins that don't fit neatly into one of these classes. For example, fish that live in frigid waters — the Antarctic notothenioids and Arctic cods — have antifreeze proteins in their circulatory systems to prevent water crystallization. Many proteins are members of more than one class, such as some cell-surface signaling receptors that are enzymes and regulatory proteins because they transmit signals from outside to inside cells by catalyzing chemical reactions. To accomplish their diverse missions efficiently, some proteins assemble into very large complexes, often called *molecular machines*.

How do proteins perform so many diverse functions? They do so by exploiting three simple mechanisms. The most fundamental is *binding*: proteins bind to one another, to other macromolecules such as DNA, and to small molecules and ions. Binding is based on molecular complementarity between a protein and its binding partner, as described in Chapter 2. A second mechanism is enzymatic *catalysis*. In enzymes, the proper conformation of the protein will place some amino acid side chains and some carboxyl and amino groups of its backbone into positions that permit the catalysis of covalent bond rearrangements in other molecules (called *substrates* of the enzyme). A third mechanism is *regulation* of protein activity. Typically, a protein's shape or activity is altered

through noncovalent (binding) or covalent (catalysis) association of molecules or ions with the protein. In many cases, this binding or catalysis induces a conformational change in the protein and that influences its activity.

A complete understanding of how proteins permit cells to live and thrive requires that we identify and characterize all of the proteins used by a cell. In a sense, molecular cell biologists want to compile a complete protein "parts list" and then construct a "user's manual" that describes how these proteins work. Compiling a comprehensive inventory of proteins has become feasible in recent years thanks to the sequenced genomes — complete sets of genes — of many organisms. From a computer analysis of a genome's sequence, researchers can deduce the amino acid sequences and approximate number of different types of proteins the genome encodes (see Chapter 6). It is also possible to determine the sequences and relative amounts of a substantial fraction of the messenger RNAs (mRNAs) in individual cells or most of the mRNAs (the transcriptome) from a collection of similar cells and thus deduce what subset of proteins encoded in the genome is made (*expressed*) in a given type of cell. DNA and mRNA sequencing are indirect methods of characterizing the potential collection of proteins in cells. There are also methods that directly measure the collection of proteins in samples of cells, which we discuss in Sections 3.5 and 3.6.

The term *proteome* was coined to refer to the entire protein complement of an organism, organ, or a particular type of cell within an organism. The human genome contains about 21,500 genes that encode proteins. However, variations in mRNA production, such as alternative splicing (see Chapter 9), and more than a hundred types of protein modifications may generate hundreds of thousands of distinct types of human proteins. By comparing the sequences and structures of proteins of unknown function with those of proteins of known function, scientists can often deduce much about what the unknown proteins do. In the past, a protein's function was characterized by genetic, biochemical, or physiological methods before the particular protein had been identified. In the modern genomic and proteomic era, a protein is often identified before its function is determined.

In this chapter, we begin our study of how the structure of a protein gives rise to its function, a theme that recurs throughout this book (Figure 3-1). In Section 3.1, we examine how linear chains of amino acids are arranged in a three-dimensional structural hierarchy. The next section discusses how proteins fold into these structures. We then turn to protein function, focusing on enzymes, those proteins that catalyze chemical reactions. Our focus on protein structure and function is part of a broader interest in cell biology in the principles relating biological structure and function. These principles were initially formulated by biologists Johann von Goethe (1749–1832), Ernst Haeckel (1834–1919), and D'Arcy Thompson (1860–1948),

whose work has been widely influential in biology and beyond. Indeed, their ideas greatly influenced the school of organic architecture pioneered in the early twentieth century that is epitomized by the dicta “form follows function” (Louis Sullivan) and “form is function” (Frank Lloyd Wright). After considering protein structure and function, we turn to various mechanisms that cells use to control the activities and life spans of proteins. Particularly important control mechanisms are allosteric effector-binding, covalent phosphorylation, and ubiquitinylation of proteins. The chapter concludes with a discussion of commonly used techniques for identifying, isolating, and characterizing proteins, and a discussion of the burgeoning field of proteomics.

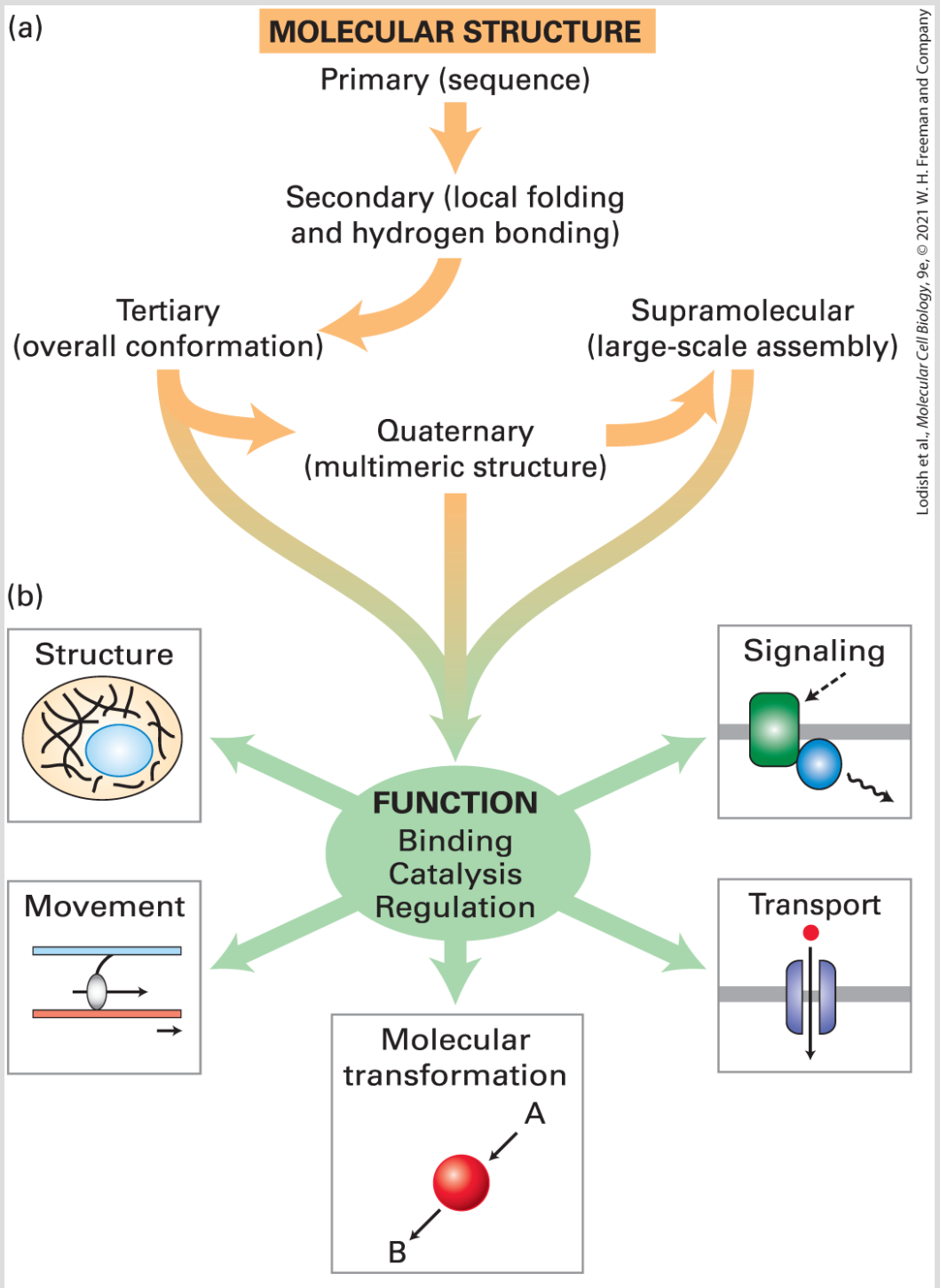


FIGURE 3-1 Overview of protein structure and function. (a) Proteins have a hierarchical structure. A polypeptide's linear sequence of amino acids linked by peptide bonds (primary structure) folds into local helices or sheets (secondary structure) that pack into a complex three-dimensional shape (tertiary structure). Some individual polypeptides associate into multichain complexes

(quaternary structure), which in some cases can be very large, consisting of tens to hundreds of subunits (supramolecular complexes). (b) Proteins perform numerous functions, including organizing the genome, organelles, cytoplasm, protein complexes, and membranes in three-dimensional space (*structure*); monitoring the environment and transmitting information (*signaling*); moving small molecules and ions across membranes (*transport*); catalyzing chemical reactions that result in the *molecular transformation* of one molecule into another; and generating force for *movement* (via motor proteins). These functions and others arise from mechanisms involving *binding*, *catalysis*, and *regulation*.

3.1 Hierarchical Structure of Proteins

Here we consider the architecture of proteins at four levels of organization: primary, secondary, tertiary, and quaternary (Figure 3-2).

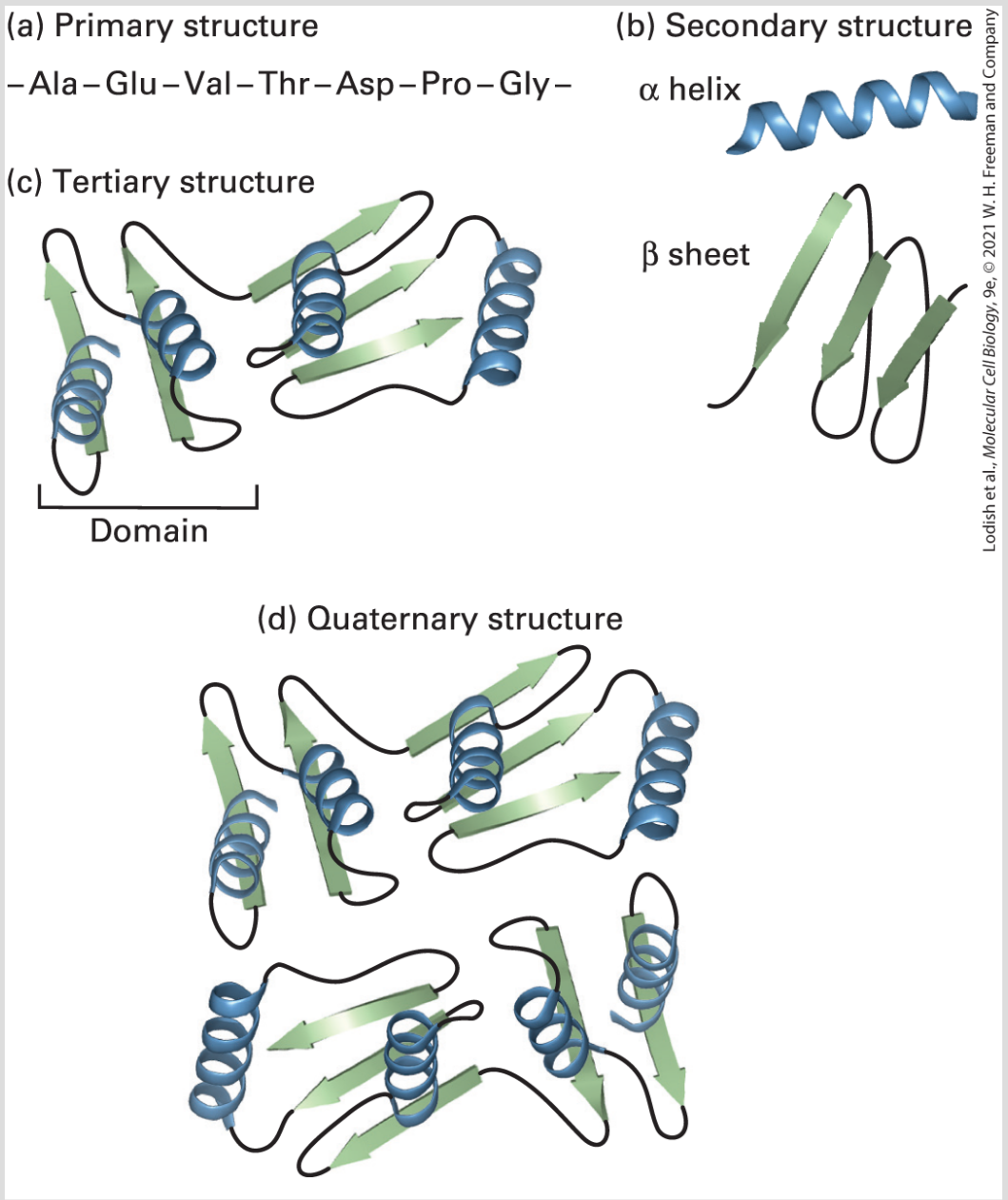
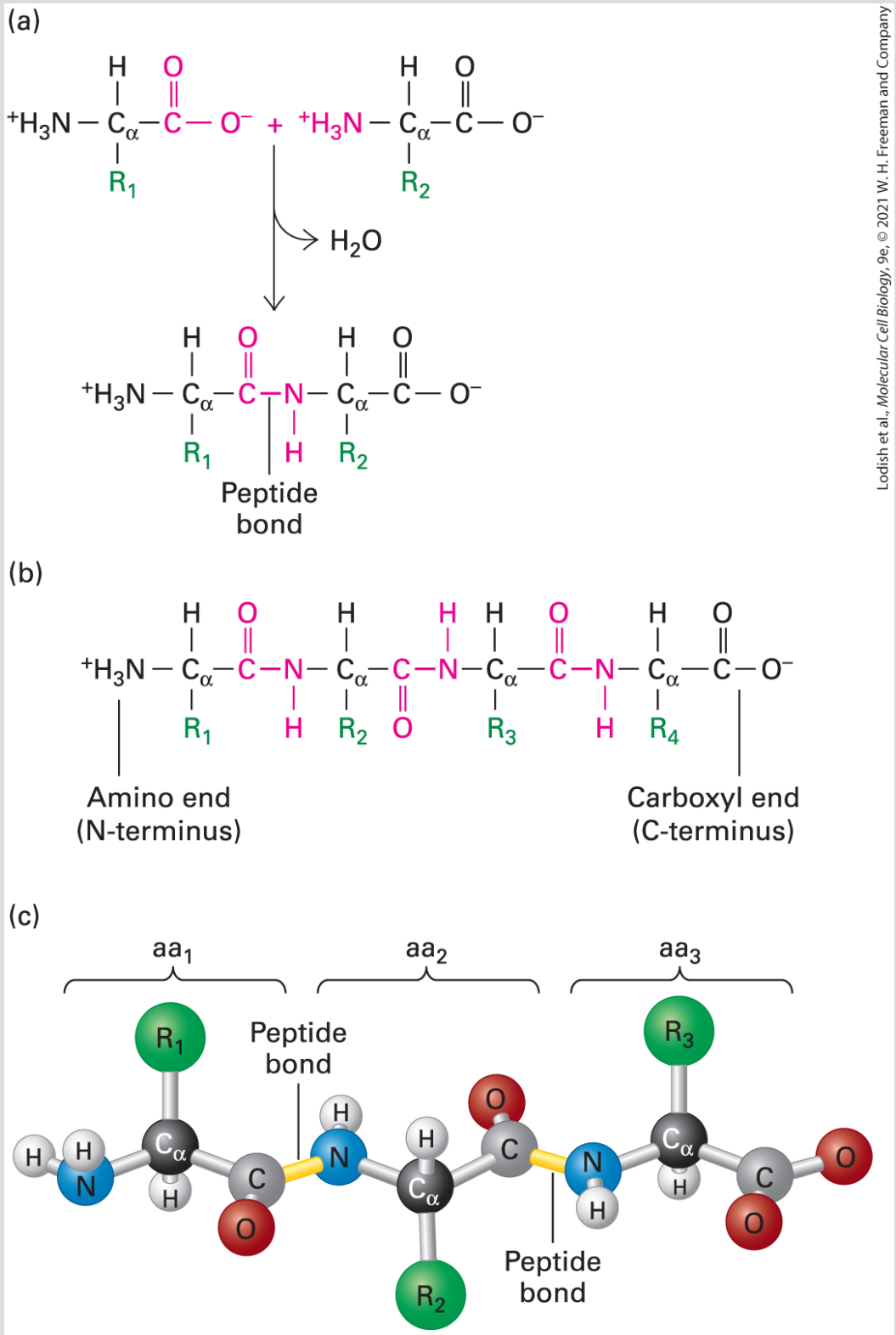


FIGURE 3-2 Four levels of protein hierarchy. (a) The linear sequence of amino acids linked together by peptide bonds is the primary structure. (b) Folding of the polypeptide chain into local α helices or β sheets represents secondary structure. (c) Secondary structural elements, together with various loops and turns in a single polypeptide chain, pack into a larger, independently stable tertiary structure, which may include distinct domains. (d) Some proteins consist of more than one polypeptide associated together in a quaternary structure.

The Primary Structure of a Protein Is Its Linear Arrangement of Amino Acids

As discussed in Chapter 2, proteins are polymers constructed out of 20 different types of amino acids. Individual amino acids are linked together in linear chains by covalent amide bonds, called **peptide bonds**, formed between the carboxyl group of one amino acid and the amino group of the adjacent amino acid (Figure 3-3a). An amino acid in a protein chain is sometimes referred to as a *residue*. The repeated amide N, α carbon (C_α), carbonyl C, and oxygen atoms of each amino acid residue form the *backbone* of a protein molecule from which the various side-chain groups project (Figure 3-3b, c). Thus one end of a protein has a free (unlinked) amino group (the *N-terminus*) and the other end has a free carboxyl group (the *C-terminus*). The **primary structure** of a protein is simply the linear covalent arrangement, or sequence, of the amino acid residues that compose it and is conventionally written with its N-terminal amino acid on the left and its C-terminal amino acid on the right. The amino acids are numbered sequentially starting from the N-terminus.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 3-3 Structure of a polypeptide. (a) Individual amino acids are linked together by peptide bonds, which form via reactions that result in a loss of

water. R_1, R_2, and so on, represent the side chains (R groups) of amino acids. (b) Linear polymers of amino acids linked by peptide bonds are called *polypeptides*, which have a free (unlinked) amino end (N-terminus) and a free carboxyl end (C-terminus). (c) A ball-and-stick model shows peptide bonds (yellow) linking the amino nitrogen atom (blue) of one amino acid (aa) with the carbonyl carbon atom (gray) of an adjacent one in the chain. The R groups (green) extend from the α-carbon atoms (black) of the amino acids. These side chains largely determine the distinct properties of individual proteins.

The first primary structure of a protein determined was that of insulin in the early 1950s. Today the number of known protein sequences exceeds 10 million and is growing daily. Many terms are used to denote the chains formed by the polymerization of amino acids. A short chain of amino acids linked by peptide bonds and having a defined sequence is called an **oligopeptide**, or simply a **peptide**; longer chains are referred to as **polypeptides**. Peptides generally contain fewer than 20–30 amino acid residues, whereas polypeptides are often 200–500 residues long and can be longer. The longest protein described to date is the muscle protein titin, some forms of which are more than 34,000 residues long. We generally reserve the term **protein** for a polypeptide (or complex of polypeptides) that has a well-defined three-dimensional structure, although there are exceptions described later in this chapter.

The size of a protein or a polypeptide is expressed either as its mass in **daltons** (a dalton is 1 atomic mass unit) or as its molecular weight (MW), which is a dimensionless number equal to the mass in daltons. For example, a 10,000-MW protein has a mass of 10,000

daltons (Da), or 10 kilodaltons (kDa). In Section 3.5, we will consider different methods for measuring the sizes and other physical characteristics of proteins. The precise molecular weight of a protein that has not been covalently modified is readily determined by summing up the weights of all of its constituent amino acids as determined from its amino acid sequence. The proteins encoded by the yeast genome, for example, have an average molecular weight of 52,728 and contain, on average, 466–amino acid residues. The average molecular weight of amino acids in proteins is 113, taking into account their average relative abundances. This value can be used to estimate the number of residues in a protein of unknown sequence if you know its molecular weight or, conversely, to estimate from the number of residues in a protein its likely molecular weight. Covalent modification of one or more amino acids in a protein — for example, by phosphorylation or glycosylation (see Chapters 2 and 13) — alters the mass of those residues and thus the mass of the protein in which they reside.

In the introduction we noted that a eukaryotic cell, on average, contains close to 7.9×10^9 molecules of protein. How do we estimate this? Let's do a simple calculation for a hepatocyte (a major type of cell in the mammalian liver). A hepatocyte, roughly a cube 15 μm (0.0015 cm) on a side, has a volume of 3.4×10^{-9} cm^3 (or milliliters, ml). Assuming a cell density of 1.03 g/ml, the cell would weigh 3.5×10^{-9} g. Since protein accounts for approximately 20 percent of a cell's weight, the total weight of cellular protein is 7×10^{-10} g.

Assuming that an average protein has a molecular weight of about 53,000 g/mol, we can calculate the total number of protein molecules per hepatocyte as about 7.9×10^9 from the total protein weight and Avogadro's number, the number of molecules per mole of any chemical compound (6.02×10^{23}).

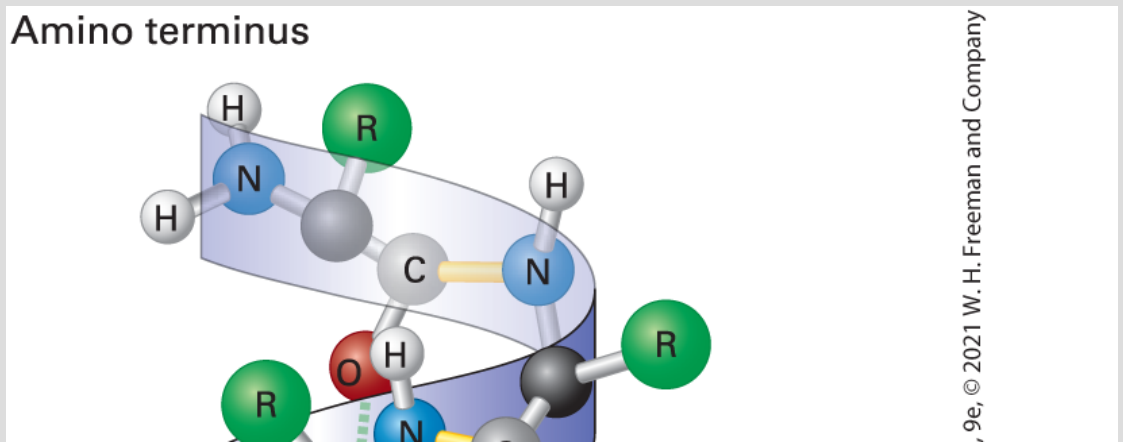
Secondary Structures Are the Core Elements of Protein Architecture

The second level in the hierarchy of protein structure is **secondary structure**. Secondary structures are stable spatial arrangements of segments of a polypeptide chain held together by hydrogen bonds between backbone amide and carbonyl groups and often involving repeating structural patterns. The propensity of a segment of a polypeptide chain to form any given secondary structure depends on its amino acid sequence (see Section 3.2). A single polypeptide may contain multiple types of secondary structure in various portions of the chain, depending on its sequence. The principal secondary structures are the **alpha (α) helix**, the **beta (β) sheet**, and the short, U-shaped **beta (β) turn**. Parts of a polypeptide that don't form these secondary structures may nevertheless have a well-defined, stable shape and are said to have an *irregular* structure. The term *random coil* applies to highly flexible parts of a polypeptide chain that have no stable, fixed three-dimensional structure. Random coils are said to be *disordered*. In an average protein, 60

percent of the polypeptide chain exists as α helices and β sheets; the remainder of the molecule is in irregular structures, coils, and turns or is disordered. Thus α helices and β sheets are the major internal supportive elements in most proteins. Here we explore the shapes of secondary structures and the forces that favor their formation. In later sections, we examine how arrays of secondary structure fold together into larger, more complex arrangements called *tertiary structure*.

The α Helix

In a polypeptide segment folded into an α helix, the backbone forms a spiral structure in which the carbonyl oxygen atom of each peptide bond is hydrogen-bonded to the amide hydrogen atom of the amino acid four residues farther along the chain in the direction of the C-terminus (Figure 3-4). Within an α helix, all the backbone amino and carboxyl groups are hydrogen-bonded to one another (conferring substantial stability) except at the very beginning and end of the helix.



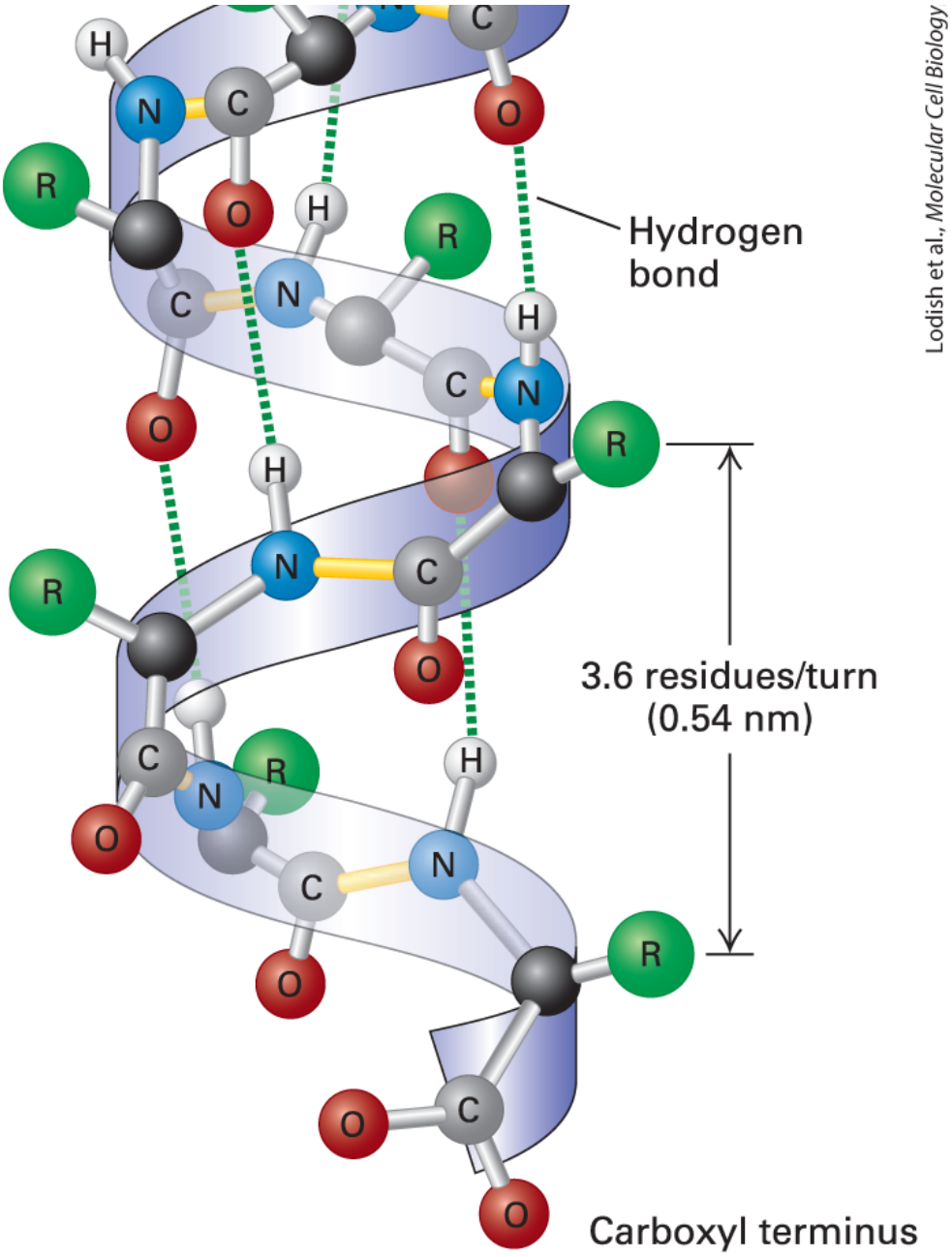


FIGURE 3-4 The α helix, a common secondary structure in proteins. The polypeptide backbone (highlighted as a ribbon) is folded into a spiral that is held in place by hydrogen bonds between backbone oxygen and hydrogen atoms linked to nitrogen. Only hydrogens involved in bonding are shown. The outer surface of the helix is covered by the side-chain R groups (green). There is a complete turn of the spiral every 3.6 residues. An α helix 36 amino acids long has 10 turns of the helix and is 5.4 nm long (0.54 nm per turn).

The stable arrangement of hydrogen-bonded amino acids in the α helix holds the backbone in a straight, rodlike cylinder from which the side chains point outward. The characteristics of the side chains entirely determine the relative hydrophobic or hydrophilic quality of a particular helix within a protein. In water-soluble proteins, hydrophilic helices with polar side chains extending outward tend to be found on the outside surfaces, where they can interact with the aqueous environment, whereas hydrophobic helices with nonpolar, hydrophobic side chains tend to be buried within the core of the folded protein. Proteins embedded in the hydrophobic core of cellular membranes (see Chapter 10) often use one or more hydrophobic helices that are 20–25 residues long to cross the membrane. The amino acid proline is usually not found in α helices because the covalent bonding of its amino group with a carbon in the side chain (see Figure 2-14) prevents its participation in stabilizing the backbone through normal hydrogen bonding. While the classic α helix is the most intrinsically stable and most common helical form in proteins, there are variations, such as more tightly or loosely twisted helices. For example, in a specialized helix called a *coiled coil* (described several sections farther on), the helix is more tightly wound (3.5 residues and 0.51 nm per turn).

The β Sheet

The β sheet consists of laterally packed *β strands*. Each β strand is a short (5–8-residue), nearly fully extended polypeptide segment.

Hydrogen bonds form between the carbonyl oxygen atom of each residue in one β strand and the amide hydrogen atom of a residue in a separate, but adjacent, β strand. These hydrogen bonds are oriented perpendicularly to the chains of backbone atoms (Figure 3-5a). These distinct β strands (indicated as green and blue arrows in the figure) may be located either within a single polypeptide chain, with short or long loops between the β strand segments, or on different polypeptide chains in a protein composed of multiple polypeptides.

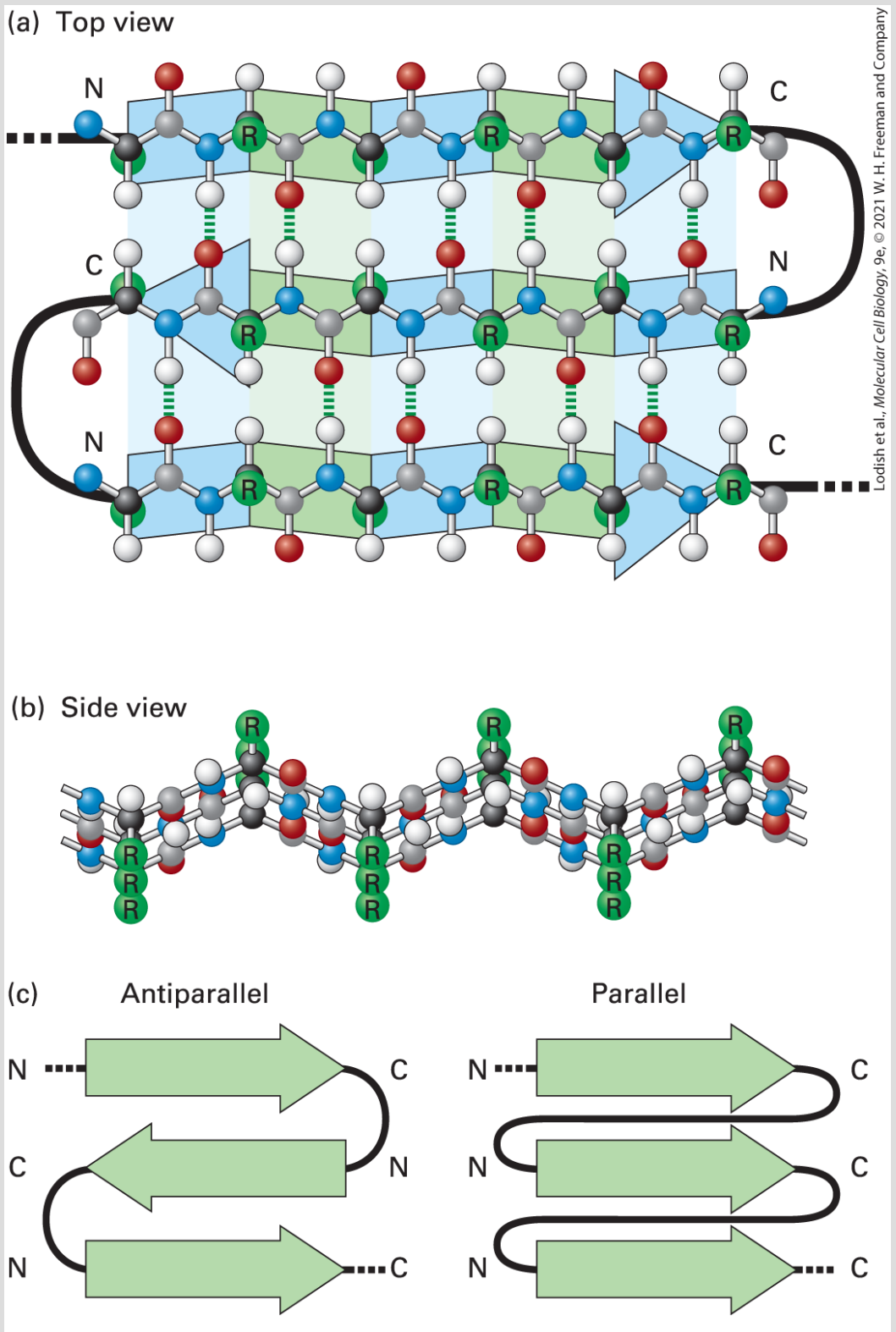


FIGURE 3-5 The β sheet, another common secondary structure in proteins. (a) Top view of a three-stranded β sheet. Each strand is highlighted by a ribbonlike

arrow with alternating blue and green segments that is pointed with an N-to-C orientation, with the loops of connecting residues indicated by thick black lines. In this antiparallel β sheet, each strand (arrow) points in the direction opposite to that of the adjacent strand. The stabilizing hydrogen bonds between the β strands are indicated by green dashed lines. (b) Side view of an antiparallel β sheet. The projection of the R groups (green) above and below the plane of the sheet is obvious in this view. The fixed bond angles in the polypeptide backbone produce a pleated contour represented in panel (a) by the alternating colored segments. (c) Top view of two β sheets, in which N-to-C orientations are represented by arrows. The individual strands are either antiparallel, in which the strands alternately point in opposite directions (*left*), or parallel, in which all strands point in the same direction (*right*).

Figure 3-5b shows how two or more β strands align into adjacent rows, forming a nearly two-dimensional β pleated sheet (or simply *pleated sheet*). Hydrogen bonds within the plane of the sheet hold the β strands together as the side chains stick out above and below the plane. Adjacent β strands can be oriented (N-terminus to C-terminus) in alternating opposite (antiparallel) directions (see Figures 3-5a and 3-5c, *left*) or in the same (parallel) direction (Figure 3-5c *right*). In some proteins, the β sheets curve around and form a cylinder, called a *β barrel*. When these proteins are embedded in membranes, the cylindrical beta sheet can form a hydrophilic central pore through which ions and small molecules may flow (see Chapter 10).

The β Turn

Composed of four residues, β turns form sharp, U-shaped bends located on the surface of a protein. Beta turns reverse the direction of the polypeptide backbone and are often stabilized by a hydrogen bond between their end residues (Figure 3-6). They help long polypeptides fold into highly compact structures. The direction of the polypeptide backbone may also be reversed by longer segments of the polypeptide that form bends or loops with varying conformations.

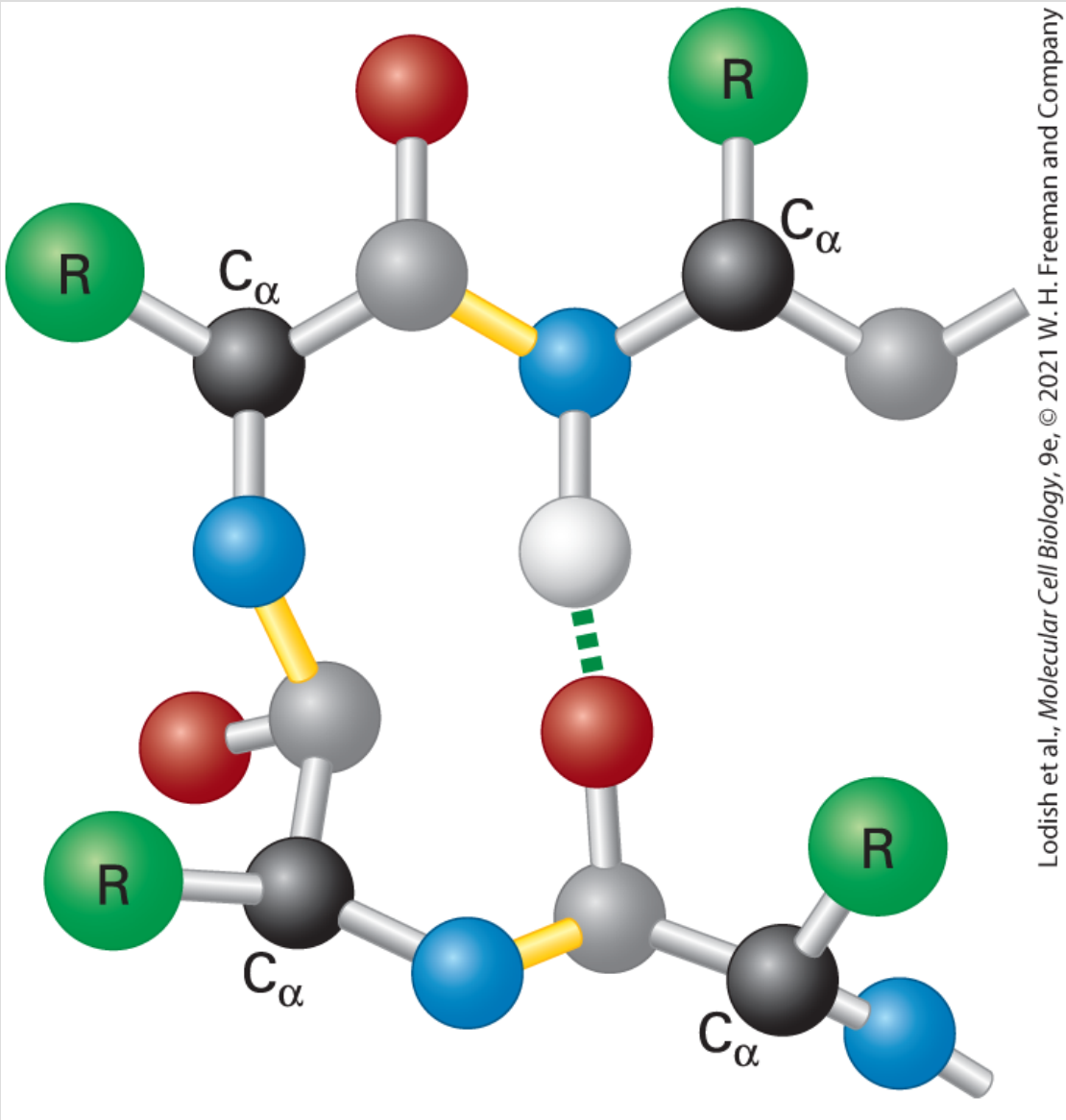


FIGURE 3-6 Structure of a β turn. Composed of four residues, β turns reverse the direction of a polypeptide chain (resulting in a 180° U-turn). Glycine and proline are commonly found in β turns. The lack of a large side chain in glycine and the presence of a built-in bend in proline allow the polypeptide backbone to fold into a tight U shape. The C_α carbons of the first and fourth residues are usually less than 0.7 nm apart, and those residues are often linked by a hydrogen bond. β turns facilitate the folding of long polypeptides into compact structures.

In many cases, the secondary structure of a polypeptide segment is determined primarily by its sequence, although in some cases long-range interactions between different parts of a polypeptide can influence the propensity of those parts to form a given secondary structure. The relationships of amino acid sequence and secondary structure have been analyzed in large numbers of known protein structures ($\sim$50,000). Investigators can use these relationships to predict the secondary structures embedded within the sequence of a polypeptide chain whose structure has not been determined ($\sim$250 million sequences are known, $\sim$86 billion amino acids!). Investigators have used a variety of sophisticated methods to predict secondary structure, including methods that consider evolutionary relationships or apply artificial intelligence (including deep learning based on artificial neural networks). These predictions are not perfect, but they are very good. When predictions of secondary structure are tested against experimentally determined structures, their accuracy can be as high as 84 percent. It is noteworthy that predictions of α helices are better than those for β sheets, which are more accurate than those for coils or loops. Such predictions can be used to predict the consequences of a mutation in a gene on the encoded protein's secondary structures, and possibly on its function.

Structural Motifs Are Regular Combinations of Secondary Structures

A particular combination of two or more secondary structures that forms a distinct three-dimensional structure, which appears in multiple proteins, is called a **structural motif**. A structural motif — an alternative name is *supersecondary structure* — is often associated with a specific function, such as binding to a particular ion or small molecule, such as calcium or ATP. Some structural motifs are stable after being isolated from the rest of a protein and are thus called *structural domains*, as we shall see shortly. However, other structural motifs do not form thermodynamically stable structures in the absence of other portions of the protein and are thus not considered independent structural domains.

One common structural motif is the α helix–based *coiled coil*. In this motif, α helices from two, three, or even four separate polypeptide chains coil about one another — resulting in a coil of coils; hence the name (Figure 3-7a). Many proteins, including fibrous proteins and DNA-regulating proteins called *transcription factors* (see Chapter 8) assemble into dimers or trimers by using a coiled-coil motif. The individual helices bind tightly to one another because each helix has a strip of aliphatic (hydrophobic, but not aromatic) side chains (leucine, valine, etc.) running along one side of the helix that interacts with a similar strip in the adjacent helix, thus sequestering the hydrophobic groups away from water and stabilizing the assembly of multiple independent helices. These hydrophobic strips are generated along only one side of the helix because the primary structure of each helix is composed of repeating seven–amino acid

units, called *heptads* or *heptad repeats*, in which the side chains of the first and fourth residues are aliphatic and the other side chains are often hydrophilic (see Figure 3-7a). Because hydrophilic side chains extend from one side of the helix and hydrophobic side chains extend from the opposite side, the overall helical structure is amphipathic. Because leucine frequently appears in the fourth positions and the hydrophobic side chains merge together like the teeth of a zipper, these structural motifs are also called **leucine zippers**.

A heptad repeat is an example of a *sequence motif*, a pattern of amino acids in a contiguous segment of a protein that can be found in many proteins, and sometimes in many copies within one protein. The sequence motif may be either an exact copy of the same sequence in every protein or a pattern of similar residues that do not have precisely the same sequence, such as the common spacing of aliphatic residues in a heptad repeat.

Many different structural motifs contain α helices. A common calcium-binding motif called the **EF hand** contains two short α helices connected by a loop (Figure 3-7b). This structural motif, one of several **helix-turn-helix** and **helix-loop-helix** structural motifs, is found in more than a hundred proteins and is used for sensing calcium levels. Ca^{2+} ion binds to oxygen atoms in conserved residues in the loop when the concentration of Ca^{2+} in the cell is high enough. Sometimes the binding induces a conformational

change in the protein bearing the EF hand, altering the protein's activity. Thus calcium concentrations can directly control the structures and functions of certain proteins. Somewhat different helix-turn-helix and **basic helix-loop-helix (bHLH)** structural motifs are used for protein binding to DNA and, consequently, for the regulation of gene activity (see Chapter 8). Yet another structural motif commonly found in proteins that bind RNA or DNA is the **zinc finger**. This motif contains three secondary structures — an α helix and two β strands with an antiparallel orientation — that form a fingerlike bundle held together by a zinc ion (Figure 3-7c).

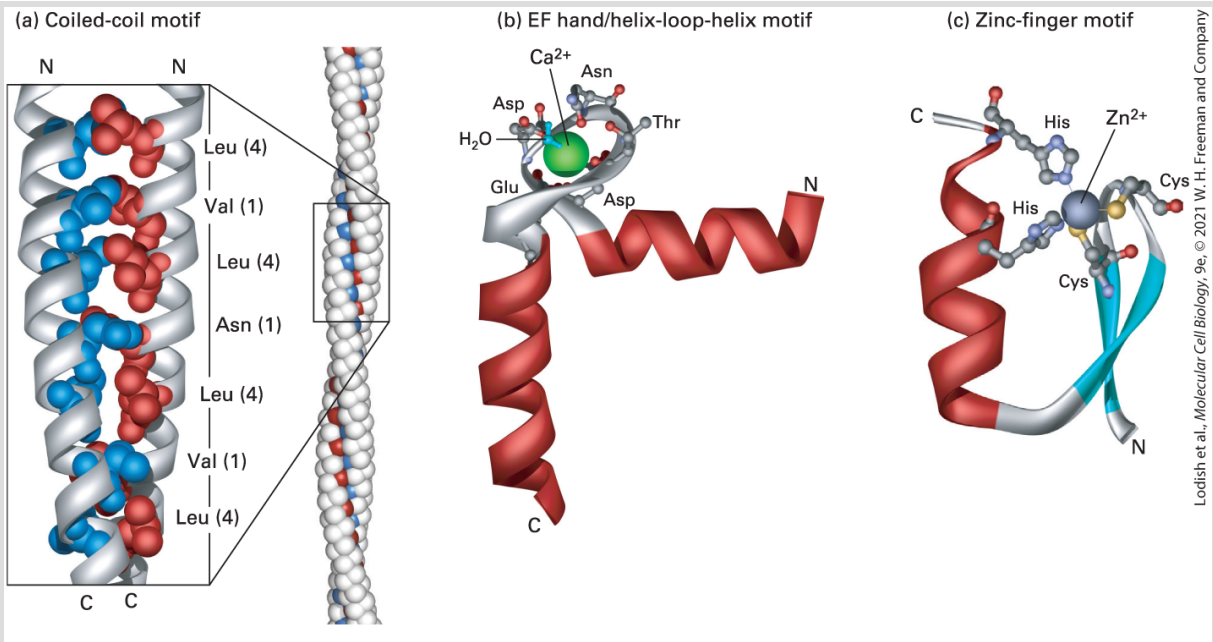


FIGURE 3-7 Motifs of protein secondary structure. (a) This parallel, two-stranded coiled-coil motif (*left*) is characterized by two α helices wound around each other. The backbone of the residues is represented as a ribbon, as illustrated in Figure 3-4, with atoms of only some of the side chains represented as balls. Helix packing is stabilized by interactions between hydrophobic side chains (red and blue) present at regular intervals along each strand and found along the seam of the intertwined helices. Each α helix exhibits a characteristic heptad repeat sequence with a hydrophobic residue often, but not always, at positions 1

and 4, as indicated. The coiled-coil nature of this structural motif is more apparent in long coiled coils containing many such sequence motifs (*right*). (b) An EF hand, a type of helix-loop-helix motif, consists of two helices connected by a short loop in a specific conformation. The backbone of the residues is represented as a ribbon with only some of the side chains represented as balls (atoms) and sticks (bonds). This structural motif is common to many proteins, including many calcium-binding and DNA-binding regulatory proteins. In calcium-binding proteins such as calmodulin, oxygen atoms from five residues in the acidic glutamate- and aspartate-rich loop and one water molecule form ionic bonds with a large Ca^{2+} ion (green). (c) The zinc-finger motif is present in many DNA-binding proteins that help regulate transcription. A Zn^{2+} ion is held between a pair of β strands (blue) and a single α helix (red) by a pair of cysteine residues and a pair of histidine residues. The two invariant cysteine residues are usually at positions 3 and 6, and the two invariant histidine residues are at positions 20 and 24 in this 25-residue motif.

[Part (a) Data from L. Gonzalez, Jr., D. N. Woolfson, and T. Alber, 1996, *Nat. Struct. Biol*. **3**:1011–1018, PDB IDs 1zik and 2tma. Part (b) Data from R. Chattopadhyaya et al., 1992, *J. Mol. Biol*. **228**:1177–1192, PDB ID 1cll. Part (c) Data from S. A. Wolfe, R. A. Grant, and C. O. Pabo, 2003, *Biochemistry* **42**:13401–13409, PDB ID 1llm.]

The relationship between the primary structure of a polypeptide chain and the structural motif into which a portion of it folds is not always straightforward. The amino acid sequences responsible for any given structural motif in different proteins may be very similar to one another. In other words, a common *sequence motif* can result in a common structural motif. This is the case for the heptad repeats that form coiled coils. However, it is also possible for seemingly unrelated amino acid sequences to fold into a common structural motif, so it is not always possible to predict which amino acid sequences will fold into a given structural motif. Conversely, it is possible that a commonly occurring sequence motif will not fold into a well-defined structural motif. Sometimes short sequence

motifs that have an unusual abundance of a particular amino acid, such as proline or aspartate or glutamate, are called *domains*; however, these and other short contiguous segments are more appropriately called *sequence motifs* than domains, as the latter term has a distinct meaning that we will define shortly.

We will encounter numerous additional motifs in our discussions of proteins in this and other chapters. The presence of the same structural motif in different proteins with similar functions clearly indicates that these useful combinations of secondary structures have been conserved in evolution.

Tertiary Structure Is the Overall Folding of a Polypeptide Chain

Tertiary structure refers to the overall conformation of a polypeptide chain — that is, the three-dimensional arrangement of all its amino acid residues. In contrast to secondary structures, which are stabilized only by hydrogen bonds, tertiary structure is stabilized primarily by hydrophobic interactions between nonpolar side chains, together with van der Waals interactions and hydrogen bonds involving both polar side chains and backbone amino and carboxyl groups. These stabilizing forces hold together elements of secondary structure — α helices, β strands, turns, and loops — and structural motifs. Because the interactions stabilizing tertiary

structures are often weaker than those stabilizing secondary structure, the tertiary structure of a protein is not necessarily rigidly fixed, but can undergo continual minute fluctuations. Some segments within the tertiary structure of a protein can be so mobile that they are considered to be disordered — that is, lacking a single, well-defined, stable, three-dimensional structure. This variation in structure has important consequences for the function and regulation of proteins.

The chemical properties of amino acid side chains help define tertiary structure. In some proteins — for example, those that are secreted from cells or are cell-surface proteins that face the extracellular environment — *disulfide bonds* form between the side chains of cysteine residues. Disulfide bonds can covalently link regions of the proteins, thus restricting the proteins' flexibility and increasing the stability of their tertiary structures. Amino acids with charged hydrophilic polar side chains tend to be on the outer surfaces of water-soluble proteins; by interacting with water, they increase the protein's solubility in water. Moreover, these polar side chains can form noncovalent interactions with other water-soluble molecules, including other proteins. In contrast, amino acids with hydrophobic nonpolar side chains are usually sequestered away from the water-facing surfaces of a protein, in many cases forming a water-insoluble central core. This observation led to what's known as the *oil drop model* of protein conformation because the core of a protein is relatively hydrophobic, or *oily* (Figure 3-8). Hydrophilic

polar side chains that are uncharged are found both on the surface and in the inner core of water-soluble proteins.

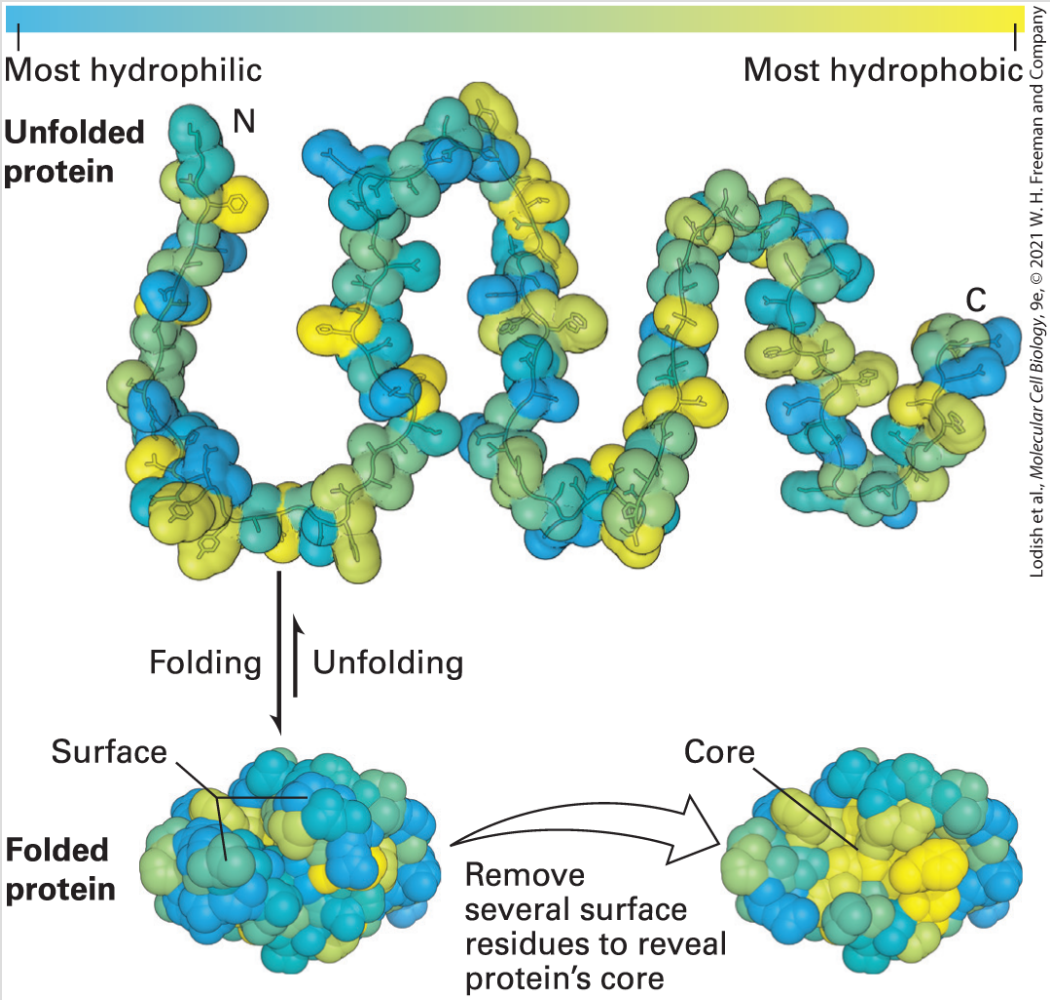


FIGURE 3-8 The oil drop model of protein folding. The hydrophobic and hydrophilic residues of a polypeptide chain can be distributed throughout its linear sequence as illustrated in the unfolded protein (*top*). The color scale denotes the most hydrophilic residues (blue) to the most hydrophobic (yellow). When the protein folds (*bottom left*), hydrophilic (charged and uncharged polar) side chains will often be exposed on the protein's surface, where they can form stabilizing interactions with surrounding water and ions. In contrast, the hydrophobic residues tend to cluster together in the inner core, somewhat like drops of oil in an aqueous liquid, driven away from the aqueous surroundings

by the hydrophobic effect (see Chapter 2). These core residues are more easily seen when several surface residues are removed (*bottom right*).

[Data from M. C. Vaney et al., 1996, *Acta Crystallogr., Sect. D*. **52**:505, PDB ID 193l.]

As you might expect, our ability to predict the tertiary structure for most proteins from primary sequence is not as advanced as our ability to predict secondary structure from sequence. One method to assess the quality of a tertiary structure prediction is to use the total score of the global distance test (a scale from 0 to 100), which compares the prediction with an experimentally determined structure. A score of <20 is a clearly incorrect prediction, scores of 80–90 are close to identifying the positions of individual atoms, and 100 is perfect. Current methods result in predictions around 57, which provide a good representation of the overall chain folding, but often have serious errors when compared to experimental results. At this time, predictions are better for shorter (length <150 residues) polypeptides.

Different Ways of Depicting the Conformation of Proteins Convey Different Types of Information

The simplest way to represent three-dimensional protein structure is to trace the course of the backbone atoms, sometimes only the C_α

atoms, with a solid line — or narrow tube — called a C_α *backbone trace* (Figure 3-9a); the most complex representation, called a *ball-and-stick model*, shows every atom (Figure 3-9b). The C_α backbone trace shows the overall folding of the polypeptide chain without consideration of the amino acid side chains; the ball-and-stick model (with balls representing atoms and sticks representing bonds) details the interactions between side-chain atoms, including those that stabilize the protein's conformation and interact with other molecules, as well as the atoms of the backbone. Even though both views are useful, the elements of secondary structure are not always easily discerned in them. Another type of representation, called a *ribbon diagram*, uses common shorthand symbols for depicting secondary structure — for example, coiled ribbons or solid cylinders for α helices, flat ribbons or arrows for β strands, and flexible thin strands for β turns, coils, and loops (Figure 3-9c). In a variation of the basic ribbon diagram, ball-and-stick or space-filling models of side chains are attached to the backbone ribbon, as we have seen in Figure 3-7. In this way, side chains of interest can be visualized clearly in the context of the secondary structure that is represented by the ribbons.

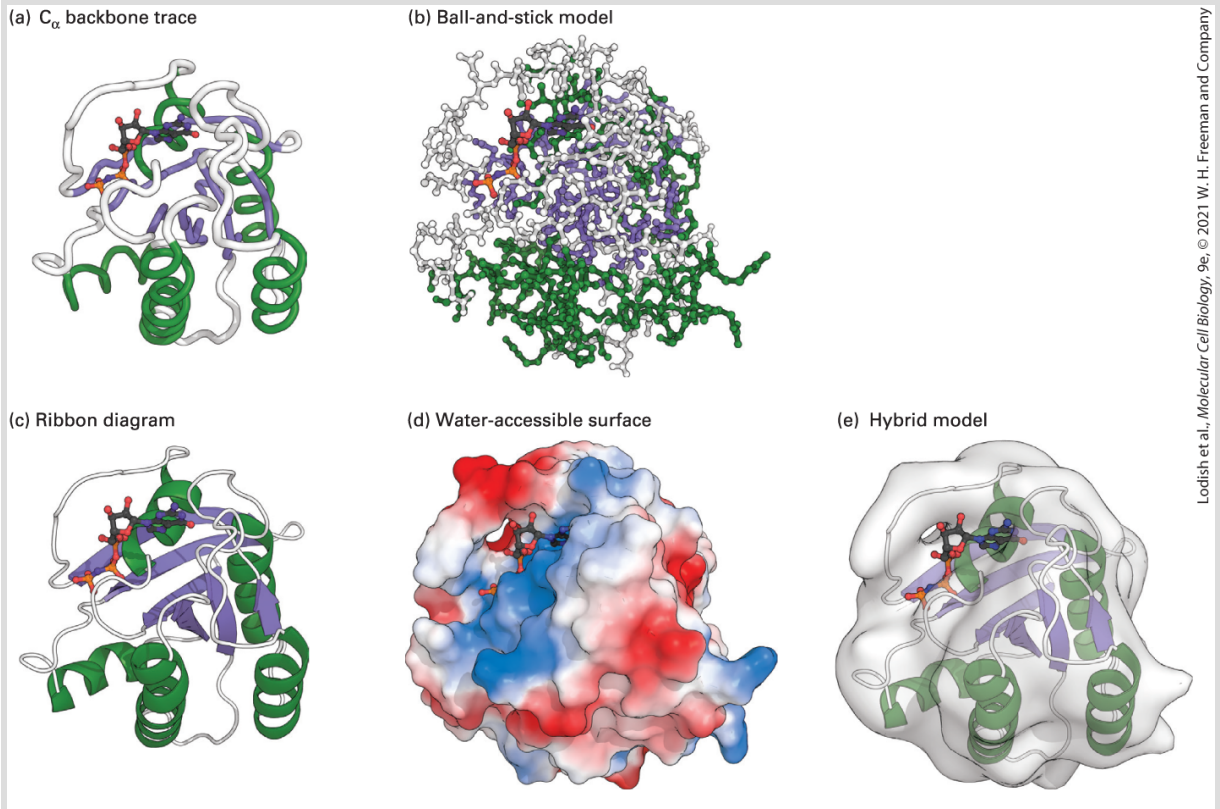


FIGURE 3-9 Five ways to visualize the protein Ras with its bound GDP. (a) The C_α backbone trace demonstrates how the polypeptide is tightly packed into a small volume. (b) A ball-and-stick representation reveals the locations of all atoms. (c) Turns and loops connect pairs of helices and strands. (d) A water-accessible surface reveals the numerous lumps, bumps, and crevices on the protein surface. Regions of positive charge are shaded blue, and regions of negative charge are shaded red. (e) Hybrid model combines ribbon and transparent surface models.

[Data from E. F. Pai et al., 1990, *EMBO J.* **9**:2351–2359, PDB ID 5p21.]

However, none of these three ways of representing protein structure conveys much information about the atoms that are on the protein's surface and in contact with the protein's environment. The surface is of interest because it is where other molecules usually bind to a protein. Thus a useful alternative way to represent proteins is to show only the surface and use colors to highlight regions having a

common chemical character, such as hydrophobicity or hydrophilicity, and charge characteristics, such as positive (basic) or negative (acidic) side chains (Figure 3-9d). Such models reveal the topography of the protein surface and the distribution of charge, both important features of binding sites, as well as clefts in the surface where other molecules may bind. This view represents a protein as it is "seen" by other molecules. Hybrid models (Figure 3-9e) combine features of multiple types of protein representation.

Domains Are Modules of Tertiary Structure

Distinct regions of protein structure are often referred to as **domains**. There are three main classes of protein domains: functional, structural, and topological. A *functional domain* is a region of a protein that exhibits a particular activity characteristic of that protein, usually even when isolated from the rest of the protein. For instance, a particular region of a protein may be responsible for a catalytic activity (e.g., a kinase domain that covalently adds a phosphate group to another molecule) or its binding ability (e.g., a DNA-binding domain or a membrane-binding domain). Functional domains are often identified experimentally by whittling down a protein to its smallest active fragment with the aid of **proteases**, enzymes that cleave one or more peptide bonds in a target polypeptide. Alternatively, the DNA encoding a protein can be

modified so that when the modified DNA is used to generate a protein, only a particular region, or domain, of the full-length protein is made. Thus it is possible to determine if specific parts of a protein are responsible for particular activities exhibited by the protein. Indeed, functional domains are often also associated with corresponding structural domains.

A *structural domain* is a region about 40 or more amino acids in length, arranged in a single, stable, and distinct structure often comprising one or more secondary structures or structural motifs. Many structural domains can fold into their characteristic structures independently of the rest of the protein in which they are embedded. As a consequence, distinct structural domains can be linked together — sometimes by stretches of polypeptide chain called *spacers* — to form a large multidomain protein (somewhat like beads on a string). Each of the polypeptide chains in the trimeric flu virus hemagglutinin, for example, contains a globular domain and a fibrous domain (Figure 3-10a).

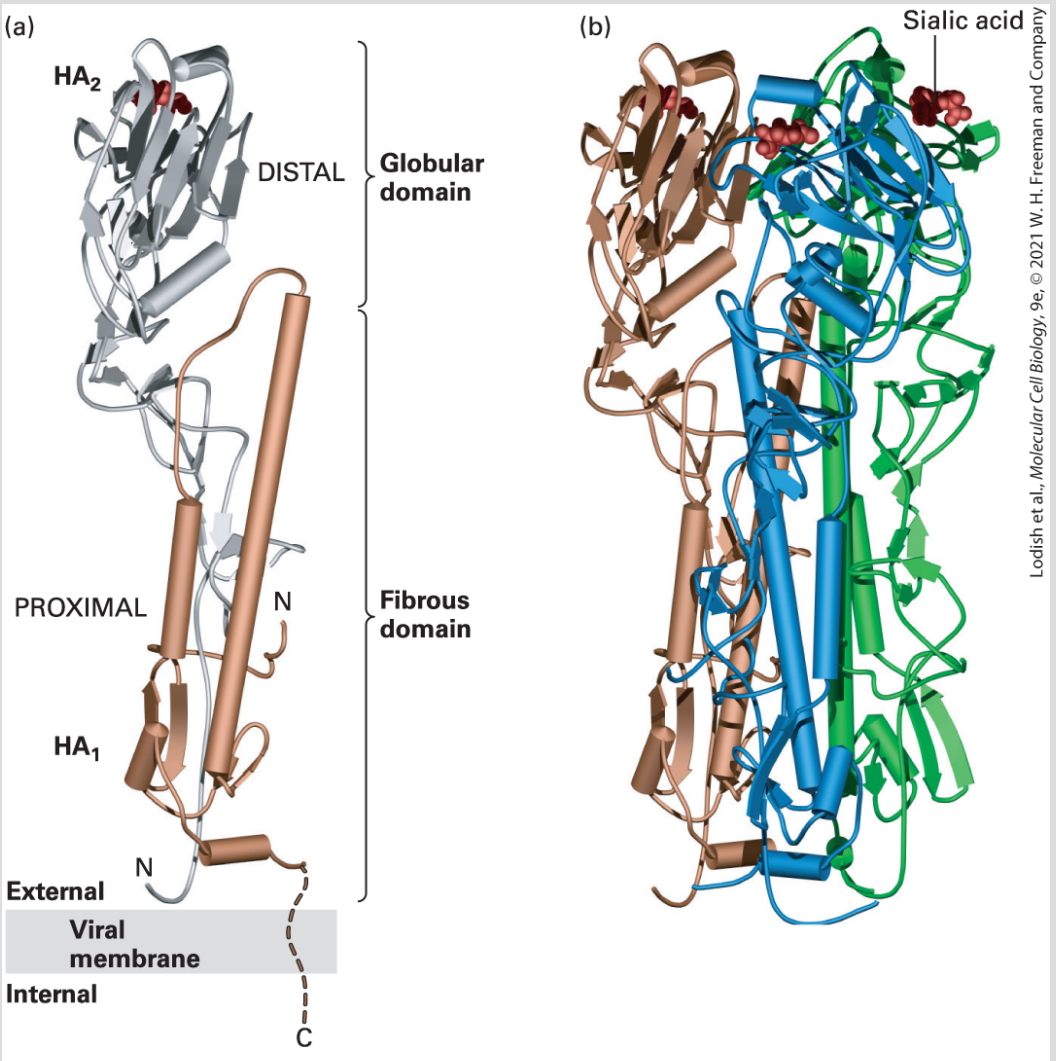


FIGURE 3-10 Tertiary and quaternary levels of structure. The protein pictured here, hemagglutinin (HA), is found on the surface of the influenza virus. This long molecule has three identical subunits, each composed of two polypeptide chains, HA_1 and HA_2. (a) The tertiary structure of each HA subunit comprises the folding of its helices and strands into a compact structure that is 13.5 nm long and divided into two domains. The membrane-distal domain (silver) is folded into a globular conformation. The membrane-proximal domain (gold) has a fibrous, stemlike conformation owing to the alignment of two long α helices (cylinders) of HA_2 with β strands in HA_1. Short turns and longer loops, many of them at the surface of the molecule, connect the helices and strands in each chain. (b) The quaternary structure of HA is stabilized by lateral interactions between the long helices (cylinders) in the fibrous domains of the three subunits (gold, blue, and green), forming a triple-stranded coiled-coil

stalk. Each of the distal globular domains in HA binds sialic acid (red) on the surface of target cells. Like many membrane proteins, HA contains several covalently linked carbohydrate chains (not shown).

[Data from S. J. Gamblin et al., 2004, *Science* **303**:1838–1842, PDB ID 1ruz.]

Structural domains frequently are also functional domains in that they can have an activity independent of the rest of the protein. Structural domains can be incorporated as modules into different proteins. The modular approach to protein architecture is particularly easy to recognize in large proteins, which tend to be mosaics of different domains that confer distinct activities and thus can perform different functions simultaneously. As many as 80 percent of the proteins in eukaryotes have multiple structural domains. The organization (order) of domains within a multidomain protein is called the *domain architecture*. A useful simile is that a protein is like a sentence. The amino acid residues are like letters, spaces, and punctuation marks, and the structural domains are like words, each having its distinct activity, and, when joined together into a *sentence*, they generate a protein with a distinctive function.

Thousands of structural domains have been identified in proteins. Some of these are not very common, whereas others are found in many different proteins. Indeed, by some estimates, only nine major types of structural domains account for as much as a third of all the structural domains in all proteins. Structural domains can be recognized in proteins whose structures have been determined by x-ray crystallography or nuclear magnetic resonance (NMR) analysis

or in images captured by electron microscopy (see Section 3.5). For example, the epidermal growth factor (EGF) domain is a structural domain that is present in several proteins (Figure 3-11). EGF is a small, soluble peptide hormone that binds to cells in the embryo and in skin and connective tissue in adults, causing them to divide. It is generated by proteolytic cleavage (breaking of a peptide bond) between repeated EGF domains in the EGF precursor protein, which is anchored in the plasma membrane by a membrane-spanning domain. EGF domains with sequences similar to, but not identical to, that of the EGF peptide hormone are present in other proteins and can be liberated by proteolysis. These proteins include tissue plasminogen activator (TPA), a protease that is used to dissolve blood clots in heart attack victims; Neu protein, which takes part in embryonic differentiation; and Notch protein, a receptor protein in the plasma membrane that functions in signaling and is important for development (see Chapter 16). Besides the EGF domain, these three proteins have other domains in common with other proteins. For example, TPA possesses a trypsin domain, a functional domain found in some proteases.

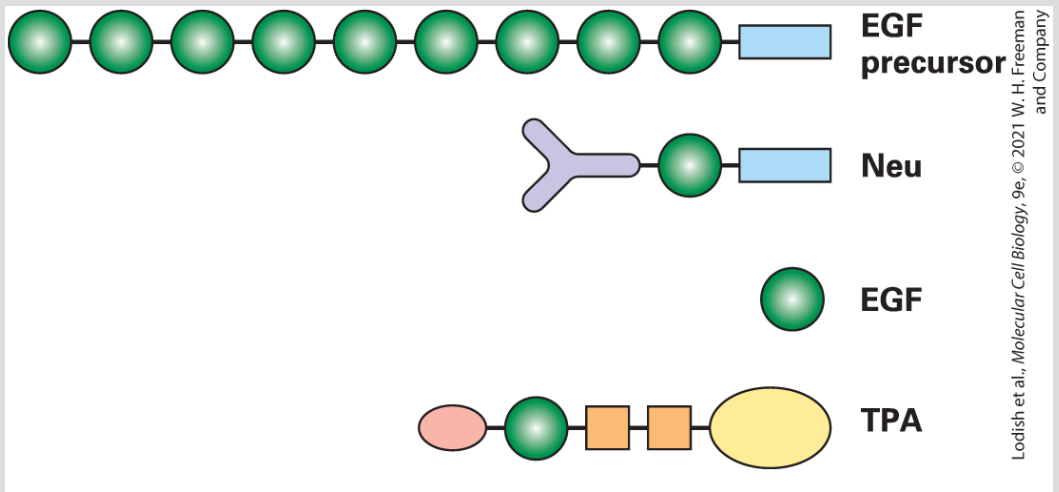


FIGURE 3-11 Modular nature of protein domains. Epidermal growth factor (EGF) is generated by proteolytic cleavage of a precursor protein containing multiple EGF domains (green) and a membrane-spanning domain (blue). An EGF domain is also present in the Neu protein and in tissue plasminogen activator (TPA). These proteins also contain other widely distributed domains, indicated by shape and color. See I. D. Campbell and P. Bork, 1993, *Curr. Opin. Struc. Biol.* **3**:385.

Regions of proteins that are defined by their distinctive spatial relationships to the rest of the protein are *topological domains*. For example, some proteins associated with cell-surface membranes have a part extending inward into the cytoplasm (cytoplasmic domain), a part embedded within the phospholipid bilayer (membrane-spanning domain), and a part extending outward into the extracellular space (extracellular domain). Each of these topological domains can comprise one or more structural and functional domains. It is also common for the region of a protein at the N-terminus or C-terminus to be called the *N-terminal* or *C-terminal domain*, respectively.

In Chapter 7, we will consider the mechanism by which the gene segments that correspond to domains became shuffled in the course of evolution, resulting in their appearance in many proteins. Once a functional, structural, or topological domain has been identified and characterized in one protein, it is possible to use that information to search for similar domains in other proteins and to suggest potentially similar functions for those domains in those proteins.

Comparing Protein Sequences and Structures Provides Insight into Protein Function and Evolution

Analyses of many diverse proteins have conclusively established a relation between the amino acid sequence, three-dimensional structure, and function of proteins. A good rule of thumb is that the greater the similarity of the sequences of two polypeptide chains, the more likely they are to have similar three-dimensional structures and similar functions. Although this comparative approach is very powerful, caution must always be exercised when attributing to one protein, or a part of a protein, a function or structure similar to that of another protein based only on amino acid sequence similarities. There are examples in which proteins with similar overall structures display different functions, as well as cases in which proteins unrelated in amino acid sequence or function nevertheless have very similar folded tertiary structures. In

many cases, however, such sequence comparisons provide important insights into protein structure and function.

Investigators have increased the use of protein sequence comparisons to study protein structure and function as the genomes and transcriptomes of more and more organisms are sequenced, because a vast array of protein sequences ($\sim$250 million) has been deduced from the nucleic acid sequences. Indeed, the molecular revolution in biology during the last decades of the twentieth century created a new scheme of biological classification based on similarities and differences in the amino acid sequences of proteins. Proteins that have a common ancestor are referred to as **homologs**. The main evidence for **homology** among proteins, and hence for their common ancestry, is similarity in their sequences, which is often reflected in similar structures. We can describe homologous proteins as belonging to a *family* and can trace their lineage — how closely or distantly they are related to one another in an evolutionary sense — from comparisons of their sequences. Generally more closely related proteins exhibit greater sequence similarity than more distantly related proteins because, over evolutionary time, alterations in the DNA sequences accumulate in the genes encoding these proteins. The folded three-dimensional structures of homologous proteins may be similar even if some parts of their primary structure show little evidence of sequence homology.

Based on their sequences, structural similarities, evolutionary histories and functions, related proteins have been classified as belonging to hierarchical sets of proteins called *superfamilies, groups, subgroups, families*, and *subfamilies*. The sequences of members of each subfamily are more similar to each other than they are to those of members of other subfamilies within the same family. Initially, proteins with relatively high sequence similarities (>50 percent exact amino acid matches, or *identities*) and related functions or structures were defined as an evolutionarily related family, while a superfamily encompassed two or more families in which the interfamily sequences matched less well (∼30–40 percent identities) than within one family. It is generally thought that proteins with about 30 percent sequence identity are likely to have similar three-dimensional structures; however, such high sequence identity is not required for proteins to share similar structures.

The kinship among homologous proteins is most easily visualized by a tree diagram (in essence, a family tree) based on sequence analyses. For example, the amino acid sequences of globins — the proteins hemoglobin and myoglobin and their relatives from bacteria, plants, and animals — suggest that they evolved from an ancestral monomeric oxygen-binding protein (Figure 3-12a). With the passage of time, the gene for this ancestral protein slowly changed due to random changes in the DNA sequence, initially diverging into lineages leading to animal and plant globins. Subsequent changes gave rise to myoglobin and, after a gene

duplication, to the α and β subunits of the tetrameric hemoglobin molecule ($\alpha_2\beta_2$) of the vertebrate circulatory system.

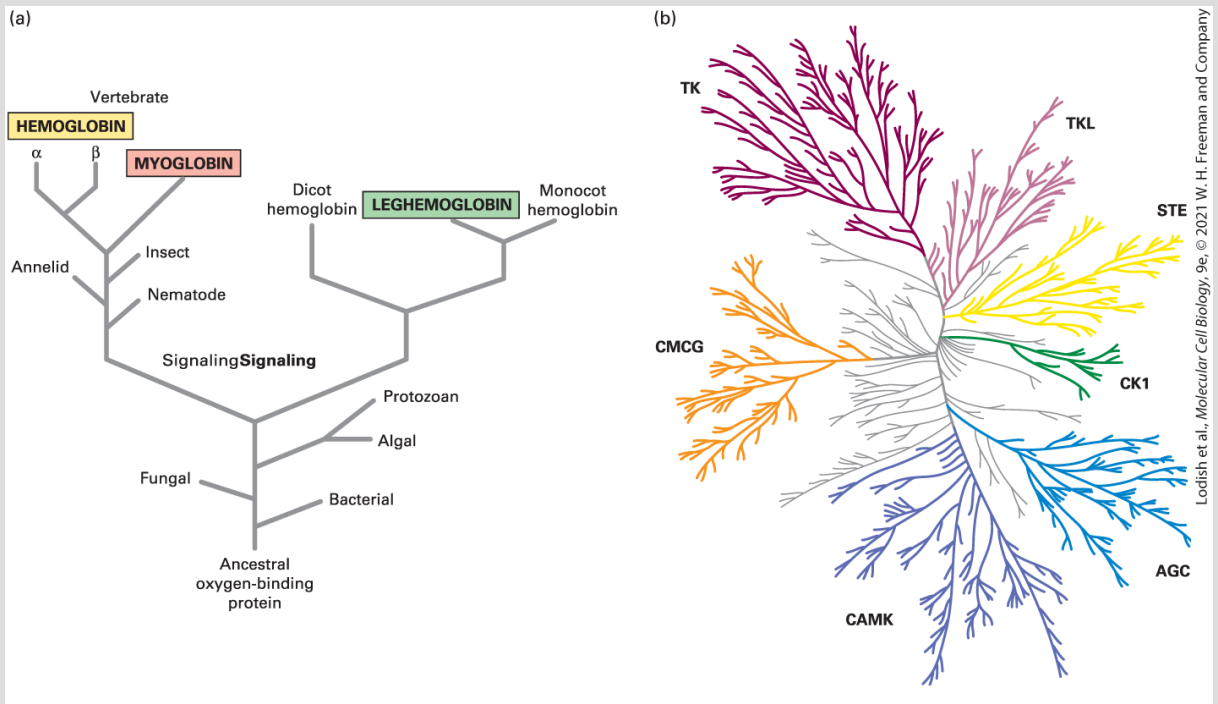


FIGURE 3-12 Evolution of protein families. (a) A primitive monomeric oxygen-binding globin is thought to be the ancestor of modern-day blood hemoglobins, muscle myoglobins, and plant leghemoglobins. Sequence comparisons have revealed that the evolution of the globin proteins parallels the evolution of animals and plants. Major changes occurred with the divergence of plant globins from animal globins and of myoglobin from hemoglobin. Later, gene duplication gave rise to the α and β subunits of hemoglobin. See R. C. Hardison, 1996, *Proc. Nat'l. Acad. Sci. USA* **93**:5675. (b) The family tree of 478 human protein kinases is divided into seven major groups of protein kinase families (comprising 405 kinases), designated: CMGC, CAMK, AGC, CK1, STE, TKL, and TK. Each of these contains families and subfamilies. We will describe the critical roles of kinases from many of these groups in subsequent chapters. For example, the members of the TK (tyrosine kinase) group are distinctive because they phosphorylate the side chains of tyrosine residues. One TK member is the receptor for the hormone insulin (Chapter 16).

[Part (b) Republished with permission of American Association for the Advancement of Science, from G. Manning et al., 2002, "The Protein Kinase Complement of the Human

Genome," *Science* **298**:1912–1934; permission conveyed through Copyright Clearance Center, Inc.]

A series of duplications and divergences can lead to the generation of a large, highly branched family of genes and corresponding proteins within a single organism. Figure 3-12b shows such a family tree for the 478 human protein kinase enzymes that evolved from one ancient precursor. All of these kinases catalyze the covalent addition of a phosphate group from ATP to the hydroxyl group of the side chains of serine, threonine, or tyrosine (∼15 percent phosphorylate tyrosines) (see Figure 2-15). A small number phosphorylate both serine (or threonine) and tyrosine are called *dual-specificity* kinases. Forty additional human kinases evolved from other precursors, and thus there is a total of 518 kinases in what is called the *human kinome*. In such a large family tree, the sum of the lengths of the lines between one member and another is generally proportional to the evolutionary distance between the two proteins — the shorter the distance on the tree between any two members of the family, the more similar they are in sequence and typically in structure and function. Often analysis of the branching patterns of proteins in such family trees is used to subdivide the members into groups, families, and subfamilies (Figure 3-12b) rather than relying on a precise formula using an arbitrary percent identity of sequence to define subfamilies, families, and so on.

There Are Four Broad Structural Categories of Proteins

Proteins usually fall into one of four broad structural categories based on their tertiary structure: *globular proteins, fibrous proteins, integral membrane proteins*, and *intrinsically disordered proteins*. These four broad categories of proteins are not mutually exclusive — some proteins are made up of combinations of segments that fall into two or more of these categories. *Globular proteins* are generally water-soluble, compactly folded structures, often spheroidal, that comprise a mixture of secondary structures (e.g., see Figure 3-9). *Fibrous proteins* are large, elongated, often stiff molecules. Some fibrous proteins are composed of a long polypeptide chain comprising many tandem copies of a short amino acid sequence motif that forms a single repeating secondary structure. These fibrous proteins are often made up of helical polypeptide chains, including α helices, triple helices (see the structure of collagen, the most abundant protein in mammals, in Figure 20-26) and helical coil coils with two or more strands (see Figure 3-7). Other fibrous proteins are composed of repeating globular protein subunits, such as the helical array of G-actin protein monomers that forms F-actin microfilaments (see Chapter 17). Fibrous proteins, which can aggregate into large multiprotein fibers that do not readily dissolve in water, usually play a structural role or participate in cellular movements. *Integral membrane proteins* are embedded within the phospholipid bilayer of the membranes that enclose cells and

organelles. Their membrane-spanning domains often comprise one or more ~20-residue long α helices and occasionally β barrels; they are discussed in detail in Chapter 10.

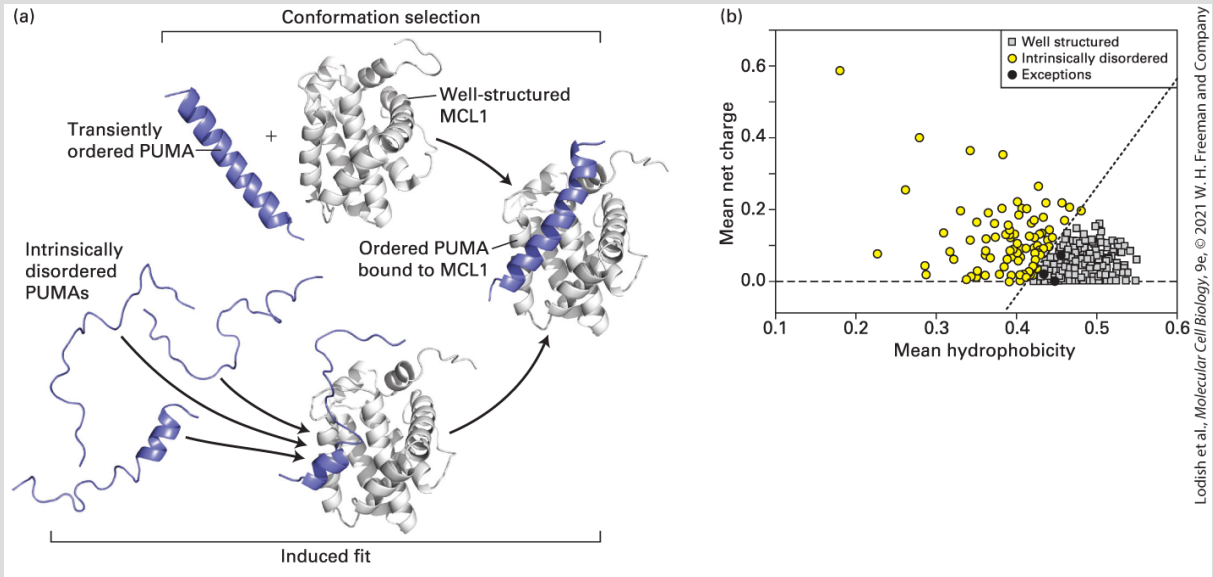
Regardless of a protein's category of tertiary structure, one or more segments of the protein may be *disordered* meaning that they do not form thermodynamically stable structures. Disordered protein segments are exceptionally flexible in conformation, and their flexibility appears to be key to their functional activities. Regions of disorder can provide flexible links, or tethers, between well-ordered regions of a protein; serve as binding sites for other proteins (Figure 3-13a); serve as sites of some types of post-translational protein modification [e.g., covalent addition of phosphate groups (phosphorylation) or sugars (glycosylation)]; serve as targets of protease digestion that regulates protein activity; inhibit the activity of the protein in which they are embedded (autoinhibition sites); or serve as signals for intracellular sorting of proteins (see Chapter 13). The activities of many proteins containing disordered segments are described in subsequent chapters. For example, phosphorylation of the disordered C-terminal domain (CTD) of RNA polymerase II (see Figure 8-12), which is composed of multiple repeats of a seven–amino acid sequence containing proline, threonine, and serine, regulates key steps in the synthesis of mRNA (see Chapters 8 and 9). The N-termini of histone proteins that control DNA organization in chromatin (see Chapter 7) are sites of important post-translational modifications, and the disordered, proline-rich FH1 region in the

protein formin controls the assembly of actin filaments (see Chapter 17).

Many proteins we consider in this book adopt only one or a few very closely related conformations when they are in their normal functional state, called the *native state*. However, in some proteins, the *entire* polypeptide chain is disordered. These proteins do not have well-ordered structures in their native, functional states; instead, their polypeptide chains are very flexible with no fixed conformation. These proteins are called *intrinsically disordered proteins (IDP)*. Intrinsically disordered proteins typically serve as signaling molecules, regulators of the activities of other molecules, or as scaffolds for multiple proteins, small molecules, and ions (e.g., binding ions via multiple charged residues). Disordered segments may also be present in otherwise well-structured proteins; these disordered segments are called **intrinsically disordered regions (IDR)**.

Intrinsically disordered proteins and disordered segments can be identified experimentally using various biochemical techniques, such as tests of sensitivity to protease digestion (disordered regions usually exhibit greater protease sensitivity), and a wide variety of biophysical techniques, including spectroscopy. The disorder of these protein segments or of an entire protein apparently arises as a consequence of their having a sequence that, relative to well-ordered proteins, is richer in polar amino acids, proline, and net

charge, and poorer in hydrophobic residues (Figure 3-13b). Algorithms primarily based on calculations of amino acid composition — particularly net charge and hydrophobicity — are used to predict which proteins or segments of proteins are disordered. By some estimates, about 30 percent or more of eukaryotic proteins have at least one segment of 50 or more consecutive residues that is disordered. In some cases, an intrinsically disordered protein (or region) can transition into a highly ordered structure (see Figure 3-13a).



EXPERIMENTAL FIGURE 3-13 Intrinsically disordered proteins: mechanisms of binding to well-ordered proteins and identification based on hydrophobicity and net charge. (a) The binding of an intrinsically disordered protein (PUMA, blue) to a well-ordered protein (MCL1, gray) results in the formation of a well-defined structure in the previously disordered protein. Two mechanisms have been proposed for generating a bound complex in which both proteins are structured: conformational selection (*top pathway*) and induced fit (*bottom pathway*). In conformational selection, the disordered protein (PUMA) occasionally and transiently adopts in solution the structure it would have in the bound state. The well-ordered binding partner (MLC1) can then bind to (select) PUMA in that transient, ordered conformation, forming a relatively stable bound complex. In induced fit, the disordered

protein begins to bind to the well-ordered partner while still disordered and then, while bound, is induced to form the ordered conformation present in the relatively stable, heterodimeric complex. Experiments suggest that the induced fit mechanism best describes the binding of PUMA and MCL1. (b) The sequences of 275 well-ordered, monomeric globular proteins (gray squares) and 91 intrinsically disordered proteins (black and yellow circles) were used to calculate the mean hydrophobicity per residue in each protein using a scale of 0 (least hydrophobic) to 1 (most hydrophobic, *x* axis), and the mean net charge per residue at pH 7.0 (*y* axis). With only three exceptions (black circles), the proteins define two distinct distributions: low hydrophobicity, high net charge (intrinsically disordered, yellow circles) and high hydrophobicity, low net charge (well-ordered, gray squares). The three disordered proteins (black circles) that overlap with the well-ordered population each contain substantial segments predicted to be disordered (low hydrophobicity, high net charge) that apparently overwhelm the rest of the proteins' sequences that might otherwise result in a well-ordered conformation.

[Part (a) From J. M. Rogers, A. Steward, and J. Clarke, 2013, "Folding and Binding of an Intrinsically Disordered Protein: Fast, but Not 'Diffusion-Limited,' " *J. Am. Chem. Soc.* **135**(4):1415–1422; https://doi.org/10.1021/ja309527h. Part (b) Data from V. N. Uversky, J. R. Gillespie, and A. L. Fink, 2000, *Proteins* **41**:415–427.]

Multiple Polypeptides Assemble into Quaternary Structures, Supramolecular Complexes, and Biomolecular Condensates

Multimeric proteins consist of two or more polypeptide chains that in this context are referred to as *subunits*. A fourth level of structural organization, **quaternary structure**, describes the number (stoichiometry), relative positions, and often the tertiary structures

of the subunits in multimeric proteins. Multimeric proteins are composed of various numbers of identical (homomeric) or different (heteromeric) subunits, typically held together by noncovalent bonds. Flu virus hemagglutinin, for example, is a trimer of three identical subunits (a homotrimer) (see Figure 3-10). In many cases, the individual monomer subunits of a multimeric protein cannot function normally unless they are assembled into the multimeric protein. In other cases, assembly into a multimeric protein permits proteins that act sequentially in a pathway to increase their efficiency of operation owing to their juxtaposition in space, a phenomenon referred to as *metabolic coupling*. Classic examples of metabolic coupling are the fatty acid synthases of fungi, which are enzymes that synthesize fatty acids, and the polyketide synthases, which are large multiprotein complexes in bacteria that synthesize a diverse set of molecules called *polyketides* often used as drugs, including the antibiotic erythromycin.

Supramolecular Complexes

The highest level in the hierarchy of protein structure is the association of proteins into supramolecular complexes. Typically, such structures are very large, containing tens to hundreds of polypeptide chains and sometimes other biopolymers such as nucleic acids. In some cases, supermolecular structures exceed 1 megadalton (MDa) in mass, approaching 30–300 nm in size. The capsid that encases the nucleic acids of a viral genome is an

example of a supramolecular complex with a structural function. The bundles of cytoskeletal filaments that support and give shape to the plasma membrane are another example.

Other supramolecular complexes act as molecular machines. These complexes carry out complex cellular processes by integrating multiple proteins, each with distinct functions, into one large assembly. For example, a transcriptional machine is responsible for synthesizing messenger RNA (mRNA) using a DNA template. The operational details of this transcriptional machine are discussed in Chapters 5 and 8. It consists of RNA polymerase, itself a multimeric protein, and at least 50 additional components, including general transcription factors, promoter-binding proteins, helicase, and other protein complexes (Figure 3-14). Ribosomes, also discussed in Chapter 5, are complex machines that synthesize proteins and are formed of multiple proteins and nucleic acids. One of the most complex multiprotein assemblies is the nuclear pore, a structure that allows macromolecules to pass between the nucleus and the cytoplasm (see Chapter 13). It is composed of multiple copies of about 30 distinct proteins and forms an assembly with an estimated mass of 50 MDa. The fatty acid synthases and polyketide synthases referred to above are also molecular machines, as is the mitochondrial *supercomplex* $I/III_2/IV$ that integrates multiple steps in the electron-transport chain that provides energy for synthesizing ATP. This supercomplex (Figure 12-23) comprises three

supramolecular complexes and more than 64 polypeptide chains (see Chapter 12).

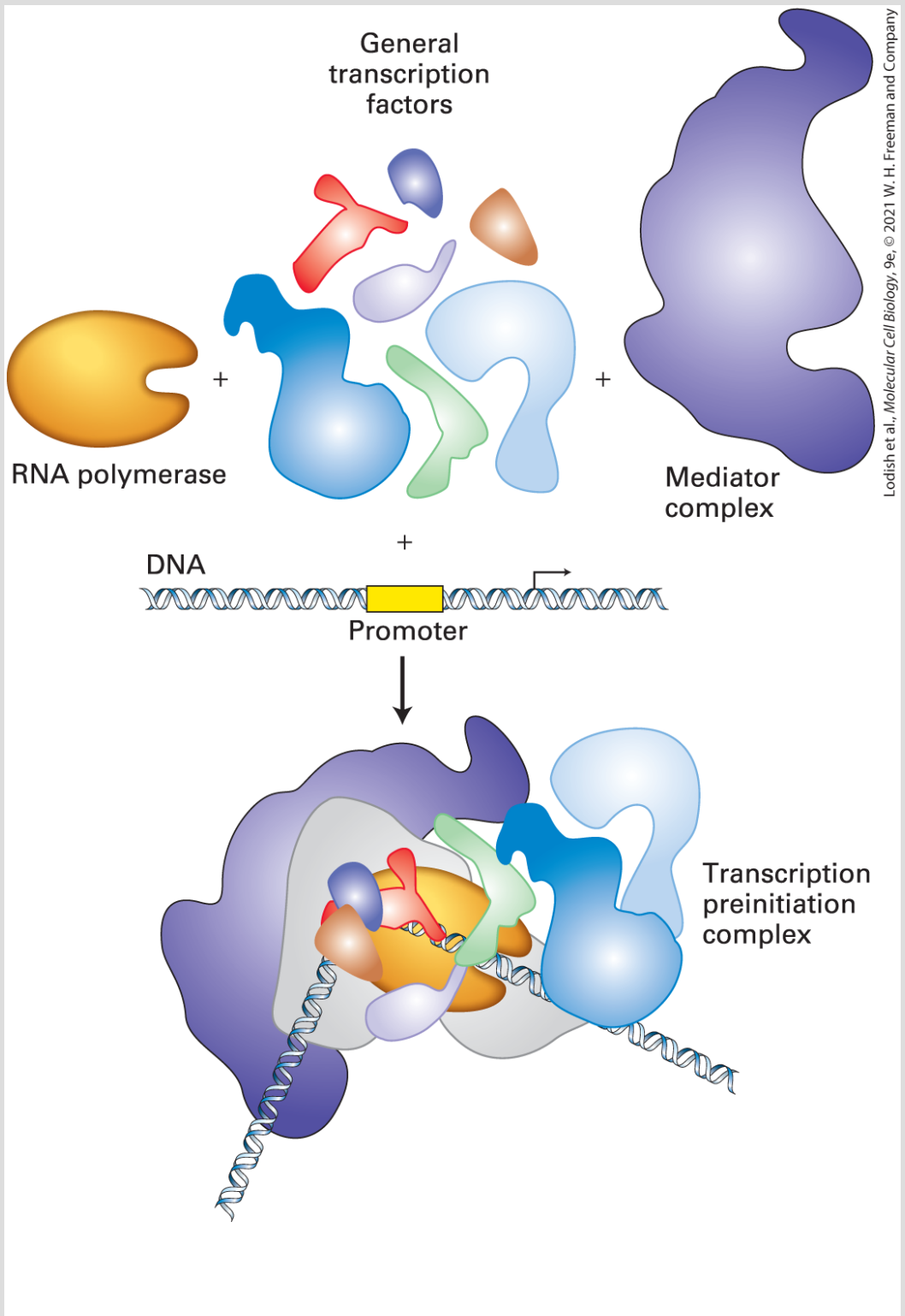
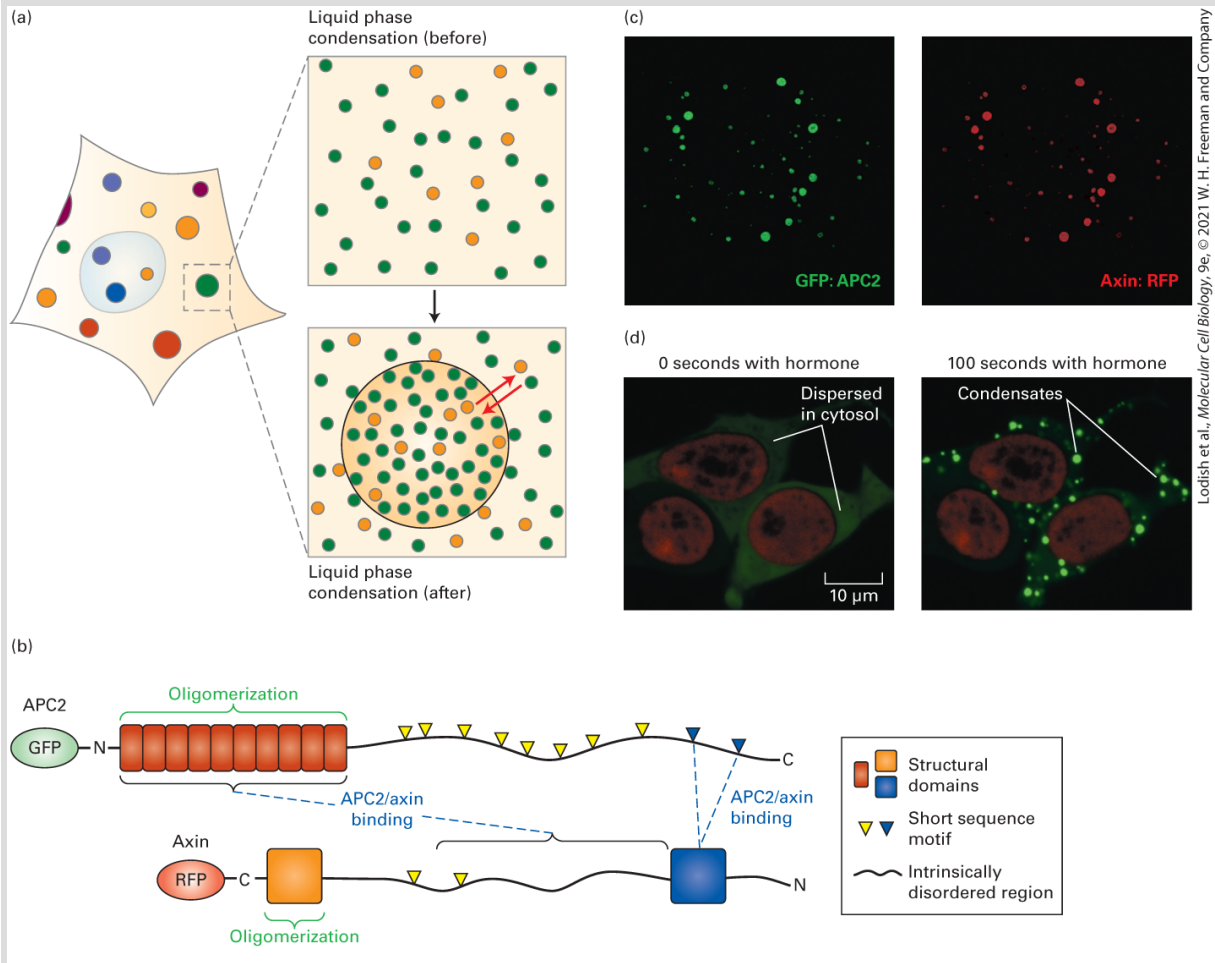


FIGURE 3-14 A molecular machine: the transcription initiation complex. The core RNA polymerase, general transcription factors, a mediator complex containing about 20 subunits, and other protein complexes not depicted here assemble at a promoter in DNA prior to the synthesis of the corresponding RNA. The polymerase carries out transcription of DNA; the associated proteins are required for initial binding of the polymerase to a specific promoter. Note that the DNA is bent in the transcription preinitiation complex. The multiple components function together as a molecular machine.

Biomolecular Condensates

A rather different approach to organizing together large collections of macromolecules is the generation of **biomolecular condensates** that are likely to play important roles in multiple aspects of cell biology. Biomolecular condensates are membraneless compartments in cells, often compared to liquid droplets, that are chemically and physically distinct from their surroundings. In biomolecular condensates, multiple copies of one or more macromolecules separate from the bulk surrounding fluid in a process called *liquid-liquid phase separation*. Through such phase separation, macromolecules, often proteins or proteins combined with RNAs, assemble or condense into a region of higher concentration than in the bulk surrounding fluid of the cytosol or nucleoplasm (Figure 3-15). Unlike the supramolecular complexes described above, biomolecular condensates can vary in size (100s to 1000s of nanometers), and their components generally do not have a fixed stoichiometry or quaternary structural arrangement.



EXPERIMENTAL FIGURE 3-15 Multivalent proteins with intrinsically disordered regions and other characteristics can assemble into biomolecular condensates. (a) Cells contain a variety of membraneless condensates, resembling liquid droplets, in the cytoplasm and nucleus (colored circles, *left*). Macromolecules (small circles in expanded rectangle before condensation, *upper right*) separate from the bulk surrounding fluid and concentrate in a biomolecular condensate due to liquid-liquid phase separation (large green circle after condensation, *lower right*). These macromolecules can diffuse into and out of the condensate (red arrows). (b) Models of the fruit fly (*Drosophila melanogaster*) proteins APC2 and Axin, two large proteins that participate in the Wnt signaling pathway (see Chapter 16). Many other species, including mammals, express APC2 and Axin homologues. These multivalent proteins contain domains (rectangles and squares), intrinsically disordered regions (black lines), and short sequence motifs (triangles) that perform key functions. The motifs designated by yellow triangles are involved with binding additional Wnt-associated signaling molecules (Chapter 16). When APC2 and Axin are synthesized in cultured cells with covalently attached fluorescent proteins (ovals), they can be visualized using confocal

fluorescence microscopy (Chapter 4). Each protein can oligomerize into larger complexes via the domains indicated by green brackets. (c) When expressed together, APC2 and Axin bind to each other [sites of binding indicated by blue dotted lines in panel (b)] forming biomolecular condensates that contain both proteins (both colors found in all condensates of various sizes, green GFP:APC2 on the *left*, red Axin:RFP on the *right*). (d) Two multivalent, artificial proteins, one of which is fluorescent (green) and can be easily detected by microscopy, were expressed together in human cells. These two proteins, which reside in the cytoplasm, cannot bind to one another unless the fluorescent protein is phosphorylated by the enzyme Protein Kinase A (PKA, see Section 3.4 and Chapter 15). In the absence of the hormone isoprenaline (*left panel*, 0 seconds with hormone), PKA is inactive, the two proteins do not bind to one another, and the green fluorescent protein is dispersed throughout the cytosol (*left panel*). In this condition, the proteins do not enter the nucleus (red ovals in the centers of the cells, the DNA was stained red by a fluorescently labeled DNA-binding protein called *histone 2B*, see Chapter 7). Addition of isoprenaline for 100 seconds (*right panel*) activates PKA, resulting in phosphorylation of the fluorescent protein, binding of the two proteins and formation in the cytosol of biomolecular condensates in which the proteins are concentrated (green dots that are brighter than the dispersed green in the *left panel*).

[Panel (a) Data from Y. Shin and C. P. Brangwynne, 2017, *Science* **357**(6357):eaaf4382. Panel (b) Data from Figure 1, M. I. Pronobis et al., 2017, *Mol. Biol. Cell* **28**:41–53; and Figure 3, K. N. Schaefer and M. Peifer, 2019, *Dev. Cell* **48**(4):429–444. Panel (c) Photos courtesy of Xiaokun Shu, from Q. Zhang et al., 2018, *Mol. Cell. (left)*; D. Yu et al., 2015, *Nat. Methods (right)*. Panel (d) Republished with permission from Elsevier, Inc., from K. N. Shaefer and M. Peifer, 2019, "Wnt/Beta-Catenin Signaling Regulation and a Role for Biomolecular Condensates," *Dev. Cell.* **48**(4):429–444; permission conveyed through Copyright Clearance Center, Inc.]

A biomolecular condensate can be thought of as an alternative to membrane-bound organelles (such as the ER, Golgi apparatus, lysosome, or a mitochondrion; see Chapter 1) for concentrating molecules in a defined region of the cell to facilitate integrated regulation of various cellular activities in time and space. We shall see in Chapter 14 that membrane-bound compartments, such as small transport vesicles, can fuse together to mix their contents and

that such vesicles can bud off, or break away from, larger structures (e.g., membrane vesicles break away from the ER and subsequently fuse with the Golgi apparatus). Biomolecular condensates can also break apart into smaller liquid droplets (disperse) or fuse (coalesce) into larger ones, just like drops of oil in water, such as in salad dressing. They also have the potential to undergo phase transitions from disordered liquid to a more highly ordered gel or even a solid-like state.

The capacity of a protein to form a condensate depends on its structure and concentration and on the conditions in its surrounding fluid, like pH, ionic strength, temperature, and the presence of other molecules. Within a cell, a set of such conditions is sometimes called the *physiologic state* of the cell. Thus alterations in conditions inside the cell appear to be able to regulate the formation of condensates. One characteristic of macromolecules that form biomolecular condensates is that they are multivalent — they have multiple sites that are able to bind to other molecules. In particular, they contain multiple domains that have the ability to bind to regions of other proteins or nucleic acids. When a protein binds to other copies of the same protein, it is said to oligomerize and the binding sites are called *oligomerization sites*. Another key characteristic of many proteins that form biomolecular condensates is that they have intrinsically disordered regions. The amino acid sequence of an IDR can influence the probability of a protein participating in the formation of a condensate and the nature of the

condensate formed. Currently, however, it can't be assumed that a protein having multiple domains or intrinsically disordered segment(s) will definitely incorporate into a condensate in vivo, but these features appear to play key roles for those proteins that do.

It is possible to readily detect condensates in cells using fluorescence microscopy (see Chapter 4) when one or more of the components contains a covalently attached fluorescent molecule. For example, in the Wnt signaling pathway that is described in detail in Chapter 16, several proteins assemble into biomolecular condensates, including two called *APC2* and *Axin* that are multivalent and have multiple protein-binding domains and IDRs (see Figure 3-15b). In the experiment shown in Figure 3-15c, the fluorescently labeled versions of the *Drosophila melanogaster* APC2 (fluoresces green) and Axin (fluoresces red) were expressed together in mammalian cells and assembled into nearly spherical condensates that contain both fluorescent molecules.

The types of proteins (and in some cases RNAs) that assemble to form a condensate control the properties of the condensate, as can the external conditions (pH, etc.). For example, Figure 3-15d shows cells that express two artificial proteins designed to form condensates. One of these is fluorescently labeled. In untreated cells the two cannot assemble into a condensate, so the fluorescent protein is distributed evenly throughout the cytosol (*left panel*, 0 sec, relatively low green fluorescence intensity). After stimulation with a

hormone for 100 seconds (*right panel*), the fluorescent protein is chemically modified (phosphorylated by an enzyme called a *kinase*; described in Section 3.4). As a consequence, it can then assemble with the other protein, which has a phosphoprotein-binding domain, into easily detected condensates (*right*, punctate green droplets that are especially bright because the concentration of the fluorescent molecules in the condensates is greater than when they are distributed through the cytosol). The condensates fuse into larger droplets as the incubation proceeds. The presence in a cell of spherical structures that (1) can fuse and break apart dynamically and (2) have components that can rapidly diffuse within the sphere and diffuse between the sphere and the surrounding liquid provides evidence that the labeled molecule participates in condensate formation. Examples of biomolecular condensates include nucleoli (sites of ribosomal subunit synthesis; Chapter 9), P-bodies (sites of translational repression and mRNA degradation; Chapter 9), and signaling complexes such as in the Wnt pathway (Chapter 16).

KEY CONCEPTS OF SECTION 3.1

Hierarchical Structure of Proteins

- Proteins are linear polymers of amino acids linked together by peptide bonds. A protein can have a single polypeptide chain or multiple polypeptide chains. There are four levels of organization of protein structure: primary, secondary, tertiary, and quaternary. The primary structure of a polypeptide chain is the sequence of covalently linked amino acids that compose the chain. Various mostly noncovalent interactions between amino acids in the linear sequence stabilize a protein's specific folded three-dimensional structure or conformation.

- Secondary structures are stabilized by hydrogen bonds between atoms of the peptide backbone. The α helix, β strand and sheet, and β turn are the most prevalent elements of protein secondary structure (see Figures 3-4, 3-5, and 3-6). Certain combinations of secondary structures give rise to structural motifs, which are found in a variety of proteins and are often associated with specific functions (see Figure 3-7).
- Protein tertiary structure results from hydrophobic interactions between nonpolar side groups and from hydrogen bonds and ionic interactions involving polar side groups and the polypeptide backbone. These interactions stabilize the folding of a polypeptide chain, including its secondary structural elements and structural motifs, into an overall three-dimensional arrangement.
- Proteins often contain distinct domains, independently folded regions with characteristic structural, functional, or topological properties.
- The incorporation of domains as modules in different proteins in the course of evolution has generated diversity in protein structure and function.
- Proteins with similar amino acid sequences generally can be assumed to have similar three-dimensional structures and similar functions. There are also examples of polypeptide chains with dissimilar sequences folding into similar three-dimensional structures.
- Homologous proteins are proteins that evolved from a common ancestor and thus have similar sequences, structures, and functions. They can be classified into families and superfamilies.
- Entire proteins or segments of proteins usually fall into one of four broad structural categories: globular proteins, fibrous proteins, integral membrane proteins, and intrinsically disordered proteins.
- The exceptional conformational flexibilities of intrinsically disordered proteins contribute to their functions as binding partners, signaling molecules, regulators of other molecules, scaffolds, flexible links between well-ordered regions of a protein, sites of post-translational protein modification, autoinhibitors, and signals for intracellular protein sorting.
- The quaternary structure of a protein is defined by the number (e.g., dimer or trimer) and organization of individual polypeptide subunits assembled into a multimeric protein.
- Cells contain large supramolecular complexes, sometimes called *molecular machines*, in which dozens of polypeptide chains assemble to perform complex cellular processes (e.g., DNA, RNA, and protein synthesis; photosynthesis; ATP generation, signal transduction).
- Biomolecular condensates are large, dynamic assemblies of molecules that form membraneless compartments. These liquid droplet-like condensates are chemically and physically distinct, or phase-separated, from the bulk surrounding fluid (e.g.,

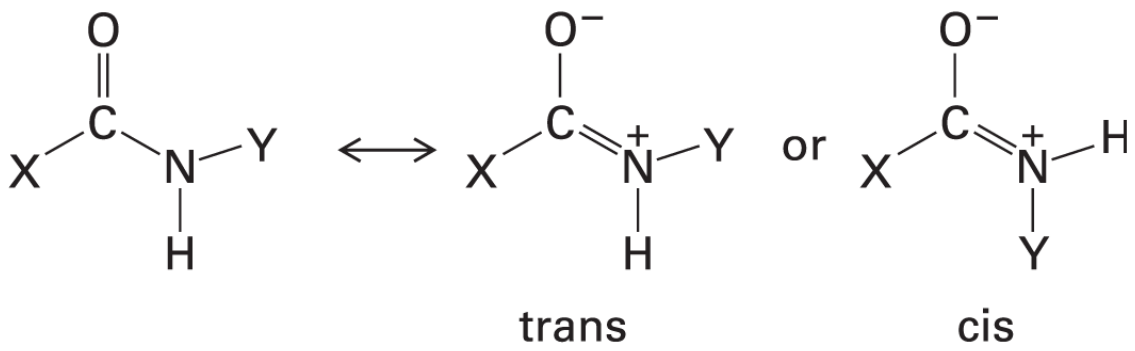
cytosol or nucleoplasm). The capacity of proteins to form condensates depends on their structures and the surrounding environment (pH, temperature, etc.). Condensate-forming proteins often contain multiple binding domains (they are multivalent) and intrinsically disordered regions.

3.2 Protein Folding

As noted previously, when it comes to the architecture of proteins, "form follows function." Thus for a polypeptide to fulfill its biological role within or outside cells, it is essential that it be synthesized with the proper amino acid sequence, and that it fold into the proper three-dimensional conformation, with the appropriate secondary, tertiary, and possibly quaternary structure. How is a protein with a proper sequence generated? A polypeptide chain is synthesized by a complex process called *translation*, which occurs in the cytoplasm on a large protein–nucleic acid complex called a *ribosome*. During translation, a sequence of *messenger RNA* (*mRNA*) serves as a template for the assembly of a corresponding amino acid sequence. The mRNA is initially generated by a process called *transcription*, whereby a nucleotide sequence in DNA is converted by transcriptional machinery in the nucleus into a sequence of mRNA. The intricacies of transcription and translation are considered in Chapter 5. Here we describe the key factors that ensure that a newly formed or forming (nascent) polypeptide chain is properly folded as it emerges from the ribosome.

Planar Peptide Bonds Limit the Shapes into Which Proteins Can Fold

A critical structural feature of polypeptides that limits how the chain can fold is the planar structure of the peptide bond. Figure 3-3 illustrates the amide group in peptide bonds in a polypeptide chain. Because the peptide bond itself behaves somewhat like a planar double bond, the portions (X and Y) of the polypeptide chain on either side of the peptide bond can be oriented in either a trans or cis configuration relative to the peptide bond.



We saw similar cis and trans isomers of carbon-carbon double bonds in unsaturated fatty acids in Chapter 2. Analysis of crystal structures indicates that in proteins, about 99.97 percent of the peptide bonds that have any residue other than proline at Y are in the trans configuration. (We will consider those with proline at Y shortly.) In a peptide bond, the carbonyl carbon and amide nitrogen and those atoms directly bonded to them must all lie in a fixed plane (Figure 3-16); little rotation about the peptide bond itself is possible. Yet a polypeptide chain must be flexible, able to twist and turn — and thus fold into different three-dimensional shapes. The only flexibility in a polypeptide chain itself is rotation of the fixed planes of adjacent peptide bonds with respect to one another about two

bonds: the C_α–amino nitrogen bond (rotational angle called Φ) and the C_α–carbonyl carbon bond (rotational angle called Ψ).

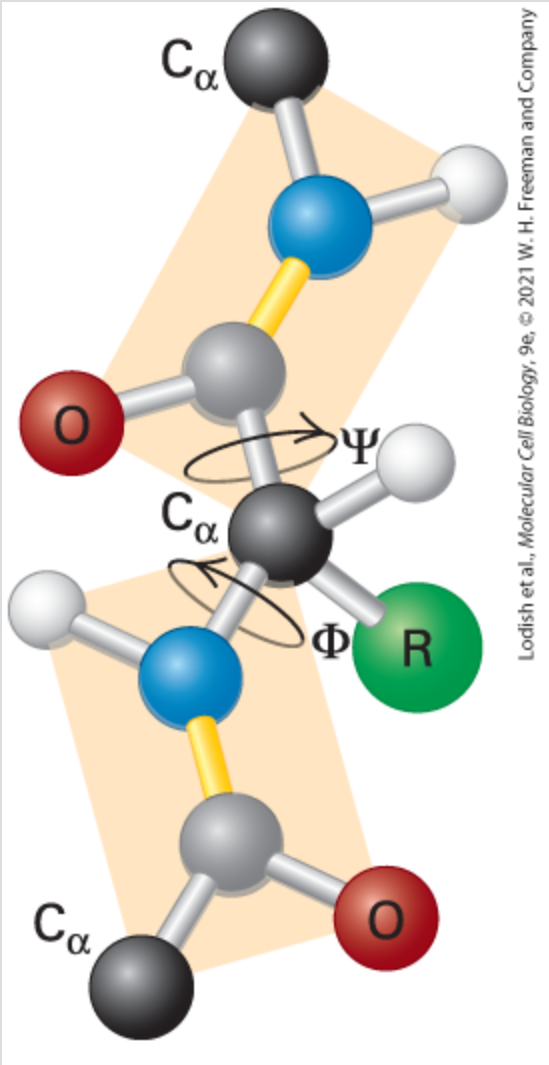


FIGURE 3-16 Rotation between planar peptide groups in proteins. In a peptide bond, the carbonyl carbon and amide nitrogen and those atoms directly bonded to them all lie in a fixed plane (orange rectangles). Rotation is possible about the C_α–amino nitrogen bond (the Φ angle) and the C_α–carbonyl carbon bond (the Ψ angle).

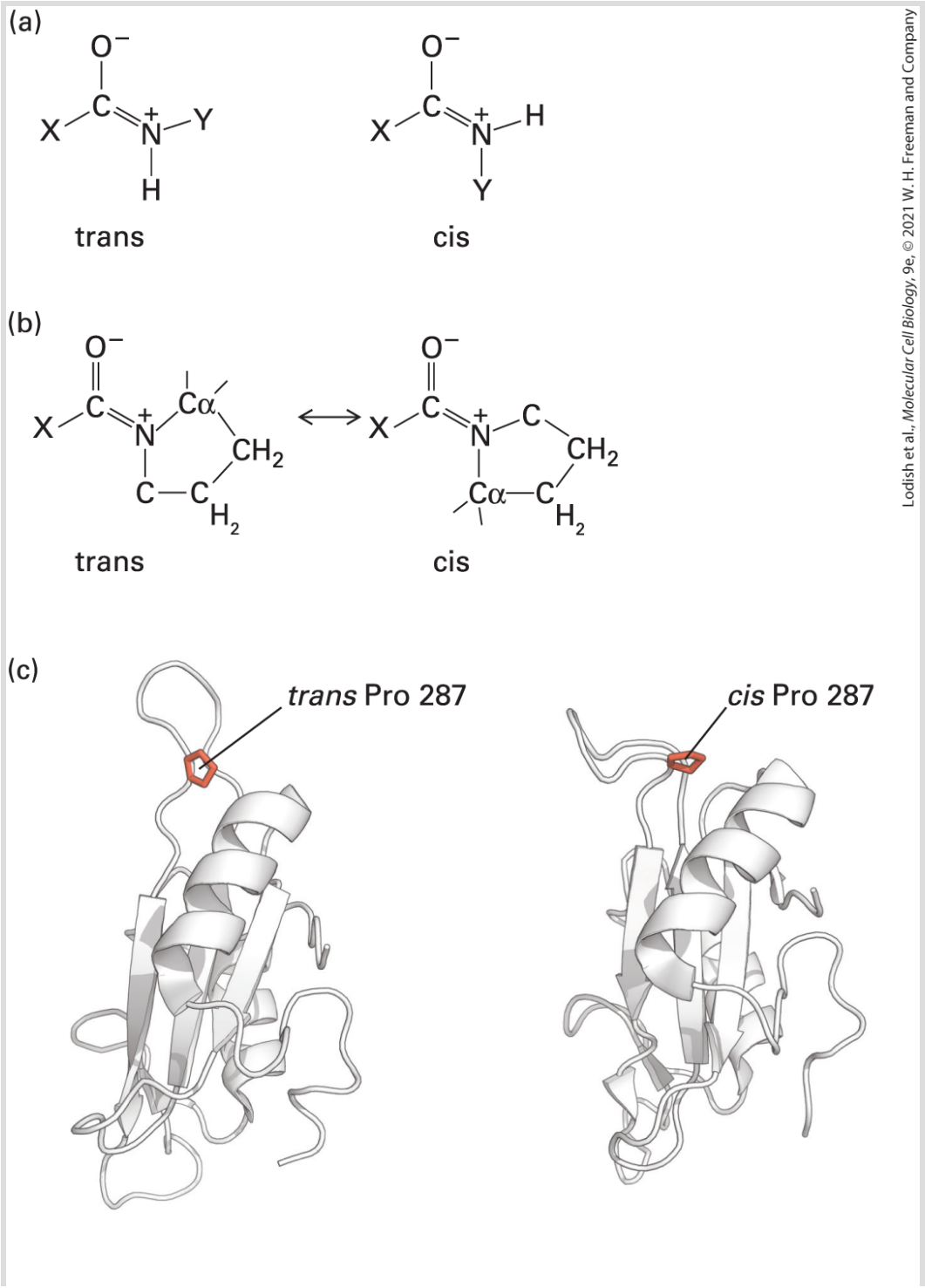
Through this rotation polypeptide backbones are able, in principle, to adopt a very large number of potential conformations. However, steric restraints due to the structure of the polypeptide backbone and the properties of the amino acid side chains dramatically restrict the potential conformations that any given protein can assume.

A further constraint imposes additional limits on the potential conformations that a polypeptide chain can adopt. Only a limited number of Φ and Ψ angles are possible because, for most Φ and Ψ angles, the backbone or side-chain atoms would come too close to one another, and thus the associated conformation would be highly unstable or even physically impossible to achieve.

Protein Folding Is Promoted by Proline Isomerases

As we have seen, the portions of the polypeptide chain on either side of a peptide bond (X and Y) are almost always oriented in a trans configuration (Figure 3-17a). However, the trans configuration is not dramatically more energetically favorable than a cis configuration when there is a proline at Y (Figure 3-17b). Among those folded proteins whose structures have been determined, about

5–7 percent of peptide bonds with proline at Y exhibit the cis configuration, as compared with 0.03 percent of all other peptide bonds without proline at Y.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 3-17 Proline cis/trans isomerizations influence protein folding and structure. (a) Due to the planar, double bond–like character of peptide bonds, the portions of the polypeptide chain on either side (X and Y) may have cis or trans configurations. The trans configuration is present in about 99.97 percent of all peptide bonds in well-ordered proteins when Y is a residue other than proline. (b) When Y is proline, about 5–7 percent of peptide bonds are in the cis configuration. Proline isomerases catalyze the cis/trans isomerization to facilitate protein folding. (c) The structure of a portion of a protein, here an SH2 protein domain (see Chapter 16), can be dramatically altered by the cis/trans isomerization of a single proline, and this structural change can influence the protein's activity.

[Part (c) Trans data from E. V. Pletneva et al., 2006, *J. Mol. Biol.* **357**:550–561, PDB ID 2etz; cis data from R. J. Mallis et al., 2002, *Nat. Struct. Biol*. **9**:900–905, PDB ID 1lui.]

The spontaneous rate of isomerization between the cis and trans configurations of proline-containing peptide bonds is relatively slow (milliseconds to seconds), whereas many protein-folding events occur more quickly (microseconds to milliseconds). Thus cells use peptidyl-proline isomerase (PPIases) proteins to catalyze these cis/trans isomerizations so that the proline in the folding protein quickly forms the proper isomer. Cis/trans isomerization of proline also can act as a switch to alter the conformation, and thus the activity, of an already stably folded protein. Indeed, such isomerizations can substantially alter the structure of some proteins (Figure 3-17c). Humans have 45 different PPIases that catalyze these critical folding and conformational switch reactions.

The Amino Acid Sequence of a Protein Determines How It Will Fold

While the constraints of backbone bond angles seem very restrictive, any polypeptide chain containing only a few residues could, in principle, still fold into many conformations. For example, if the Φ and Ψ angles were limited to only 8 combinations, an n-residue-long peptide would potentially have 8^n conformations; for even a small polypeptide of only 10 residues, that's about 8.6 million possible conformations! In general, however, the native state of any particular protein that is not intrinsically disordered adopts only one or a few very closely related conformations; for the vast majority of these proteins, the native state is a stably folded form of the molecule and the one that permits it to function normally. In thermodynamic terms, the native state is usually the conformation with the lowest free energy (G) (see Chapter 2).

What features of natively well-ordered proteins limit their folding from so many potential conformations to just one or a few? The properties of the side chains (e.g., size, hydrophobicity, charge, ability to form hydrogen and ionic bonds), together with their particular sequence along the polypeptide backbone, impose key restrictions. For example, a large side chain, such as that of tryptophan, might sterically block one region of the chain from packing closely against another region, whereas a side chain with a positive charge, such as that of arginine, might attract a segment of

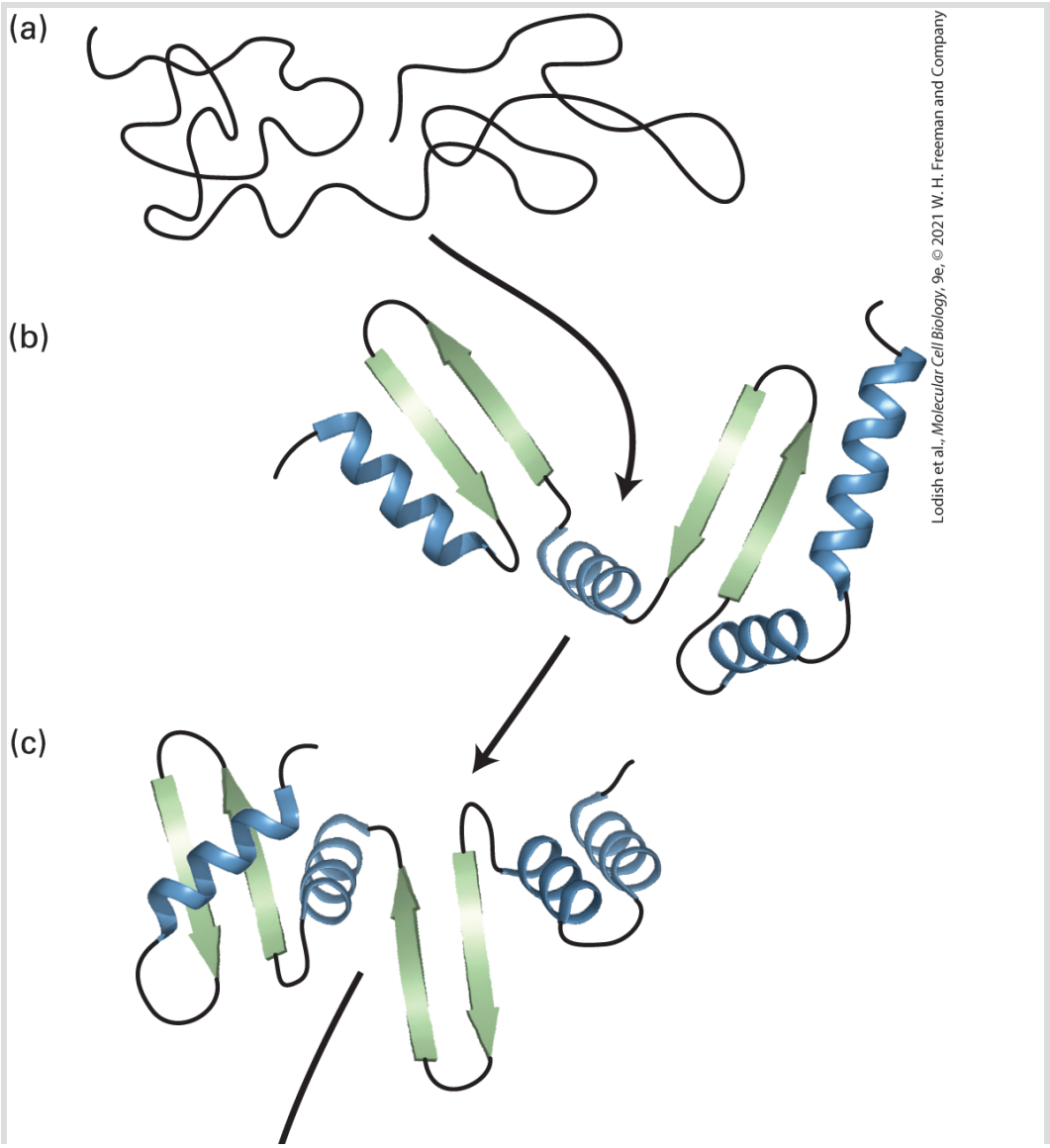
the polypeptide that has a complementary negatively charged side chain (e.g., aspartic acid). Another example we have already discussed is the effect of the aliphatic side chains in heptad repeats in promoting the association of helices and the consequent formation of coiled coils. *Thus a polypeptide's primary structure determines its secondary, tertiary, and quaternary structures*.

The initial evidence that the information necessary for a protein to fold properly is encoded in its amino acid sequence came from in vitro (in test tubes) studies on the refolding of purified proteins, especially the Nobel Prize–winning studies in the 1960s by Christian Anfinsen of the refolding of ribonuclease A, an enzyme that cleaves RNA. Others had previously shown that various chemical and physical perturbations can disrupt the weak noncovalent interactions that stabilize the native conformation of a protein, leading to the loss of its normal tertiary structure. The disruption of a protein's structure (and this can include secondary as well as tertiary structure) is called **denaturation**. Denaturation can be induced by thermal energy from heat, extremes of pH that alter the charges on amino acid side chains, or exposure to *denaturants* such as urea or guanidine hydrochloride at concentrations of 6–8 M, all of which disrupt structure-stabilizing noncovalent interactions. Treatment with reducing agents, such as β-mercaptoethanol, that break disulfide bonds can further destabilize disulfide-containing proteins. Under denaturing conditions, a population of uniformly folded protein molecules is destabilized and converted into a

collection of many unfolded, or denatured, molecules that have many different non-native and biologically inactive conformations. As we have seen, a large number of possible non-native conformations exist (e.g., $8^n - 1$). Two broad classes of non-native conformations are seen in proteins: (1) monomeric unfolded or denatured structures and (2) aggregates, which can either be amorphous or have a well-organized structure, as is the case for the disease-associated amyloid fibrils described later in this chapter. In principle, aggregates can comprise many copies of a single protein (homogeneous aggregates) or contain a mixture of distinct proteins (heterogeneous aggregates).

The spontaneous unfolding of proteins under denaturing conditions is not surprising, given the substantial increase in entropy that occurs because a denatured protein can adopt many non-native conformations (increased disorder). What is striking, however, is that when a pure sample of a single type of unfolded protein in a test tube is shifted back very carefully to normal conditions (body temperature, normal pH levels, reduction in the concentration of denaturants), some denatured polypeptides can spontaneously refold into their native, biologically active states, as in Anfinsen's experiments. This kind of refolding experiment, as well as studies showing that proteins synthesized in a test tube can fold properly, established that the information contained in a protein's primary structure can be sufficient to direct correct refolding. Newly synthesized proteins appear to fold into their proper conformations

just as denatured proteins do. The observed similarity in the folded, three-dimensional structures of proteins with similar amino acid sequences, noted in Section 3.1, provided additional evidence that the primary sequence also determines protein folding in vivo (in live organisms). It appears that secondary structures and structural motifs form early in the folding process, followed by assembly of more complex structural domains, which then associate into more complex tertiary and quaternary structures (Figure 3-18).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

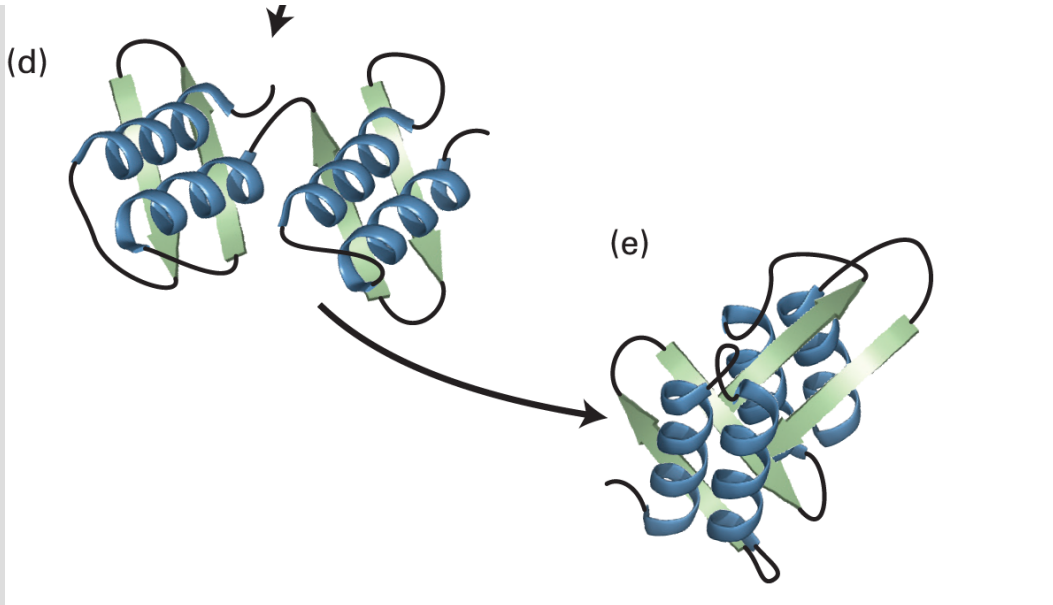


FIGURE 3-18 Hypothetical protein-folding pathway. Folding of a monomeric protein follows the structural hierarchy of primary (a) → secondary (b–d) → tertiary (e) structure. Formation of small structural motifs (c) appears to precede formation of domains (d) and the final tertiary structure (e).

Folding of Proteins In Vivo Is Promoted by Chaperones

The conditions under which a purified, denatured protein refolds in a test tube differ markedly from the conditions under which a newly synthesized polypeptide folds in a cell. The presence of other biomolecules, some of which are themselves nascent and in the process of folding, can potentially interfere with the autonomous, spontaneous folding of an otherwise natively well-ordered protein by forming aggregates. The cytosolic concentrations of some proteins are very high, and the total cytosolic protein concentration

can be ~300 mg/ml in mammalian cells. These high-protein concentrations favor the formation of aggregates by increasing the chances a nascent protein will encounter other proteins prior to completing its folding. Unfolded and partly folded proteins tend to aggregate into large, often water-insoluble masses, from which it is extremely difficult for a protein to dissociate and then fold into its proper conformation on its own. In part, this aggregation is due to the exposure of hydrophobic side chains that have not yet had a chance to be buried in the inner core of the folded protein. Exposed hydrophobic side chains on different molecules will stick to one another, owing to the hydrophobic effect (see Chapter 2) and thus promote aggregation. The risk of such aggregation is especially high for newly synthesized proteins that have not yet completed their proper folding. Intrinsically disordered proteins are much less likely to form such deleterious aggregates because, at least in some cases, they have relatively fewer hydrophobic side chains. Although some proteins can fold into a well-ordered native state in vitro, not all unfolded molecules are able to fold properly in a timely fashion because of the very large number of potentially incorrect, intermediate conformations into which the protein might fold.

Given such impediments, cells require faster, more efficient mechanisms for folding natively well-ordered proteins into their correct shapes than sequence alone provides. Without such help, cells might waste much energy in the synthesis of improperly folded, nonfunctional proteins, which would have to be destroyed to

prevent their disrupting cell function. Cells clearly have such mechanisms, since more than 95 percent of the proteins present within cells have been shown to be in their native conformations. Proteins that do not or cannot fold properly — for example, those encoded by genes with mutations that alter the amino acid sequence — are often recognized as unfolded and rapidly degraded by enzymes.

The explanation for the cell's remarkable efficiency in promoting proper protein folding is that cells make a set of proteins, called **chaperones**, that facilitate proper folding of nascent proteins. One way that chaperones facilitate proper folding is that they prevent aggregation by binding to the target polypeptide or sequestering it from other partially or fully unfolded proteins, thereby giving the nascent protein time to fold properly without colliding into other unfolded proteins. The importance of chaperones is highlighted by the observation that many are evolutionarily conserved. Chaperones are found in all organisms from bacteria to humans, and some are homologs with high sequence similarity that use almost identical mechanisms to assist protein folding. In eukaryotes, chaperones are located not only in the cytosol but also within organelles, like the endoplasmic reticulum, mitochondria, and the nucleus.

Chaperones help fold newly made proteins into functional conformations or refold misfolded or unfolded proteins into functional conformations. Sometimes in a cycle of chaperone-assisted protein folding, the details of which are described below,

the target protein fails to adopt a mature, stable conformation. Chaperones can then re-engage with the target for additional cycles of chaperone-mediated folding until proper folding is achieved. Chaperones can also disassemble potentially toxic protein aggregates that form due to protein misfolding; assemble and dismantle large multiprotein complexes; and mediate transformations between inactive and active forms of some proteins (such as the open and closed forms of some protein kinases, which we discuss shortly). Thus chaperones are very important!

Chaperones bind to the target proteins — also called *substrates* or *client proteins* — whose folding they will assist. Chaperones use a cycle of ATP binding, ATP hydrolysis to ADP, and exchange of a new ATP molecule for the ADP to induce a series of conformational changes that are essential for their function. There are several different classes of chaperones with distinct structures, but all use ATP binding and hydrolysis in a variety of ways, which include (1) to enhance the binding of the target protein and (2) to switch their own conformation. This ATP-dependent conformational switching is used in turn (1) to optimize folding, (2) to return the chaperone to its initial state so that it is available to help fold another molecule, and (3) to set the time permitted for refolding, which can be determined by the rate of ATP hydrolysis.

Two general families of chaperones have been identified:

- Molecular chaperones, which bind to a short segment of a protein substrate and stabilize unfolded or partly folded proteins, thereby preventing these proteins from aggregating and being degraded. Molecular chaperones bind to the nascent chain of a protein as it is synthesized and as it exits the ribosome and therefore begin binding to a new protein even before its synthesis is complete (see Chapter 13).
- Chaperonins, which form folding chambers into which all or part of an unfolded protein can be sequestered, giving it time and an appropriate environment to fold properly.

Molecular Chaperones

Two major types of molecular chaperone are Hsp70 and Hsp90. The substrates of Hsp70 are unfolded proteins, such as newly synthesized proteins. Hsp70 can function in concert with Hsp90 to help fold partially folded proteins or to remodel or otherwise change the conformation of certain folded proteins. The mechanisms by which Hsp70 and Hsp90 participate in folding other proteins are shown in Figure 3-19.

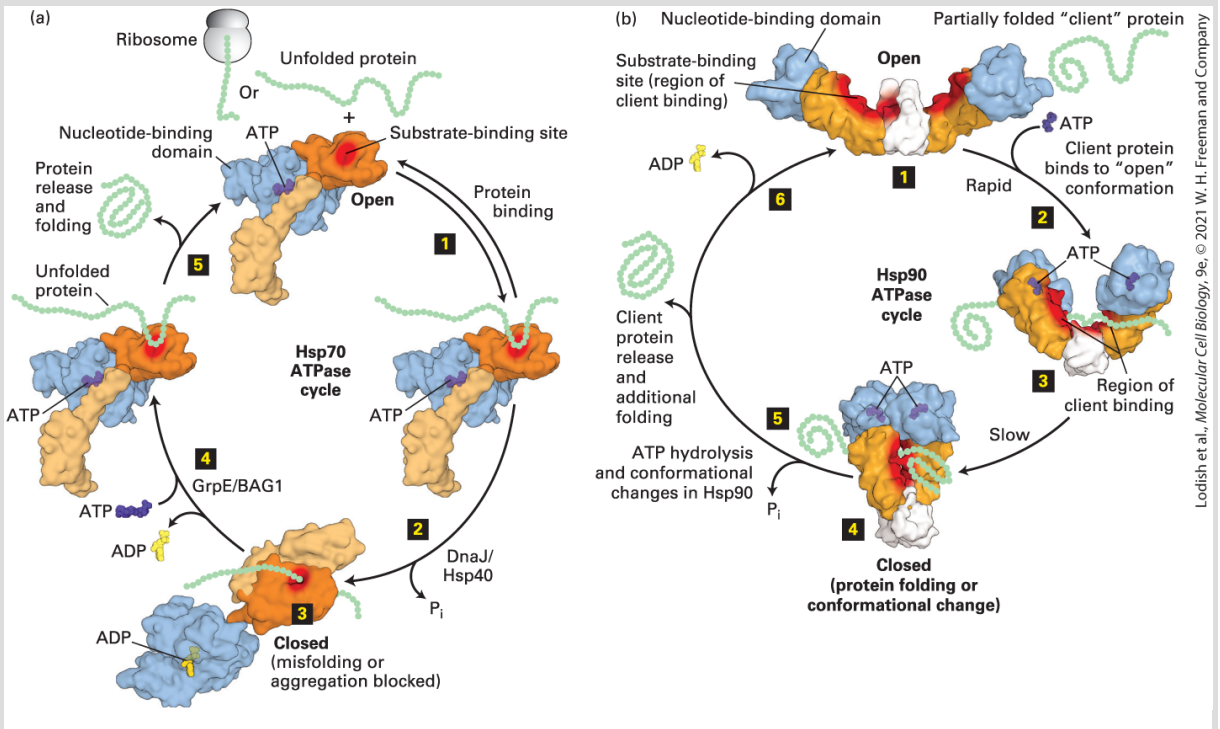


FIGURE 3-19 Protein folding mediated by molecular chaperones. (a) Many proteins fold into their proper three-dimensional structures with the assistance of Hsp70 or one of several Hsp70-like proteins. These molecular chaperones transiently bind to a nascent polypeptide as it emerges from a ribosome or to a protein that is unfolded. In the Hsp70 cycle, an unfolded substrate protein binds in equilibrium (step 1) to Hsp70's substrate-binding site (red) in the open conformation of its substrate-binding domain (light and dark orange) when an ATP (purple) is bound at Hsp70's nucleotide-binding domain (light blue). The substrate-binding domain comprises two subdomains (light and dark orange) that change relative positions and conformations during the cycle. Co-chaperone accessory proteins (DnaJ/Hsp40) stimulate the hydrolysis of ATP to ADP (yellow) that induces a large conformational change in the substrate-binding domain, resulting in the closed conformation, in which the substrate is tightly locked into the substrate-binding domain, preventing inappropriate folding or aggregation (steps 2 and 3). Exchange of cytosolic ATP for the bound ADP, stimulated by other accessory, co-chaperone proteins (GrpE/BAG1), converts the Hsp70 back to the open conformation (step 4), releasing the substrate and freeing it to continue folding (step 5). With the substrate released, Hsp70 can then interact with additional substrates. If the released substrate does not fold properly, it can rebind to another chaperone and the process is repeated. (b) In many, but not all, cases, a substrate protein partially folded by Hsp70 will be transferred to Hsp90 for further folding. Three conformational states of the dimeric Hsp90 molecular chaperone are thought to be involved

in substrate (also called *client*) protein remodeling. Client proteins bind at the substrate-binding site (red surface) shared by the substrate-binding (orange) and C-terminal dimerization (white) domains and are thought to change conformation in response to ATP binding and hydrolysis. The Hsp90 cycle begins when there is no nucleotide bound to the nucleotide-binding domains (light blue) and the dimer is in a very flexible, open configuration (step **1**) that can bind a client. Rapid ATP binding leads to a conformational change (step **2**) in which the nucleotide-binding domains and the substrate-binding domains move together (intermediate shown in step **3**) into a closed conformation in which the nucleotide-binding domains are bound to each other (step **4**). With Hsp90 in its closed conformation, the client protein may undergo folding. In some cases, when a folded protein is bound to the closed Hsp90 (step **4**) it can undergo a conformational change. The precise locations in Hsp90 at which clients bind apparently vary for different clients. ATP hydrolysis results in a conformational change in Hsp90 (step **5**) that may include a highly compact form, folding of the client, client protein release and additional folding of the unbound client. The ADP-bound form of Hsp90 can adopt several conformations, including a highly compact form. Release of ADP (step **6**) completes the cycle by regenerating the initial flexible open state, which can then interact with additional clients. The order of substrate release, ATP hydrolysis, and ADP release remain to be established definitively. See E. D. Kirschke et al., 2014, *Cell* **157**:1685; and M. Taipale, D. F. Jarosz, and S. Lindquist, 2010, *Nat. Rev. Mol. Cell Biol.* **11**:515.

[Solvent-accessible surface model of HSP90 courtesy of Elaine Kirschke and David A. Agard, UCSF. Open (ATP) PDB ID 2ior, closed (ATP) PDB ID 2cg9, closed (ADP) based on PDB ID 2cg9.]

Hsp70

The heat-shock protein Hsp70 in the cytosol and its homologs were first identified by their rapid appearance after a cell had been stressed by an increase in temperature, that is, heat shock (*Hsp* stands for "*h*eat-*s*hock *p*rotein"). The homologs of cytosolic Hsp70 include Hsp70 in the mitochondrial matrix, BiP in the endoplasmic reticulum, and DnaK in bacteria. Hsp70 and its homologs are the

major chaperones in all organisms that use an ATP-dependent cycle to facilitate proper folding of their substrates (Figure 3-19a). The five steps in the Hsp70 cycle are:

1. Substrate binding to the "open" ATP-bound Hsp70.
2. Substrate binding induces ATP hydrolysis to ADP that leads to a "closed" ADP-bound Hsp70 conformation that binds substrate much more tightly.
3. Inhibition of substrate misfolding or aggregation while bound tightly to the closed Hsp70. In a sense, Hsp70 acts by inhibiting inappropriate protein folding.
4. Exchange of ATP for ADP on the Hsp70 with its conversion to the "open" conformation.
5. Release of the substrate protein and its subsequent folding.

If the target is now properly folded, it cannot rebind to an Hsp70. If it remains at least partially unfolded, it can bind again to give a chaperone another chance to help fold it properly. A variety of co-chaperone proteins facilitate this process. As we will see later in this chapter, a variety of proteins use a cycle of trinucleotide hydrolysis to a dinucleotide, followed by dinucleotide/trinucleotide exchange, to control their activities. We will discuss a group of proteins called *GTPases* that depend on the exchange of GTP, rather than ATP, for bound GDP (instead of ADP) to induce conformational changes that dramatically influence the proteins' activities and the subsequent hydrolysis of the bound GTP to GDP.

Additional proteins, such as the co-chaperone Hsp40 in eukaryotes (DnaJ in bacteria), help increase the efficiency of the Hsp70-mediated folding of many proteins not only by stimulating the binding of substrate but also by increasing the rate of hydrolysis of ATP by 100- to 1000-fold (see step **2** in Figure 3-19a). Members of four different families of nucleotide exchange factors (e.g., GrpE in bacteria; BAG, HspBP, and Hsp110 in eukaryotes) also interact with Hsp70 (or DnaK in bacteria), promoting the exchange of ATP for ADP (see step **4**). Multiple molecular chaperones are thought to bind to all nascent polypeptide chains as they are being synthesized on ribosomes. In bacteria, 85 percent of the proteins are released from their chaperones and proceed to fold normally; an even higher percentage of proteins in eukaryotes follow this pathway.

Hsp90

Hsp90-family molecular chaperones usually recognize partially folded substrate proteins, generally called *clients*. Evolutionarily related Hsp90 family members are present in all organisms except archaea. Their strong evolutionary conservation is seen in the high amino acid sequence similarity (55 percent) of the Hsp90 from the bacterium *Escherichia coli* and human Hsp90. In most eukaryotes, there are four distinct Hsp90s: two in the cytosol and one each in the endoplasmic reticulum and the mitochondrion. At 1–2 percent of total protein, Hsp90 is one of the most abundant cytosolic proteins. Although the range of protein substrates for Hsp90 chaperones is

not as broad as for some other chaperones (at least 10 percent of yeast proteins are thought to be Hsp90 substrates), the Hsp90s are essential in eukaryotes. The Hsp90s help cells fold partially folded clients (see Figure 3-19b), cope with denatured proteins generated by stress (e.g., heat shock), and ensure that some of their clients can be converted from an inactive to an active state or otherwise held in a functional conformation. In some cases, an Hsp90 forms a relatively stable complex with a client until an appropriate signal causes it to dissociate, freeing the client to perform some regulated function in the cell. Hsp90 clients include transcription factors such as the receptors for many steroid hormones, like estrogen and testosterone. These steroid receptors regulate sexual development and function by controlling the activities of many genes (see Chapter 8). Another type of Hsp90 client is the set of enzymes called *kinases*, which control the activities of many other proteins by phosphorylation (see Section 3.4 and Chapters 15 and 16).

Unlike monomeric Hsp70, Hsp90 functions as a dimer in a cycle in which ATP binding, hydrolysis, and ADP release are coupled to major conformational changes and to binding, folding or activation, and release of clients (see Figure 3-19b). Although much about the mechanism of Hsp90 remains to be learned, it is clear that partially folded clients bind to the substrate-binding domains when the chaperone is in the "open" conformation (step **1** in Figure 3-19b), that ATP binding leads to interaction of the ATP-binding domains and formation of a "closed" conformation (steps **1**–**4** in Figure 3-

19b), and that hydrolysis of ATP plays an important role in the activation of some client proteins and their subsequent release from the Hsp90 (steps **5** and **6**). The order of substrate release, ATP hydrolysis and ADP release remain to be established definitively. We also know that there are at least 20 co-chaperones that can profoundly affect the activity of Hsp90, including modulating its ATPase activity and determining which proteins will be clients (client specificity). Co-chaperones can also help coordinate the activities of Hsp90 and Hsp70. For example, Hsp70 can help begin the folding of a client that is then handed off by a co-chaperone to Hsp90 for additional processing. Hsp90 activity can also be influenced by its covalent modification by small molecules. Finally, Hsp90 can help cells recognize misfolded proteins that are unable to refold and facilitate their degradation by mechanisms discussed later in this chapter. Thus, as part of the quality-control system in cells, molecular chaperones can help proteins fold properly or facilitate the destruction of those that cannot fold properly. In some cases, Hsp70s, Hsp90s, and their co-chaperones can bind to their clients in a well-defined order during the folding process.

Chaperonins

The proper folding of a large variety of newly synthesized proteins also requires the assistance of another class of proteins, the chaperonins, also called *Hsp60s*. Chaperonins are huge cylindrical supramolecular assemblies. In the center of chaperonins are

chambers into which all or part of an unfolded protein enters. The chambers provide an environment conducive to protein folding that occurs in the context of cycles of ATP binding and hydrolysis and the opening and closing of lids that seal the chambers during the folding process, as described below. The folding chambers are formed from two rings of oligomers. There are two distinct groups of chaperonins that differ somewhat in their structures, detailed molecular mechanisms, and locations. Group I chaperonins are found in prokaryotes, chloroplasts, and mitochondria. They are composed of two rings, each having seven subunits that interact with a co-chaperone *lid* that also has seven subunits. Each ring is a folding chamber into which an unfolded protein enters. The bacteria group I chaperonin, known as *GroEL/GroES*, is shown in Figure 3-20a. In the bacterium *E. coli*, GroEL is thought to participate in the folding of about 10 percent of all proteins. Group II chaperonins are found in the cytosol of eukaryotic cells and in archaea. These chaperonins can have eight to nine homomeric or heteromeric subunits in each ring, and the lid function is incorporated into those subunits themselves — no separate lid protein is needed. It appears that ATP hydrolysis triggers the closing of the lid of group II chaperonins.

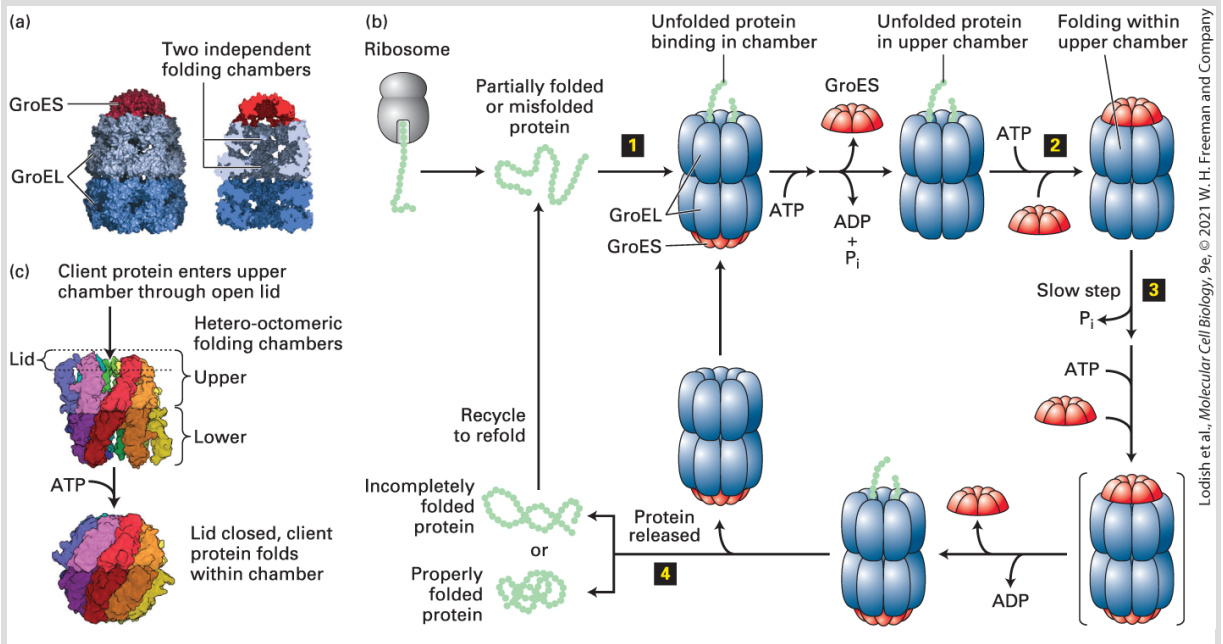


FIGURE 3-20 Protein folding mediated by chaperonins. Proper folding of some proteins depends on chaperonins, such as the prokaryotic group I chaperonin GroEL and the group II chaperonin TRiC. (a) GroEL is a barrel-shaped complex of fourteen identical ~60,000-MW subunits, arranged in two stacked rings (blue) of seven subunits each that form two distinct internal polypeptide folding chambers. GroES lids (red) of seven identical subunits can bind to either end of the barrel and seal the chamber on that side (cutaway view on the right). (b) The GroEL-GroES folding cycle. A partly folded or misfolded polypeptide enters one of the folding chambers (step 1). The second chamber is blocked by a GroES lid. Each ring of seven GroEL subunits binds seven ATPs, hydrolyzes them, and then releases the ADPs in a set order coordinated with GroES binding and release and polypeptide binding, folding, and release. The major conformational changes that take place in the GroEL rings control the binding of the GroES lid that seals the chamber (step 2). The polypeptide remains encased in the chamber capped by the lid, where it can undergo folding until ATP hydrolysis — the slowest, rate-limiting step in the cycle ($t_{1/2}$~10 s) (step 3) — induces binding of ATP and a different GroES to the other ring (transient intermediate shown in brackets). This binding then causes the GroES lid and ADP bound to the peptide-containing ring to be released, opening the chamber and permitting the folded protein to diffuse out of the chamber (step 4). If the polypeptide has folded properly, it can proceed to function in the cell. If it remains partially folded or misfolded, it can rebind to an unoccupied GroEL and the cycle can be repeated. See D. L. Nelson and M. M. Cox, 2013, *Lehninger Principles of Biochemistry*, 6th ed., Macmillan. (c) CryoEM structures of lid open (top) and lid closed (bottom) yeast TRiC. Two hetero-octamers form the upper and lower hemispheres of eukaryotic cytosolic

group II chaperonins (TRiC). Clients enter the folding chambers through open lids (dashed black lines). The binding and hydrolysis of ATP induces lid closing and folding of the client within the sealed chamber of a hemisphere. Release of inorganic phosphate and ADP opens the lid, releasing the client, and resets the chamber for additional rounds of folding (not shown).

[Part (a) Data from Z. Xu, A. L. Horwich, and P. B. Siegler, 1997, *Nature* **388**:741–750, PDB ID 1aon. Part (c) Data from movie S1, from M. Jin et al., 2019, *Proc. Nat'l. Acad. Sci. USA* **116**(39):19513–19522.]

Figure 3-20b illustrates the GroEL/GroES cycle of protein folding. A partially folded or misfolded polypeptide of less than 60 kDa in mass is captured by hydrophobic residues near the entrance of the GroEL chamber and enters one of the folding chambers (upper chamber in Figure 3-20b). The second chamber is blocked by a GroES lid. Each of the 14 subunits of GroEL can bind ATP, hydrolyze it, and subsequently release ADP. These reactions are concerted for each set of seven subunits in a single ring and lead to major conformational changes. The binding of ATP causes the GroES lid to seal the chamber and alters the environment of the chamber in which polypeptide folding takes place. The polypeptide remains encased in the chamber capped by the lid. There it can undergo folding until ATP hydrolysis in that chamber, which is the slowest, rate-limiting step in the cycle, induces ATP and a different GroES to bind to the other ring. This then causes the GroES lid and ADP bound to the peptide-containing ring to be released, opening the chamber and permitting the folded protein to diffuse out. If the polypeptide is folded properly, it can proceed to function in the cell.

If it remains partially folded or misfolded, it can rebind to an unoccupied GroEL and the cycle can be repeated.

There is a reciprocal relationship between the two rings in one GroEL complex. The capping of one chamber by GroES to permit substrate folding to proceed in that chamber is accompanied by the release of substrate polypeptide from the other chamber (simultaneous binding, folding, and release from the second chamber is not illustrated in Figure 3-20b). There is a striking similarity between the capped-barrel design of GroEL/GroES, in which proteins are sequestered for folding, and the structure of the 26S proteasome that participates in protein degradation (discussed in Section 3.4). In addition, a group of proteins that are part of the AAA^+ family of ATPases are composed of hexameric rings with a central pore into which substrates can enter for folding or unfolding or in some cases proteolysis; examples of these will be discussed in Section 3.4 and in Chapter 13.

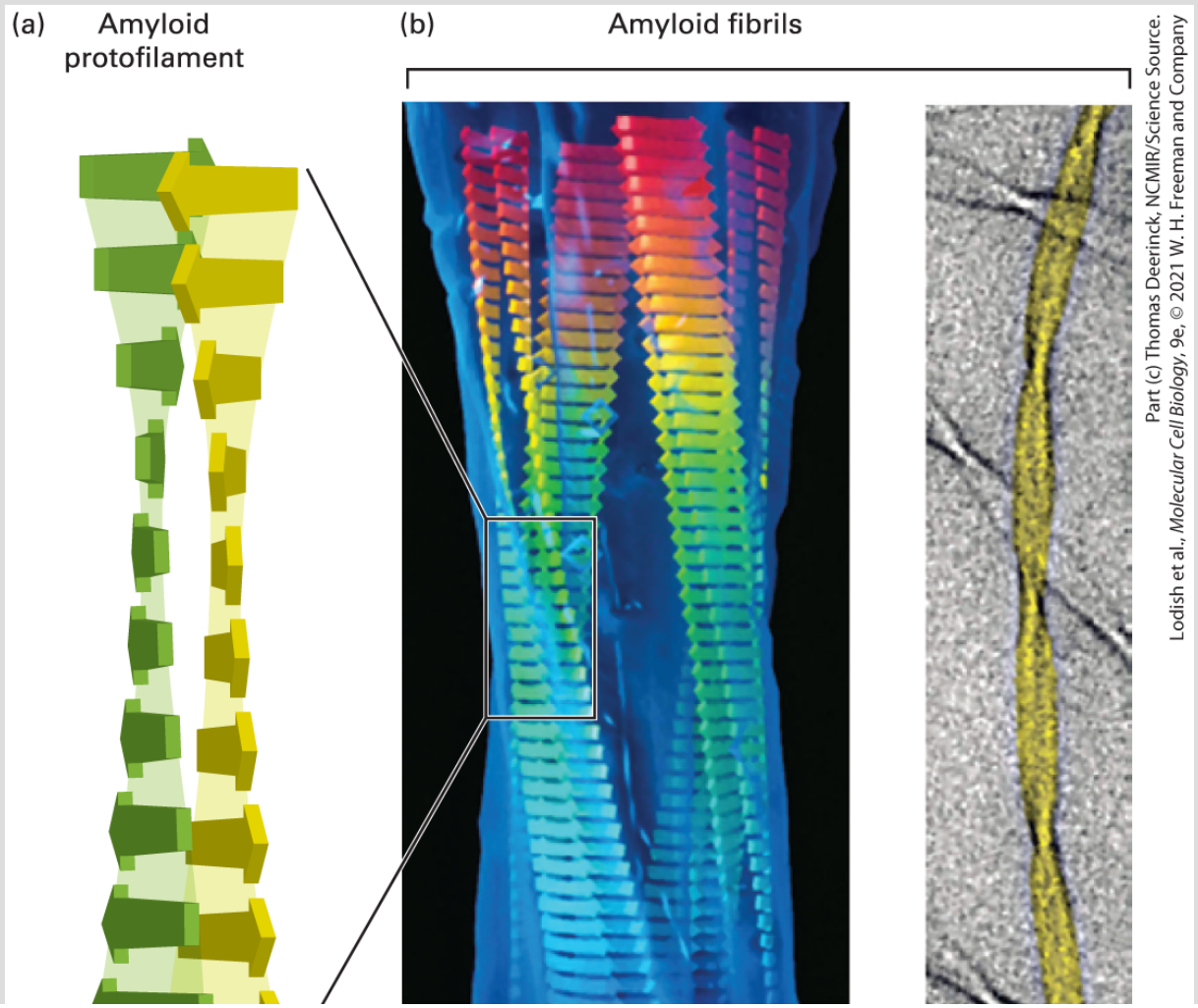
In the Group II chaperonins present in the cytosol of eukaryotes (e.g., yeast TRiC in Figure 3-20c), the consequences of ATP binding and hydrolysis differ from those in GroEL/GroES. For TriC, ATP binding and hydrolysis in the presence of bound client protein leads to closing of the lid and folding of the client in the sequestered environment within the folding chamber. Release of the inorganic phosphate opens the lid and releases the substrate and ADP release resets the TriC chamber so that it can bind another client molecule.

Abnormally Folded Proteins Can Form Amyloids That Are Implicated in Diseases

After it is synthesized, a protein may fold into an alternative, abnormal three-dimensional structure as the result of mutations, inappropriate covalent modifications, or chemical (e.g., pH) or physical (e.g., heat) alterations in its environment. A misfolded protein may not function normally and may be marked for destruction through proteolytic degradation, as described later in this chapter. However, when degradation is incomplete or fails to keep pace with the production of misfolded protein, the misfolded protein or its proteolytic fragments can accumulate either inside or outside of cells in aggregates, or *plaques*. Such plaques may be present in various organs, including joints between bones, the liver, and the brain. Even those proteins or protein fragments that are normally highly resistant to aggregation, as is the case for intrinsically disordered proteins or protein fragments, will form aggregates if their concentrations are sufficiently elevated or when there are changes in environmental conditions.

As noted above, such aggregates can either be amorphous or have a well-organized structure, which most commonly is the *amyloid state*. Strikingly, many diverse proteins can each aggregate into amyloid fibrils that have a common structure, called a *cross-β sheet* (Figure 3-21a). Short segments, generally 6–12 residues long, in the unfolded

or misfolded proteins hydrogen-bond to each other, forming a long array, or filament, of β sheets. In these arrays, each β strand is nearly perpendicular to the long axis of the filament, and two long, nearly flat β sheets pack closely together and twist around each other to form protofilaments, which then assemble together into thicker filaments, called *amyloid fibrils*. Within each protofilament the β strands can be either parallel or antiparallel (see Figure 3-5). Although some proteins form amyloid fibrils in their native, functional states, most amyloids are considered to be consequences of protein misfolding.



Part (c) Thomas Deerinck, NCMIR/Science Source.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

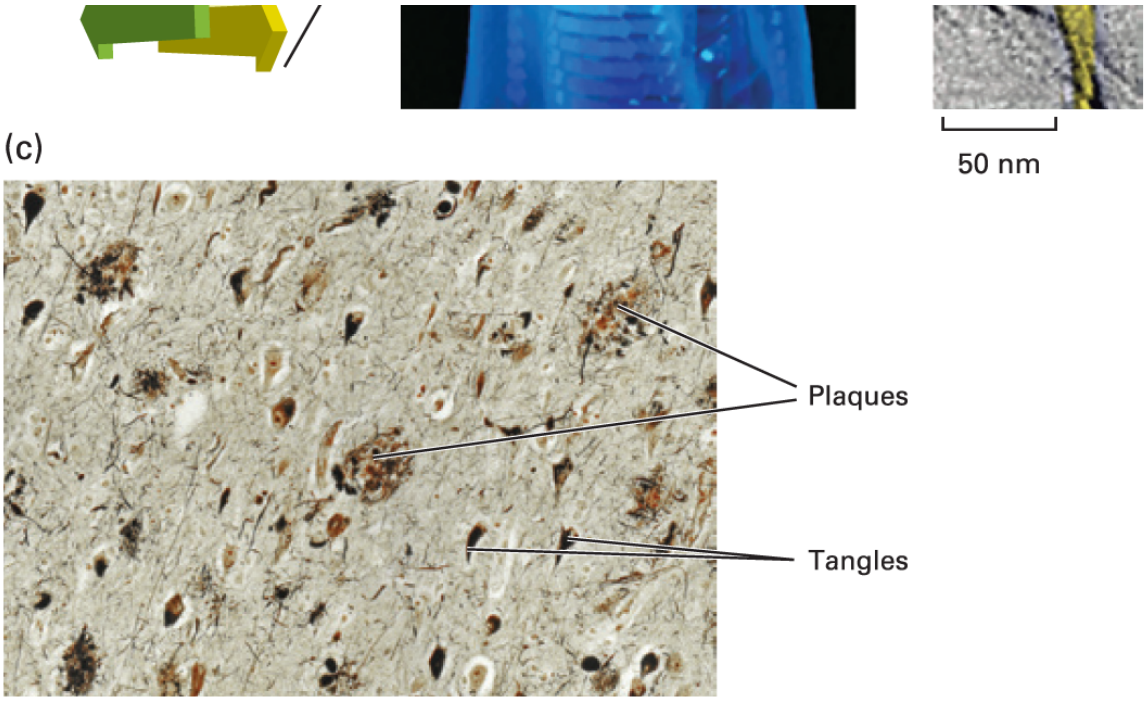


FIGURE 3-21 Misfolded proteins can form ordered amyloid aggregates based on a cross-β sheet structure. (a) In unfolded segments of proteins and polypeptides, exposed segments 6–12 residues long (short, flat arrows) can assemble into β sheets (see also Figure 3-5) in which each β strand is oriented nearly perpendicularly to the long axis (vertical in this figure) of the resultant amyloid protofilament and hydrogen-bonded (light shading) to the strands above and below. Two long, nearly flat sheets pack closely together and twist around each other to form amyloid protofilaments, which then assemble together into thicker filaments called *amyloid fibrils* (b). Amyloid fibrils can be composed of varying numbers of protofilaments. A model of a fibril containing four protofilaments fit into the electron density (blue) of acid-denatured insulin fibrils (*left*) and a cryoelectron microscopic image of fibrils containing two-protofilaments from fragments of transthyretin with an NMR-based model (yellow). Fibrils can aggregate into macroscopic plaques and tangles that are deposited in tissues and, when stained, are large enough to be visible using light microscopy. (c) Microscopic view of a section of human brain tissue from a patient with Alzheimer's disease showing multiple amyloid plaques and fibrillary tangles.

[Part (b, *left*) Republished with permission from Elsevier, from C. M. Dobson, 1999, "Protein Misfolding, Evolution and Disease," *Trends Biochem. Sci*. **24**(9):329–332, Fig. 3; permission conveyed through Copyright Clearance Center, Inc. Part (b, *right*) Reprinted with permission from Nature Publishing Group, from T. P. J. Knowles et al., 2014, "The Amyloid State and Its

Association with Protein Misfolding Diseases," *Nat. Rev. Mol. Cell Biol*. **15**(6):384–396, Fig. 3a; permission conveyed through Copyright Clearance Center, Inc.]

Amyloids were first recognized in protein aggregates that are deposited in tissues. They are resistant to enzymatic degradation and are associated with dozens of diseases, called *amyloidoses*. These diseases include neurodegenerative diseases such as Alzheimer's disease and Parkinson's disease in humans and transmissible spongiform encephalopathy ("mad cow" disease) in cows and sheep. Each of these diseases is characterized by the presence of filamentous plaques in a deteriorating brain (Figure 3-21b). Amyloidoses most commonly occur with aging; however, mutations in the genes encoding the aggregating protein can result in early amyloid formation and disease onset. The amyloid fibrils composing the plaques derive from abundant natural proteins. For example, fragments of the amyloid precursor protein, which is embedded in the plasma membrane, form the plaque found in the brains of patients with Alzheimer's disease; and prion protein, an infectious protein, forms fibrils in prion diseases. In Alzheimer's disease, a hyperphosphorylated form of the protein tau, normally a microtubule-binding protein (see Chapter 18), forms twisted fibers called *tangles*. These amyloids, either as relatively short, water-soluble protofilaments or as long, insoluble fibrils, are thought to be toxic and to contribute directly to the pathology of amyloidoses. ■

KEY CONCEPTS OF SECTION 3.2

Protein Folding

- The primary structure (amino acid sequence) of a protein determines its three-dimensional structure, which determines its function. In short, function derives from structure; structure derives from sequence.
- Because protein function derives from protein structure, newly synthesized proteins must fold into the correct shape to function properly.
- The planar structure of the peptide bond limits the number of conformations a polypeptide can have (see Figure 3-16). Peptidyl-proline isomerases facilitate formation of the proper conformation of proline-containing peptide bonds.
- The amino acid sequence of a protein dictates its folding into a specific three-dimensional conformation, the native state. Proteins will unfold, or denature, if treated under conditions that disrupt the noncovalent interactions stabilizing their three-dimensional structures.
- Two broad classes of non-native conformations are seen in misfolded or denatured proteins: (1) monomeric unfolded or denatured structures and (2) aggregates, which can either be amorphous or have a well-organized structure.
- Protein folding in vivo occurs with assistance from ATP-dependent chaperones. Chaperones can influence proteins in several ways, including by preventing misfolding and aggregation, facilitating proper folding, and maintaining an appropriate, stable structure required for subsequent protein activity (see Figure 3-19).
- There are two broad classes of chaperones: (1) molecular chaperones, which bind to a short segment of a substrate protein, and (2) chaperonins, which form folding chambers in which all or part of an unfolded protein can be sequestered, giving it time and an appropriate environment to fold properly. Cycles of ATP binding and hydrolysis, followed by exchange of the ADP produced with a new ATP molecule, play key roles in the mechanisms of protein folding by chaperones.
- Many misfolded or denatured proteins can form well-organized aggregates, called *amyloid fibrils*, made by short stretches of polypeptide that form a long array of β sheets nearly perpendicular to the fibril axis, called a *cross-β structure*. Formation of amyloid fibrils that are resistant to degradation by diverse enzymes is associated with dozens of diseases called *amyloidoses*. Examples include the neurodegenerative diseases Alzheimer's disease and Parkinson's disease.

3.3 Protein Binding and Enzyme Catalysis

Proteins perform an extraordinarily diverse array of activities both inside and outside cells, yet most of these diverse functions are based on the ability of proteins to engage in a common activity: binding. Proteins bind to one another, to other macromolecules, to small molecules, and to ions. In this section, we describe some key features of protein binding and then turn to look at one group of proteins, enzymes, in greater detail. The activities of the other functional classes of proteins (structure, movement, transport, and signaling) will be described in later chapters.

Specific Binding of Ligands Underlies the Functions of Most Proteins

The molecule to which a protein binds is called its **ligand**. In some cases, ligand binding causes a change in the shape of a protein. Such conformational changes are integral to the mechanism of action of many proteins and are important in regulating protein activity.

Two properties of a protein characterize how it binds ligands. **Specificity** refers to the ability of a protein to bind one molecule or a very small group of molecules in preference to all other molecules. *Affinity* refers to the tightness or strength of binding, usually

expressed as the dissociation constant (K_d). The K_d for a protein-ligand complex, which is the inverse of the equilibrium constant K_{eq} for the binding reaction, is the most common quantitative measure of affinity (see Chapter 2). The stronger the interaction between a protein and ligand, the lower the value of K_d. Both the specificity and the affinity of a protein for a ligand depend on the structure of the *ligand-binding site*. For high-affinity and highly specific interactions to take place, the shape and chemical properties of the binding site must be complementary to those of the ligand molecule, a property termed **molecular complementarity**. As we saw in Chapter 2, molecular complementarity allows molecules to form multiple noncovalent interactions at close range and thus stick together.

One of the best-studied examples of protein-ligand binding, involving high affinity and exquisite specificity, is the binding of antibodies to antigens. *Antibodies* are proteins that circulate in the blood and are made by the immune system in response to *antigens*, which are usually macromolecules present in infectious agents (e.g., a bacterium or a virus) or other foreign substances (e.g., proteins or polysaccharides in pollens). Different antibodies are generated in response to different antigens, and these antibodies have the remarkable characteristic of binding specifically to ("recognizing") the part of the antigen, called an *epitope*, that initially induced the production of the antibody, and not to other molecules. Antibodies act as specific sensors for antigens, forming antibody-antigen

complexes that initiate a cascade of protective reactions in cells of the immune system. Chapter 24 discusses antibodies, how they are generated, and their roles in the immune system, and later in this chapter we will discuss techniques for studying proteins that exploit antibodies. Here we briefly introduce the structure of antibodies and their binding to epitopes.

Antibodies are Y-shaped molecules, often formed from two identical longer, or *heavy*, chains and two identical shorter, or *light*, chains. In IgG antibodies (also called *immunoglobulins*, shown in Figure 3-22a and b), there are four globular domains in each heavy chain and two in each light chain, all of which are called *immunoglobulin (Ig) domains*. Each of the two branching arms of an IgG antibody contains a single light chain linked to a heavy chain by a disulfide bond, and two disulfide bonds covalently link the two heavy chains together. Near the end of each arm are six highly variable loops, called *complementarity-determining regions (CDRs)*, which form the antigen-binding sites. The sequences of the six loops are highly variable among antibodies, generating unique complementary ligand-binding sites that make them specific for different epitopes (Figure 3-22c). The intimate contact between antibody and epitope surfaces, stabilized by numerous noncovalent interactions, is responsible for the extremely precise binding specificity exhibited by an antibody.

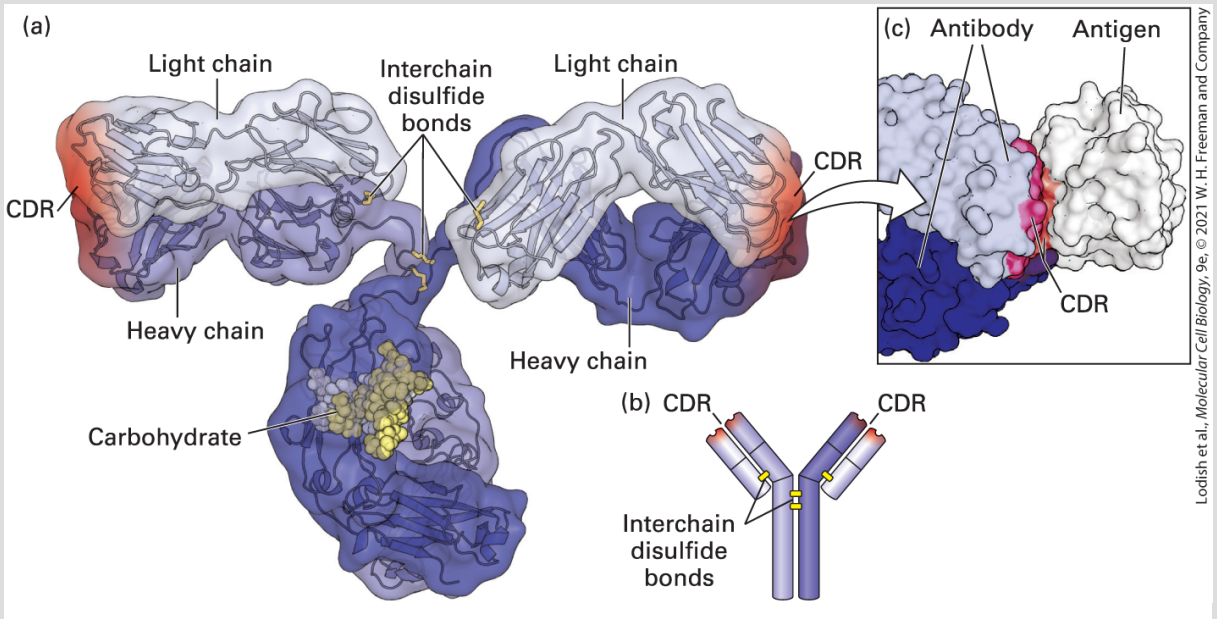


FIGURE 3-22 Protein-ligand binding of antibodies. (a) Hybrid (surface and ribbon) model of an antibody. Every antibody molecule of the immunoglobulin G (IgG) class consists of two identical heavy chains (medium and dark blue) and two identical light chains (light blue) covalently linked by disulfide bonds (yellow). The complementarity-determining regions (CDRs) that define the antigen-binding sites are represented by red shading. (b) The cartoon shows the overall structure containing the two heavy (longer) and two light (shorter) chains, with yellow bars representing disulfide bonds. (c) The hand-in-glove fit between an antibody and the site to which it binds (epitope) on its target antigen — in this case, chicken egg-white lysozyme. The antibody contacts the antigen with residues from its CDRs.

[Part (a) Data from L. J. Harris et al., 1997, *Biochemistry* **36**:1581–1597, PDB ID 1igt. Part (b) Data from E. A. Padlan et al., 1989, *Proc. Nat'l. Acad. Sci. USA* **86**:5938–5942, PDB ID 3hfm.]

The specificity of antibodies is so precise that they can distinguish between the cells of individual members of a species and in some cases between proteins that differ by only a single amino acid, or even between proteins with identical sequences that differ only in their post-translational modifications. Because of their specificity and the ease with which they can be produced (see Chapter 24),

antibodies are highly useful reagents used in many of the experiments discussed in subsequent chapters.

We will see many examples of protein-ligand binding throughout this book, including binding of hormones to receptors (see Chapter 15), binding of regulatory molecules to DNA (see Chapter 8), and binding of cell-adhesion molecules to extracellular matrices (see Chapter 20), to name just a few. Here we focus on how the binding of one class of proteins, enzymes, to their ligands results in the catalysis of the chemical reactions essential for the survival and function of cells.

Enzymes Are Highly Efficient and Specific Catalysts

Proteins that catalyze chemical reactions — the making and breaking of covalent bonds — are called **enzymes**, and the ligands of enzymes are called **substrates**. Enzymes make up a large and very important functional class of proteins — indeed, almost every chemical reaction in the cell is catalyzed by a specific catalyst, usually an enzyme. Another form of catalytic macromolecule in cells is made from RNA. These RNAs are called **ribozymes** (see Chapter 5).

Thousands of different types of enzymes, each of which catalyzes a single chemical reaction or a set of closely related reactions, have

been identified. Certain enzymes are found in the majority of cells because they catalyze the synthesis of common cellular products (e.g., proteins, nucleic acids, and phospholipids) or take part in harvesting energy from nutrients (e.g., by the conversion of glucose and oxygen into carbon dioxide and water during cellular respiration; see Chapter 12). Other enzymes are present only in a particular type of cell because they catalyze chemical reactions unique to that cell type (e.g., the enzymes in neurons that convert tyrosine into dopamine, a neurotransmitter). A particularly important class of enzymes in cell biology are those used to communicate information about the internal and external state of the cell (signaling enzymes). This group of enzymes includes kinases that attach phosphate groups to proteins and lipids and phosphatases that remove them, as well as large and small GTP-hydrolyzing proteins. Other critical enzymes in cell biology function in controlling gene expression by causing structural changes in the nuclear DNA. These enzymes include topoisomerases, polymerases, and chromatin-modifying enzymes that can acetylate, de-acetylate, or methylate DNA-binding histones. Although most enzymes are located within cells, some are secreted and function at extracellular sites, such as the blood (clotting enzymes), the digestive tract (hydrolyases that digest dietary macromolecules), or even outside the organism (e.g., toxic enzymes in the venom of poisonous snakes).

Like all *catalysts* (see Chapter 2), enzymes increase the rate of a reaction, but they do not affect the extent of a reaction, which is determined by the change in free energy (ΔG) between reactants and products, and they are not themselves permanently changed as a consequence of the reaction they catalyze. Enzymes increase the reaction rate by lowering the energy of the *transition state*, and therefore the *activation energy* required to reach it (Figure 3-23). In the test tube, catalysts such as charcoal and platinum facilitate reactions, but usually only at high temperatures or pressures, at extremes of high or low pH, or in organic solvents. Within cells, however, enzymes must function effectively in an aqueous environment at 37 °C and 1 atmosphere of pressure and at physiological pH values, usually 6.5–7.5 but sometimes lower. Enzymes exhibit immense catalytic power, in some cases accelerating the rates of reactions to $10^6 - 10^{12}$ times those of the corresponding uncatalyzed reactions under otherwise similar conditions.

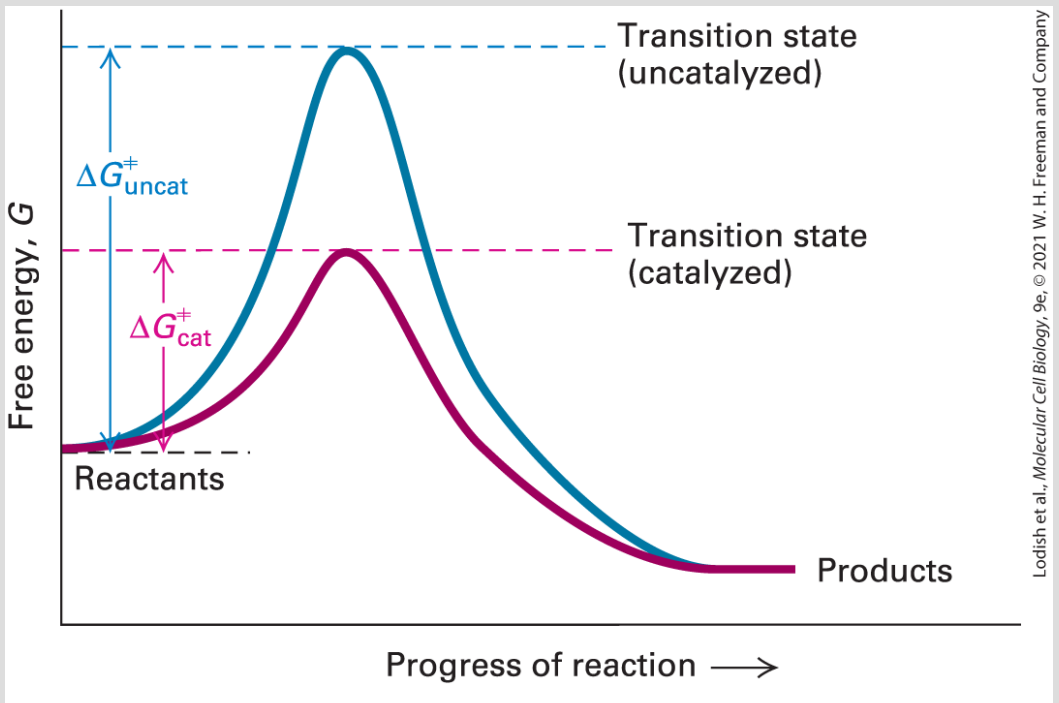


FIGURE 3-23 Effect of an enzyme on the activation energy of a chemical reaction. This hypothetical reaction pathway depicts the changes in free energy, *G*, as a reaction proceeds. A reaction will take place spontaneously only if the total *G* of the products is less than that of the reactants (negative ΔG). However, all chemical reactions proceed through one or more high-energy transition states, and the rate of a reaction is inversely proportional to the activation energy ($\Delta G^{\ddagger}$), which is the difference in free energy between the reactants and the transition state (highest point along the pathway). Enzymes and other catalysts accelerate the rate of a reaction by reducing the free energy of the transition state and thus $\Delta G^{\ddagger}$.

An Enzyme's Active Site Binds Substrates and Carries Out Catalysis

Certain amino acids of an enzyme are particularly important in determining its specificity and catalytic power. In the native

conformation of an enzyme, critically important amino acids (which usually come from different parts of the linear sequence of the polypeptide) are brought into proximity, forming a cleft in the enzyme surface called the **active site** (Figure 3-24). An active site usually makes up only a small part of the total protein; the remaining part is involved in the folding of the polypeptide, regulation of the active site, and interactions with other molecules.

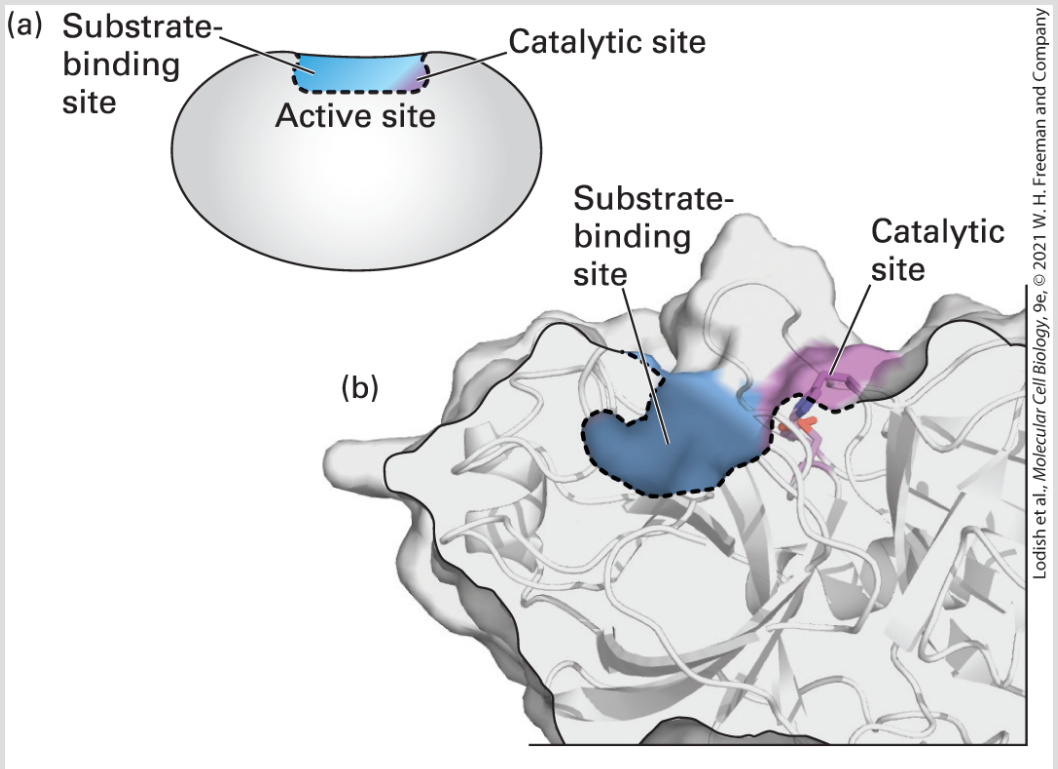


FIGURE 3-24 Active site of the enzyme trypsin. (a) An enzyme's active site (outlined by dashed line) is composed of a substrate-binding site (blue), which binds specifically to a substrate, and a catalytic site (purple), which carries out catalysis. (b) A surface-and-ribbon representation of a portion of the serine protease trypsin. Clearly visible are the active-site cleft containing the catalytic site (purple, includes the key catalytic triad of Ser-195, Asp-102, and His-57, see also Figure 3-28) and a portion of the substrate-binding site called the *side-chain-specificity binding pocket* (blue).

[Data from B. Sandler, M. Murakami, and J. Clardy, 1998, *J. Am. Chem. Soc.* **120**:595–596, PDB ID 1aq7.]

An active site consists of two functionally important regions: the *substrate-binding site,* which recognizes and binds the substrate or substrates, and the *catalytic site,* which carries out the chemical reaction once the substrate has bound. The catalytic groups in the catalytic site are amino acid side chains and backbone carboxyl and amino groups. In some enzymes, the catalytic and substrate-binding sites overlap; in others, the two regions are spatially distinct.

The substrate-binding site is responsible for the remarkable specificity of enzymes. An alteration in the structure of an enzyme's substrate by only one or a few atoms, or a subtle change in the geometry (e.g., stereochemistry) of the substrate, can result in a variant molecule that is no longer a substrate of the enzyme. As with the specificity of antibodies for antigens described above, the specificity of enzymes for substrates is a consequence of the precise molecular complementarity between an enzyme's substrate-binding site and the substrate. Usually only one or a few substrates can fit precisely into a binding site so that catalysis takes place.

The idea that substrates might bind to enzymes in the manner of a key fitting into a lock was first suggested by Emil Fischer in 1894. In the late 1940s Linus Pauling proposed that enzymes could facilitate catalysis by binding more tightly to the transition state of the

reaction than to the substrate itself, and thus lower the activation energy $\Delta G^{\ddagger}$ by stabilizing the transition state (see Figure 3-23). Daniel Koshland in 1958 proposed a variation of the lock-and-key model, called *induced fit*. The induced fit model posits that the substrate-binding site is not rigid, as a lock is, but flexible, and that substrate binding induces the enzyme to change shape and consequently bind most strongly to the transition state, thereby optimizing catalysis. As early as 1913, Leonor Michaelis and Maud Leonora Menten provided crucial evidence supporting the enzyme-substrate-binding hypothesis. They showed that the rate of an enzymatic reaction was proportional to the substrate concentration at low substrate concentrations, but that as substrate concentrations increased, the rate reached a plateau, or **V_{max} (maximal velocity)**, and became independent of the substrate concentration. The value of V_{max} was directly proportional to the amount of enzyme present in the reaction mixture (Figure 3-25).

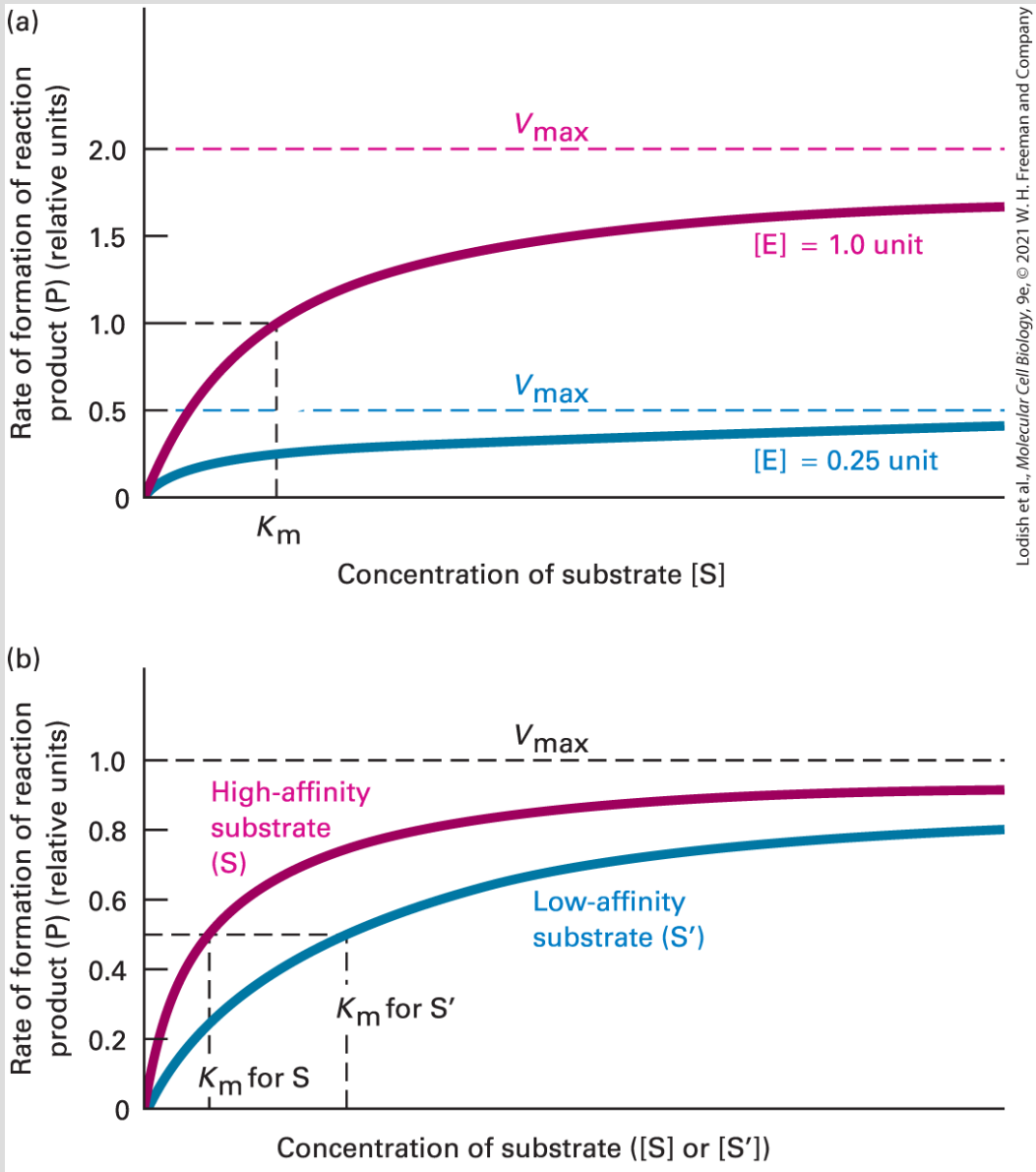
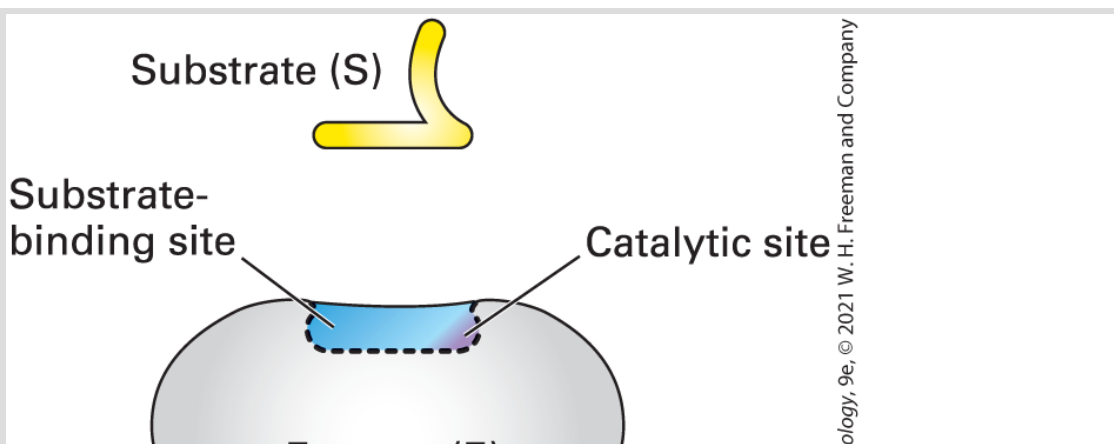


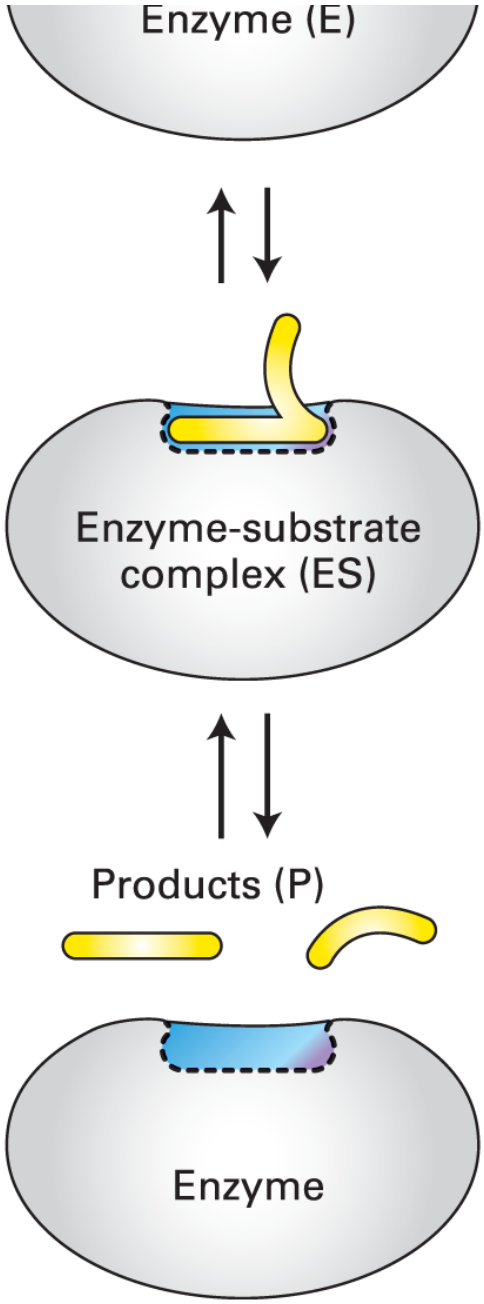
FIGURE 3-25 K_m and V_{max} for an enzyme-catalyzed reaction. K_m and V_{max} are determined by analyzing how the initial reaction rate depends on substrate concentration. The shape of these hypothetical kinetic curves is characteristic of a simple enzyme-catalyzed reaction in which one substrate (S) is converted into product (P). The initial reaction velocity is measured immediately after addition of enzyme to substrate, before the substrate concentration changes appreciably. (a) Plots of initial reaction velocity at two different concentrations of enzyme [E] as a function of substrate concentration [S]. The [S] that yields a half-maximal reaction rate is the Michaelis constant K_m, a measure of the affinity of E for turning S into P. Quadrupling the enzyme concentration causes

a proportional increase in the reaction rate, so the maximal velocity $V_{\max}$ is quadrupled; K_{m}, however, is unaltered. (b) Plots of initial reaction velocity versus substrate concentration with a substrate S for which the enzyme has a high affinity and with a substrate S' for which the enzyme has a lower affinity. Note that $V_{\max}$ is the same with both substrates because [E] is the same, but that K_{m} is higher for S', the low-affinity substrate.

Michaelis and Menten deduced that these characteristics were due to the binding of substrate molecules (S) to a fixed and limited number of sites on the enzymes (E), and they called the bound species the *enzyme-substrate (ES) complex*. At high concentrations of substrate, all the binding sites on the enzymes have substrate bound, and the substrate-binding sites are said to be *saturated* with substrate — no additional binding to active sites is possible, and the maximal velocity of the reaction is achieved. Michaelis and Menten proposed that the ES complex is in equilibrium with the unbound enzyme and substrate and is an intermediate step followed by the conversion of substrate to product (P) (Figure 3-26):

$$\mathrm{E} + \mathrm{S} \rightleftharpoons \mathrm{ES} \rightarrow \mathrm{E} + \mathrm{P}$$





Lodish et al., *Molecular Cell Bi*

FIGURE 3-26 Schematic model of an enzyme's reaction mechanism. Enzyme kinetics suggest that an enzyme (E) binds substrate molecules (S) at a fixed and limited number of sites — the enzyme's active sites. The bound species is known as an *enzyme-substrate (ES) complex*. The ES complex is in equilibrium with the unbound enzyme and substrate (double arrows) and is an intermediate step in the conversion of substrate to products (P).

In principle, all enzymatic reactions are reversible as can be seen in the glycolytic pathway (see Figure 12-3). However, because of the conditions under which they proceed or the type of chemical reaction catalyzed, many other enzymatic reactions are, in practice, virtually irreversible. The rate V_0 of formation of product at a particular substrate concentration [S] is given by what is now called the *Michaelis–Menten equation*:

$$V_0 = V_{\max} \frac{[\mathrm{S}]}{[\mathrm{S}] + K_{\mathrm{m}}} \tag{3-1}$$

where the **Michaelis constant, K_{m},** a measure of the affinity of an enzyme for its substrate, is the substrate concentration that yields a half-maximal reaction rate (i.e., $\frac{1}{2}V_{\max}$ in Figure 3-25). The K_{m} is somewhat similar in nature, but not identical, to the dissociation constant, K_{d} (see Section 2.3). The smaller the value of K_{m}, the more effective the enzyme is at making product from dilute solutions of substrate, and the lower the substrate concentration needed to reach half-maximal velocity. The smaller the K_{d}, the lower the ligand concentration needed to reach 50 percent of binding. The concentrations of the various small molecules in a cell vary widely, as do the K_{m} values for the different enzymes that act on them. A good rule of thumb is that the intracellular concentration of a substrate is often approximately the same as, or

somewhat greater than, the K_m value of the enzyme to which it binds.

The rates of reaction at substrate saturation vary enormously among enzymes. The maximum number of substrate molecules converted to product at a single enzyme active site per second, called the *turnover number*, can be less than one for very slow enzymes. The turnover number for carbonic anhydrase, one of the fastest enzymes, is 6×10^5 molecules per second.

Many enzymes catalyze the conversion of substrates to products by dividing the process into multiple, discrete chemical reactions, in which the product of one reaction is the substrate for the subsequent reaction. These sequential reactions generate multiple, distinct enzyme-substrate complexes (ES, ES', ES'', etc.) prior to the final release of the products:

$$\mathrm{E} + \mathrm{S} \rightleftharpoons \mathrm{ES} \rightleftharpoons \mathrm{ES}' \rightleftharpoons \mathrm{ES}'' \rightleftharpoons \ldots \mathrm{E} + \mathrm{P}$$

The energy profiles for such multistep reactions contain multiple hills and valleys (Figure 3-27). Methods have been developed to trap the intermediates in such reactions to learn more about the details of how enzymes catalyze reactions.

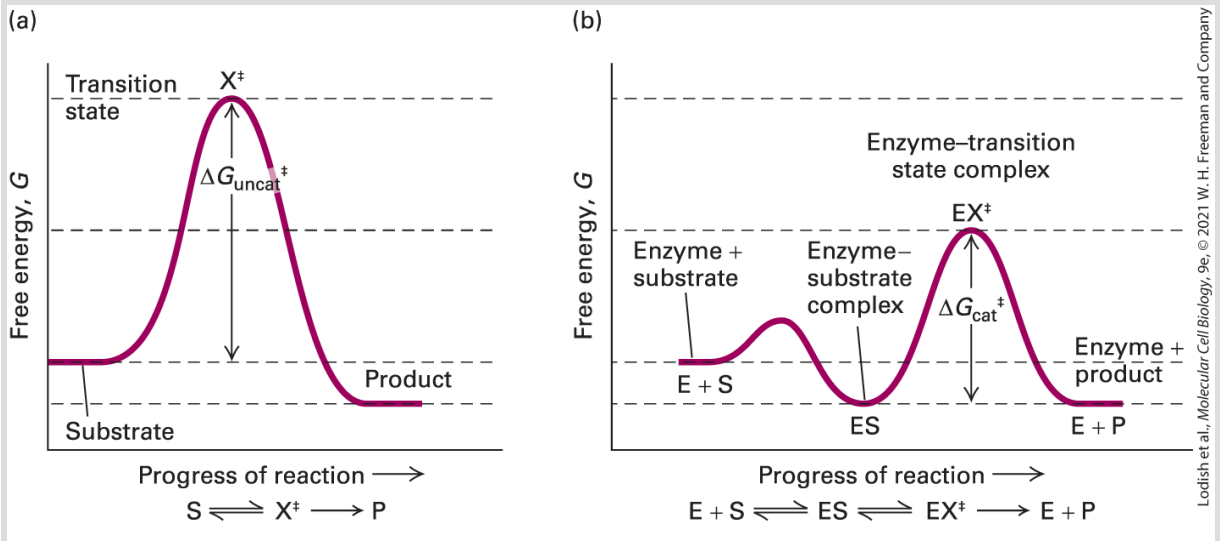
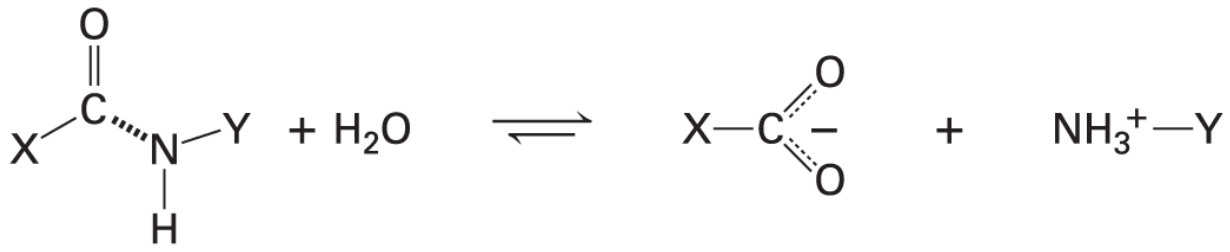


FIGURE 3-27 Free-energy reaction profiles of uncatalyzed and multistep enzyme-catalyzed reactions. (a) The free-energy reaction profile of a hypothetical simple uncatalyzed reaction converting substrate (S) to product (P) via a single high-energy transition state. (b) Many enzymes catalyze such reactions by dividing the process into multiple discrete steps, in this case, the initial formation of an ES complex followed by conversion via a single transition state ($EX^{\ddagger}$) to the free enzyme (E) and P. The activation energy for each of these steps is significantly less than the activation energy for the uncatalyzed reaction; thus the enzyme dramatically enhances the reaction rate.

Serine Proteases Demonstrate How an Enzyme's Active Site Works

Serine proteases, a large family of protein-cleaving, or proteolytic, enzymes, are used throughout the biological world — to digest meals (the pancreatic enzymes trypsin, chymotrypsin, and elastase), to control blood clotting (the enzyme thrombin), even to help silk moths chew their way out of their cocoons (cocoonase). This class of enzymes usefully illustrates how an enzyme's substrate-binding site

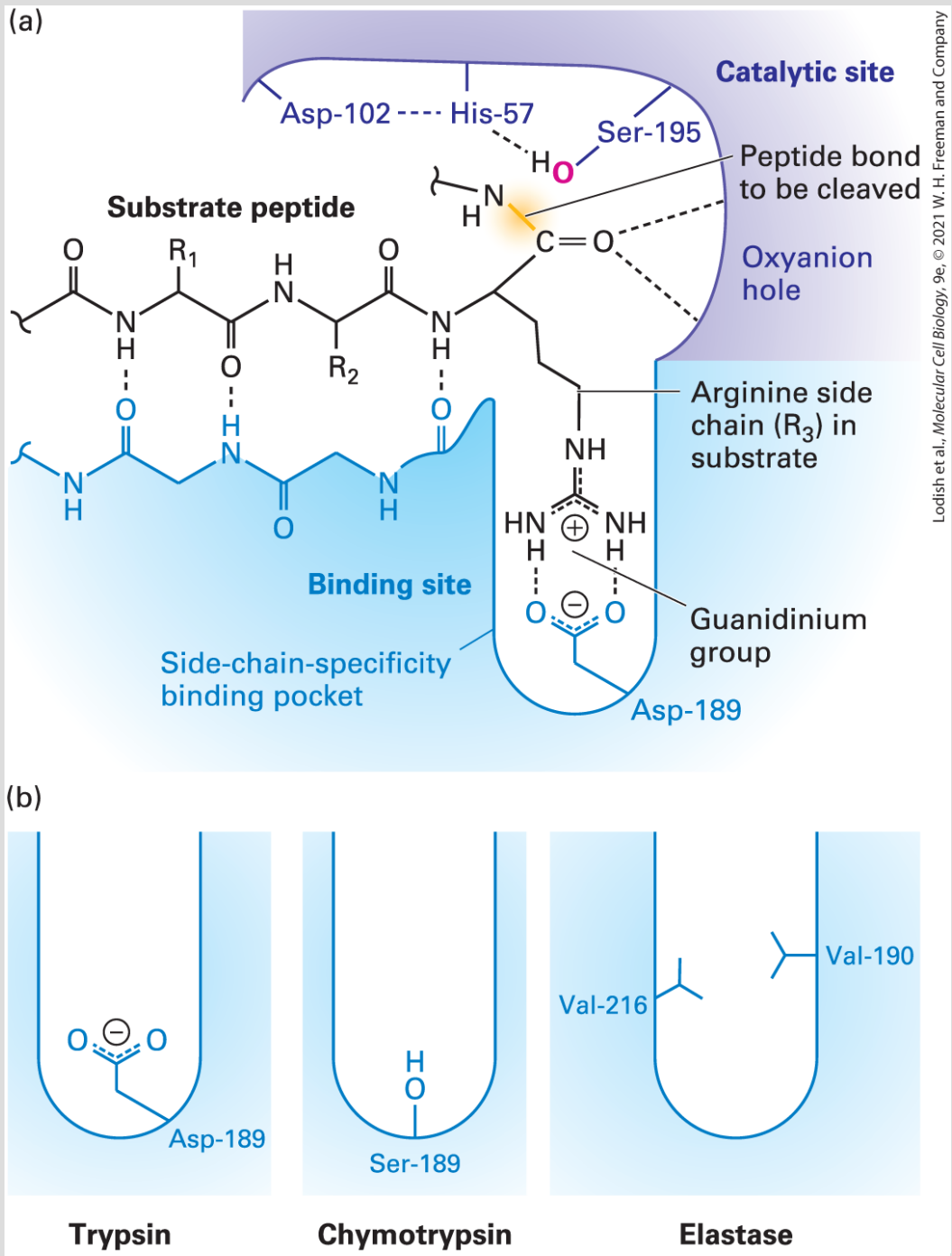
and catalytic site cooperate in multistep reactions to convert substrate to product. Here we consider how trypsin and its two evolutionarily closely related pancreatic proteases, chymotrypsin and elastase, catalyze cleavage of a peptide bond in a polypeptide substrate:



where X is the part of the protein on the N-terminal side of the peptide bond to be cleaved and Y is the part of the protein on the C-terminal side. We first consider how serine proteases bind specifically to their substrates and then show in detail how catalysis takes place.

Figure 3-28a shows how a substrate polypeptide binds to the substrate-binding site in the active site of trypsin. There are two key binding interactions. First, the substrate (black polypeptide backbone) and enzyme (blue polypeptide backbone) form hydrogen bonds that resemble those of a β sheet. Second, a key side chain of the substrate that determines which peptide in the substrate is to be cleaved extends into the enzyme's *side-chain-specificity binding pocket*, at the bottom of which resides the negatively charged side chain of trypsin's Asp-189. Trypsin has a marked preference for hydrolyzing

substrates at the carboxyl (C=O) side of an amino acid with a long, positively charged side chain (arginine or lysine) because the side chain is stabilized in the enzyme's side-chain-specificity binding pocket by the negative Asp-189.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 3-28 Substrate binding in the active site of trypsin-like serine proteases. (a) The active site of trypsin (purple and blue molecule) with a bound substrate (black molecule). The substrate forms a two-stranded β sheet with trypsin's substrate-binding site, and the side chain of an arginine (R_3) in the substrate is bound in the side-chain-specificity binding pocket of the binding site. Its positively charged guanidinium group is stabilized by the

negative charge on the side chain of the enzyme's Asp-189. This binding aligns the peptide bond of the arginine appropriately for hydrolysis catalyzed by the enzyme's active-site catalytic triad (side chains of Ser-195, His-57, and Asp-102). (b) The amino acids lining the side-chain-specificity binding pocket determine its shape and charge and thus its binding properties. Trypsin accommodates the positively charged side chains of arginine and lysine; chymotrypsin, large, hydrophobic side chains such as phenylalanine; and elastase, small side chains such as glycine and alanine.

[Panel (a) data from J. J. Perona and C. S. Craik, 1997, *J. Biol. Chem.* **272**:29987–29990.]

Slight differences in the structures of otherwise similar binding pockets help explain the differing substrate specificities of two serine proteases related to trypsin: chymotrypsin prefers large aromatic groups (as in Phe, Tyr, Trp), and elastase prefers the small side chains of Gly and Ala (Figure 3-28b). The uncharged Ser-189 in chymotrypsin allows large, uncharged, hydrophobic side chains to bind stably in the binding pocket. The specificity of elastase is influenced by the replacement of glycines in the sides of the binding pocket in trypsin with the branched aliphatic side chains of valines (Val-216 and Val-190), which obstruct the binding pocket (Figure 3-28b). As a consequence, large side chains in substrates are prevented from fitting into the binding pocket of elastase, whereas substrates with the short alanine or glycine side chains at this position can bind well and be subject to subsequent cleavage.

In the catalytic site, all three enzymes use the hydroxyl group on the side chain of a serine at position 195 to catalyze the hydrolysis of

peptide bonds in substrate proteins. A catalytic triad formed by the three side chains of Ser-195, His-57, and Asp-102 participates in what is essentially a two-step hydrolysis reaction. Polypeptide substrate binding to the binding site on the enzyme aligns the peptide bond to be cleaved with the enzyme's active-site catalytic triad. Figure 3-29 shows how the catalytic triad cooperates in breaking the peptide bond, with Asp-102 and His-57 supporting the attack of the hydroxyl oxygen of Ser-195 on the carbonyl carbon in the substrate. This attack initially forms an unstable transition state with four groups attached to the carbon (tetrahedral intermediate). Breaking of the C—N peptide bond then releases one part of the substrate protein (NH_3—Y), while the other part remains covalently attached to the enzyme via an ester bond to the serine's oxygen, forming a relatively stable acyl enzyme intermediate. The subsequent replacement of this oxygen with one from water, in a reaction involving another unstable tetrahedral intermediate, leads to release of the final product (X—COOH). The tetrahedral intermediate transition states are partially stabilized by hydrogen bonding with the enzyme's backbone amino groups in what is called the *oxyanion hole*. The large family of serine proteases and related enzymes, all of which have an active-site serine, illustrates how an efficient reaction mechanism is used over and over by distinct enzymes to catalyze similar reactions.

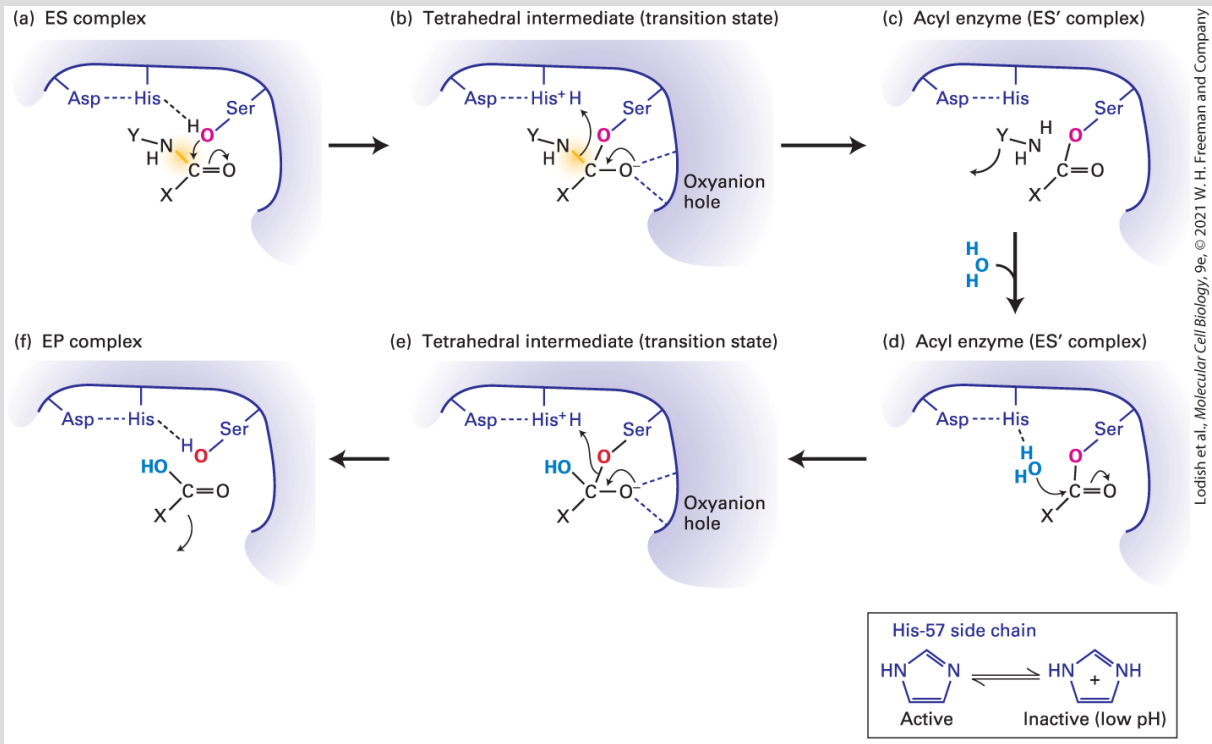


FIGURE 3-29 Mechanism of serine protease–mediated hydrolysis of peptide bonds. The catalytic triad of Ser-195, His-57, and Asp-102 in the active sites of serine proteases employs a multistep mechanism to hydrolyze peptide bonds in target proteins. (a) After a polypeptide substrate binds to the active site (see Figure 3-27), forming an ES complex, the hydroxyl oxygen of Ser-195 attacks the carbonyl carbon of the substrate's targeted peptide bond (yellow). Movements of electrons are indicated by arrows. (b) This attack results in the formation of a transition state called the *tetrahedral intermediate*, in which the negative charge on the substrate's oxygen is stabilized by hydrogen bonds formed with the enzyme's *oxyanion hole*. (c) Additional electron movements result in the breaking of the peptide bond, release of one of the reaction products (NH_2—Y), and formation of the acyl enzyme (ES′ complex). (d) An oxygen from a solvent water molecule then attacks the carbonyl carbon of the acyl enzyme. (e) This attack results in the formation of a second tetrahedral intermediate. (f) Additional electron movements result in the breaking of the Ser-195–substrate bond (formation of the EP complex) and release of the final reaction product (X—COOH). The side chain of His-57, which is held in the proper orientation by hydrogen bonding to the side chain of Asp-102, facilitates catalysis by withdrawing and donating protons throughout the reaction (*inset*). If the pH is too low and the side chain of His-57 is protonated, it cannot participate in catalysis and the enzyme is inactive.

The serine protease mechanism illustrates several key features of enzymatic catalysis. First, enzyme catalytic sites have evolved to stabilize the binding of a transition state, thus lowering the activation energy and accelerating the overall reaction. Second, multiple side chains, together with the polypeptide backbone, carefully organized in three dimensions, work together to chemically transform substrate into product, often by multistep reactions.

Third, acid-base catalysis mediated by one or more amino acid side chains is often used by enzymes, as when the imidazole group of His-57 in serine proteases acts as a base to remove the hydrogen from Ser-195's hydroxyl group. As a consequence, only a particular ionization state (protonated or nonprotonated) of one or more amino acid side chains in the catalytic site may be compatible with catalysis, and thus the enzyme's activity may be pH dependent. For example, the imidazole of His-57 in serine proteases, whose pK_a is ~ 6.8, can help the Ser-195 hydroxyl attack the substrate only if it is not protonated. Thus the activity of the protease is low at pH < 6.8, at which the imidazole is protonated, and the shape of the pH activity profile in the pH range 4–8 matches the titration of the His-57 side chain, which is governed by the Henderson-Hasselbalch equation, with an inflection near pH 6.8 (see chymotrypsin data in Figure 3-30 and see Chapter 2). The activity drops at higher pH values, generating a bell-shaped activity curve, because the proper folding of the protein is disrupted when the amino group at the protein's

amino terminus (pK_a ~9) is deprotonated; the conformation near the active site changes as a consequence.

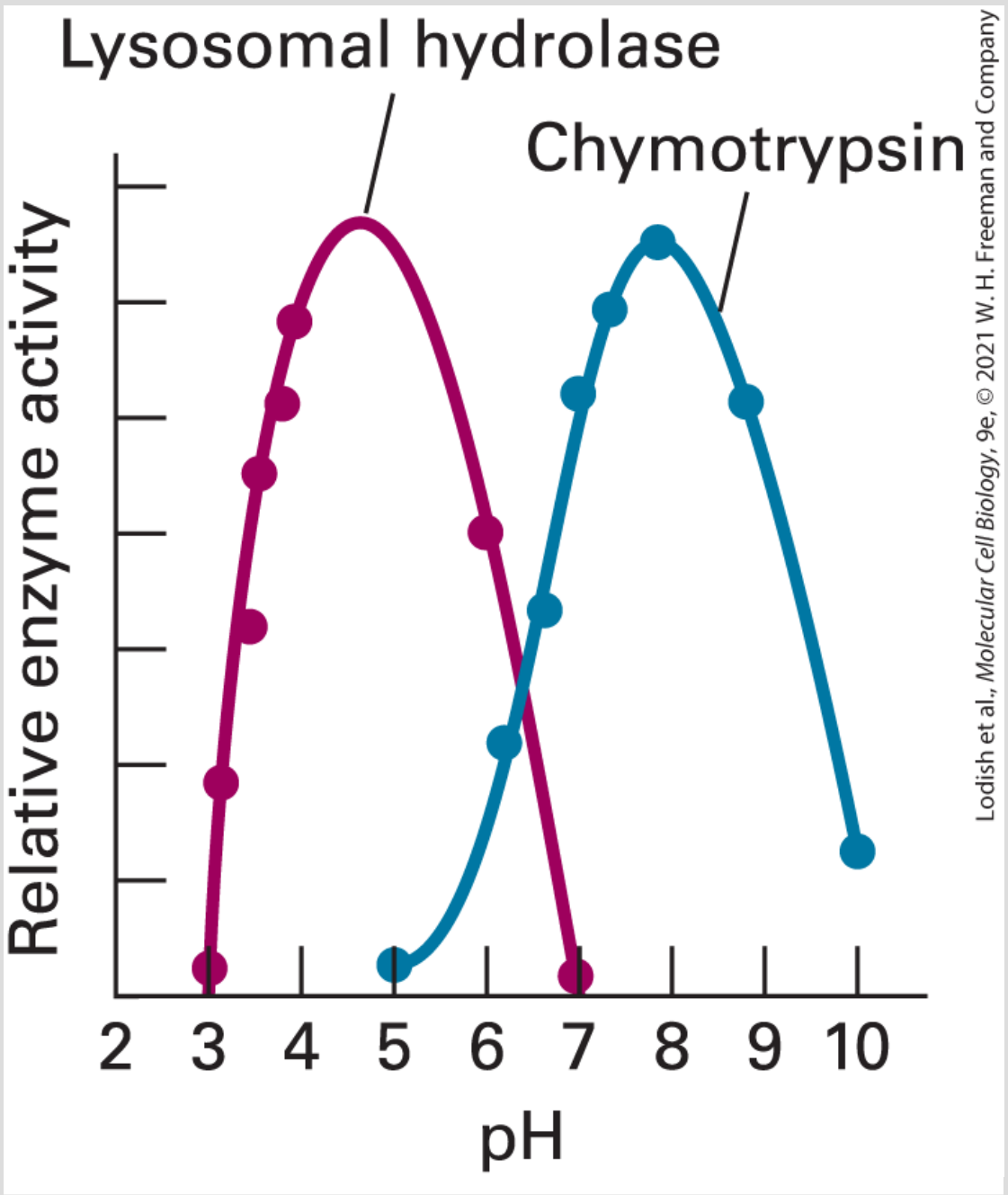


FIGURE 3-30 The pH dependence of enzyme activity. In some cases, ionizable (pH-titratable) groups in enzyme active sites or elsewhere in enzymes must be either protonated or deprotonated to permit proper substrate binding or catalysis, or to

permit the enzyme to adopt its correct conformation. Measurement of enzyme activity as a function of pH can be used to identify the pK_a's of these groups. The pancreatic serine proteases, such as chymotrypsin, exhibit maximum activity at around pH 8 because of titration of the active-site His-57 (i.e., the deprotonated imidazole side-chain of His-57, which is required for catalysis, $pK_a \sim 6.8$) and of the amino terminus of the protein (required for proper conformation, $pK_a \sim 9$). Many lysosomal hydrolases have evolved to exhibit a lower pH optimum (~ 4.5) to match the low internal pH in the lysosomes in which they function.

[Data from P. Lozano, T. De Diego, and J. L. Iborra, 1997, *Eur. J. Biochem*. **248**:80–85, and W. A. Judice et al., 2004, *Eur. J. Biochem*. **271**:1046–1053.]

The pH sensitivity of an enzyme's activity can be due to changes in the ionization of catalytic groups, groups that participate directly in substrate binding, or groups that influence the conformation of the protein. Pancreatic serine proteases evolved to function in the neutral or slightly basic conditions in the intestines; hence their pH optima are ~ 8. Proteases and other hydrolytic enzymes that function in acidic conditions must employ a different catalytic mechanism. This is the case for enzymes within the stomach (pH ~ 1), such as the protease pepsin, and for those within lysosomes (pH ~ 4.5), which play a key role in degrading macromolecules within cells (see the lysosomal hydrolase data in Figure 3-30). Indeed, lysosomal hydrolases, which degrade a wide variety of biomolecules (proteins, lipids, etc.), are relatively inactive at the pH in the cytosol (~ 7), which helps to protect a cell from self-digestion if these enzymes escape the confines of the membrane-bounded lysosome.

One key feature of enzymatic catalysis not seen in serine proteases but found in many other enzymes is a *cofactor* or *prosthetic group*. This helper group is a nonpolypeptide small molecule or ion (e.g., iron, zinc, copper, manganese) that is bound in the active site and plays an essential role in the reaction mechanism. Small organic prosthetic groups in enzymes are also called *coenzymes*. Some of these groups are chemically modified during the reaction and thus need to be replaced or regenerated after each reaction; others are not. Examples of coenzymes include NAD^+ (nicotinamide adenine dinucleotide), FAD (flavin adenine dinucleotide) (see Figure 2-33), and the heme groups that bind oxygen in hemoglobin or transfer electrons in some cytochromes (see Figure 12-17). Thus the chemical reactions catalyzed by enzymes are not restricted by the limited number of types of amino acids in polypeptide chains. Many of the vitamins [e.g., the B vitamins thiamine (B_1), riboflavin (B_2), niacin (B_3), and pyridoxine (B_6), as well as vitamin C], which cannot be synthesized in mammalian cells, function as, or are used to generate, coenzymes. That is why supplements of vitamins must be added to the liquid medium in which mammalian cells are grown in the laboratory (see Section 4.1).

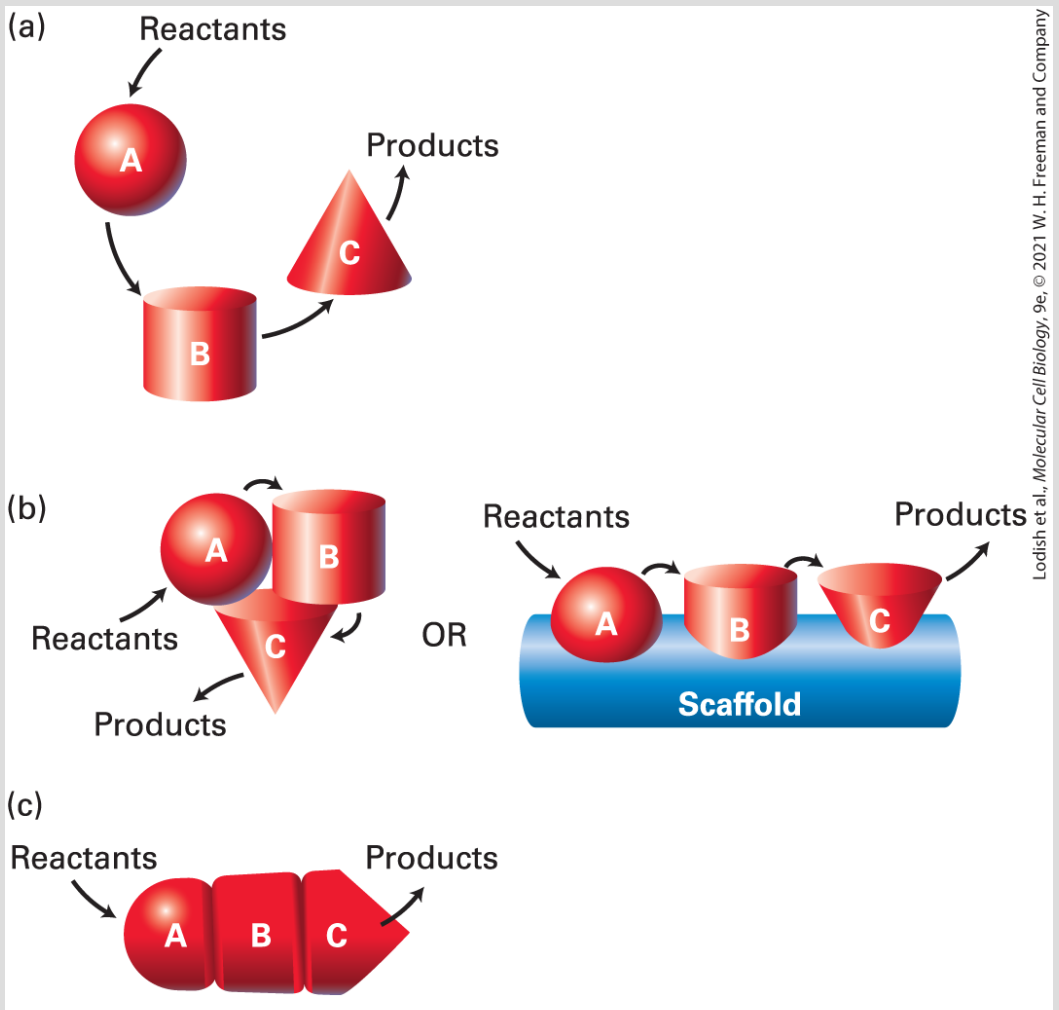
Small molecules that can bind to active sites and disrupt catalytic reactions are called *enzyme inhibitors*. Such inhibitors are useful tools for studying the roles of enzymes in cells and organisms. Inhibitors that bind directly to an enzyme's binding site and thus compete directly with the normal substrate are called *competitive*

inhibitors. Noncompetitive inhibitors are those that interfere with enzyme activity in other ways — for example, by binding to some other site on the enzyme and changing its conformation. Enzyme inhibitors complement the use of genetic mutations and a technique called *RNA interference (RNAi)* for probing an enzyme's function in cells (see Chapter 6). In all three approaches, the cellular consequences of disrupting an enzyme's activity can be used to deduce the normal function of the enzyme. The same approaches can be used to study the functions of nonenzymatic macromolecules. Interpreting the results of inhibitor studies can be complicated, however, if, as is often the case, the inhibitors disrupt the normal activity of more than one protein.

Inhibition of protein activity by small molecules is the basis for the action of many drugs as well as chemical warfare agents. Aspirin inhibits enzymes called *cyclooxygenases*, whose products can cause pain. Sarin and other nerve gases react with the active serine hydroxyl groups of both serine proteases and a related enzyme, acetylcholine esterase, which is a key enzyme in regulating nerve conduction (see Chapter 23). ■

Enzymes in a Common Pathway Are Often Physically Associated with One Another

Enzymes taking part in a common metabolic process (e.g., the degradation of glucose to pyruvate during glycolysis; see Chapter 12) are generally located in the same cellular compartment, be it in the cytosol, at a membrane, or within a particular organelle. Within this compartment, products from one reaction can move by diffusion to the next enzyme in the metabolic pathway. Diffusion, however, entails random movement and can be a slow, relatively inefficient process for moving molecules between enzymes (Figure 3-31a). To overcome this impediment, cells have evolved mechanisms for bringing enzymes in a common pathway into close proximity, a process called *metabolic coupling*.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 3-31 Assembly of enzymes into efficient multienzyme complexes. In the hypothetical reaction pathways illustrated here, the initial reactants are converted into final products by the sequential action of three enzymes: A, B, and C. (a) When the enzymes are free in solution, or even constrained within the same cellular compartment, the intermediates in the reaction sequence must diffuse from one enzyme to the next, an inherently slow process. (b) Diffusion is greatly reduced or eliminated when the individual enzymes associate into multisubunit complexes, either by themselves or with the aid of a scaffold protein. (c) The closest integration of different catalytic activities occurs when the enzymes are fused at the genetic level, becoming domains in a single polypeptide chain.

In the simplest such mechanism, polypeptides with different catalytic activities cluster closely together as subunits of a multimeric enzyme or assemble on a common *scaffold* that holds them together (Figure 3-31b). This arrangement allows the product of one reaction to be channeled directly to the next enzyme in the pathway. In some cases, independent proteins have been fused together at the genetic level to create a single multidomain, multifunctional enzyme (Figure 3-31c). Metabolic coupling usually involves large multiprotein complexes, as described earlier in this chapter.

KEY CONCEPTS OF SECTION 3.3

Protein Binding and Enzyme Catalysis

- A protein's function depends on its ability to bind other molecules, known as *ligands*. For example, antibodies bind to a group of ligands known as *antigens*, and enzymes bind to reactants called *substrates* that will be converted by chemical reactions into products.
- The specificity of a protein for a particular ligand refers to the preferential binding of one or a few closely related ligands. The affinity of a protein for a particular ligand refers to the strength of binding, usually expressed as the dissociation constant K_{d}.
- Proteins are able to bind to ligands because of molecular complementarity between the ligand-binding sites and the corresponding ligands.
- Enzymes are catalytic proteins that accelerate the rates of cellular reactions by lowering the activation energy and stabilizing transition-state intermediates (see Figure 3-23).
- An enzyme's active site, which is usually only a small part of the protein, comprises two functional parts: a substrate-binding site and a catalytic site. The substrate-binding site is responsible for the exquisite specificity of enzymes owing to its molecular complementarity with the substrate.
- The initial binding of a substrate (S) to an enzyme (E) results in the formation of an enzyme-substrate complex (ES), which then undergoes one or more reactions

catalyzed by the catalytic groups in the catalytic site until the final product (P) is formed.

- From plots of reaction rate versus substrate concentration, two characteristic parameters of an enzyme can be determined: the Michaelis constant, K_{m}, a rough measure of the enzyme's affinity for converting substrate into product, and the maximal velocity, $V_{\max}$, a measure of its catalytic power (see Figure 3-25).
- The rates of enzyme-catalyzed reactions vary enormously, with turnover numbers (numbers of substrate molecules converted to product at a single active site at substrate saturation) ranging from fewer than 1 to 6×10^5 molecules per second.
- Many enzymes catalyze the conversion of substrates to products by dividing the process into multiple discrete chemical reactions that involve multiple distinct enzyme-substrate complexes (ES', ES'' etc.).
- Serine proteases hydrolyze peptide bonds in substrate proteins using as catalytic groups the side chains of Ser-195, His-57, and Asp-102. Amino acids lining the side-chain-specificity binding pocket in the binding site of serine proteases determine the residue in a substrate protein whose peptide bond will be hydrolyzed and account for differences in protease specificity (e.g., trypsin versus chymotrypsin and elastase).
- Enzymes often use acid-base catalysis mediated by one or more amino acid side chains, such as the imidazole group of His-57 in serine proteases, to catalyze reactions. The pH dependence of protonation of catalytic groups $(\mathrm{p}K_{\mathrm{a}})$ is often reflected in the pH-rate profile of the enzyme's activity.
- Nonpolypeptide small molecules or ions, called *cofactors* or *prosthetic groups*, bind to the active sites of some enzymes and play an essential role in enzymatic catalysis. Small organic prosthetic groups in enzymes are also called *coenzymes*; many vitamins, which cannot be synthesized in higher animal cells, function as or are used to generate coenzymes.
- Enzymes in a common metabolic pathway are often located within the same cellular compartments and may be further associated as domains of a monomeric protein, subunits of a multimeric protein, or components of a protein complex assembled on a common scaffold (see Figure 3-31).

3.4 Regulating Protein Function

Most processes in cells do not take place independently of one another or at a constant rate. The activities of all proteins and other biomolecules are regulated to integrate their functions for optimal performance for survival. For example, the catalytic activity of enzymes is regulated so that the amount of reaction product is just sufficient to meet the needs of the cell. As a result, the steady-state concentrations of substrates and products may vary depending on cellular conditions. Regulation of nonenzymatic proteins — the opening or closing of membrane channels or the assembly of a macromolecular complex, for example — is also essential.

In general, there are three ways to regulate protein activity. First, cells can increase or decrease the steady-state level of the protein by altering its rate of synthesis, its rate of degradation, or both. Second, cells can change the intrinsic activity, as distinct from the amount, of the protein. For example, through noncovalent and covalent interactions, cells can change the affinity of substrate binding, or the fraction of time the protein is in an active versus an inactive conformation. Third, there can be a change in the location or the concentration within the cell of the protein itself, of the target of the protein's activity (e.g., an enzyme's substrate), or of some other molecule required for the protein's activity (e.g., an enzyme's cofactor). All three types of regulation play essential roles in the lives and functions of cells. In this section, we first discuss

mechanisms for regulating the amount of a protein, then turn to noncovalent and covalent interactions that regulate protein activity.

Regulated Synthesis and Degradation of Proteins Is a Fundamental Property of Cells

The rate of synthesis of a protein is determined by the rate at which the DNA encoding the protein is converted to mRNA (transcription), the steady-state amount of the active mRNA in the cell, and the rate at which the mRNA is converted into newly synthesized protein (translation). These important processes are described in detail in Chapter 5.

The life spans of intracellular proteins vary from as short as a few minutes for mitotic cyclins, which help regulate passage through the mitotic stage of cell division (see Chapter 19), to as long as the age of an organism for proteins in the lens of the eye. Protein life span is controlled primarily by regulated protein degradation.

Protein degradation plays two especially important roles in the cell. First, it removes proteins that are potentially toxic, improperly folded or assembled, or damaged — including the products of mutated genes and proteins damaged by chemically active cell metabolites or stress such as heat shock. Despite the existence of

chaperone-mediated protein folding, some newly made proteins are rapidly degraded because they are misfolded. This degradation might be necessary due to failure of timely engagement of the necessary chaperones to guide the folding of the proteins or due to their defective assembly into complexes. Most other proteins are degraded more slowly, undergoing about 1–2 percent degradation per hour in mammalian cells. Second, the controlled destruction of otherwise normal proteins, along with controlled rates of synthesis, provides a powerful mechanism — called **proteostasis** — for maintaining proteins and their activities at appropriate levels and for making rapid adjustments in these levels in response to changing conditions.

Eukaryotic cells have several pathways for degrading proteins. One major pathway is degradation by enzymes within lysosomes. The acidic interior (pH $\sim$4.5) of these membrane-limited organelles is filled with a host of hydrolytic enzymes. Lysosomal degradation is directed primarily toward aged or defective organelles of the cell — a process called *autophagy* (see Chapter 14) — and toward extracellular proteins taken up by the cell. Lysosomes will be discussed at length in later chapters. Here we focus on another important degradation pathway: degradation of cytoplasmic proteins by proteasomes, which can account for up to 90 percent of the protein degradation in mammalian cells.

The Proteasome Is a Molecular Machine Used to Degrade Proteins

Proteasomes are very large, multisubunit, protein-degrading molecular machines that influence many different cellular functions, including the cell cycle (see Chapter 19); transcription and DNA repair (see Chapter 5); programmed cell death, or *apoptosis* (see Chapter 22); recognition of and response to infection by foreign organisms (see Chapter 24); and removal of misfolded proteins. There are three key steps in the degradation of a protein by proteasomes. First, the protein is tagged to target it for proteasomal degradation. It is possible for cells to control the tagging of any given protein and thus regulate the rate at which that protein is degraded by proteasomes. Second, the proteasome binds to the targeted protein via the tag and unfolds the protein as it is transferred into an internal chamber. Third, protein-cutting subunits of the proteasome within the chamber degrade the target protein into small peptides, which are released into the cytosol for further processing. We will now describe in detail the structure of proteasomes and the mechanisms underlying these three steps.

There are approximately 30,000 proteasomes in a typical mammalian cell. Proteasomes consist of roughly 60 protein subunits and have a mass of about 2.4×10^6 Da. Proteasomes have a cylindrical, barrel-like catalytic core (Figure 3-32a), called the *20S proteasome* (where *S* is a Svedberg unit based on the sedimentation

properties of the particle and is proportional to its size), which is approximately 14.8 nm tall and 11.3 nm in diameter. Bound to one or both ends of this core are either one or two 19S *regulatory particle (RP)* complexes that regulate the activity of the 20S catalytic core. When the core and one or two regulatory particles are combined, they are referred to collectively as the *26S complex*, even though the two-cap-containing complex is larger (30S). A 19S regulatory particle has 19 protein subunits, six of which can hydrolyze ATP (i.e., they are AAA-type ATPases) to provide the energy needed to unfold protein substrates and selectively transfer them into the inner chamber of the proteasome's catalytic core. Genetic studies in yeast have shown that cells cannot survive without functional proteasomes. In fact, proper proteasomal activity is so important that cells will expend as much as 30 percent of the energy needed to synthesize a protein to degrade it in a proteasome. Activities of individual proteasomes can be regulated by phosphorylation — for example, to coordinate proteasome activity with the cell cycle (see Chapter 19). In addition, the amount of proteasomes in a cell is regulated by complex feedback pathways to ensure that enough are present for protein degradation.

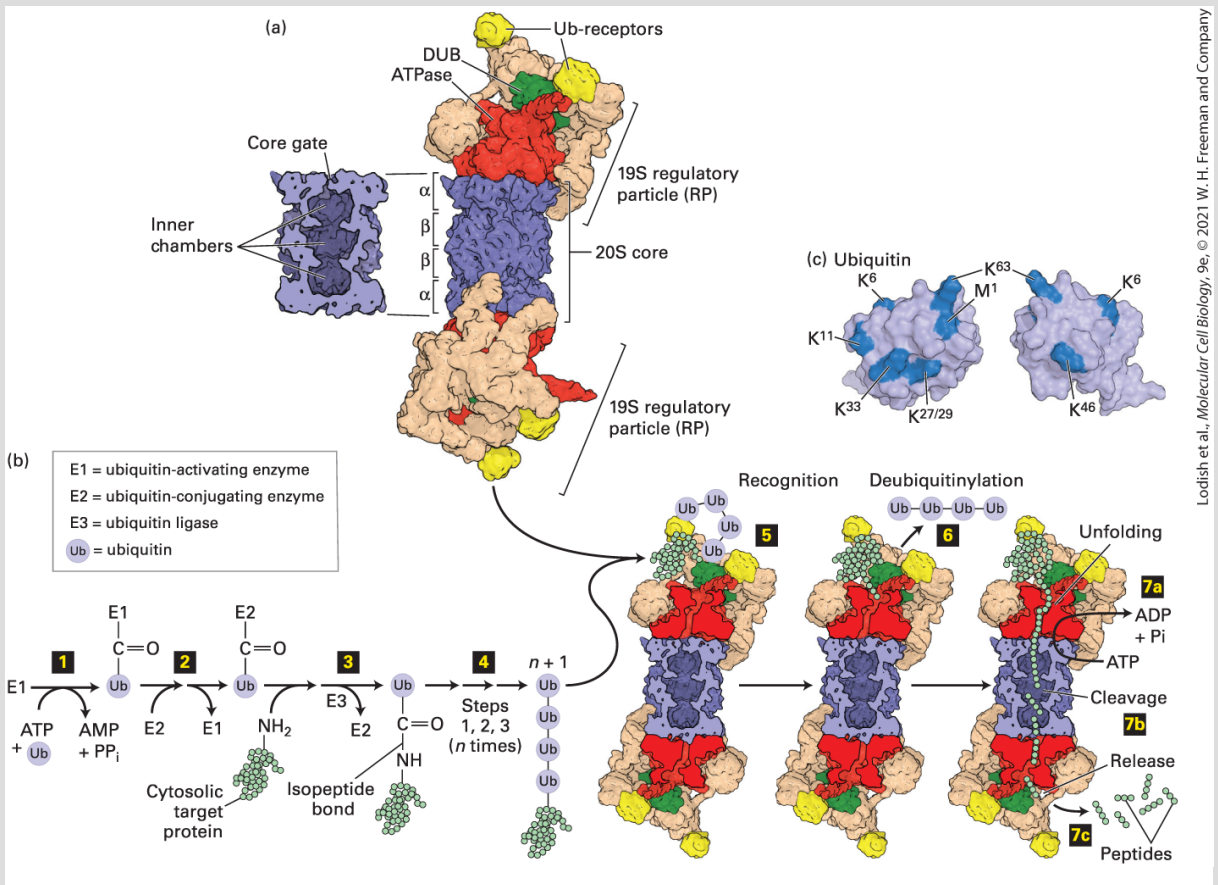


FIGURE 3-32 Ubiquitin- and proteasome-mediated proteolysis. (a) *Right:* The 26S proteasome has a cylindrical structure with a 19S regulatory particle (RP) at one or both ends of the 20S core particle. The 19 different subunits of the 19S RP (shown in multiple colors) include six AAA-ATPase subunits (Rpt1–6, red), which assemble into a heterohexameric ring; three ubiquitin (Ub) receptors (Rpn1, Rpn10, and Rpn13, yellow); and a deubiquitinase enzyme (DUB, Rpn11, green), which forms a heterodimer with its evolutionarily related counterpart Rpn8. Moreover, the 19S RP contains scaffolding and other proteins (tan). The two 19S RPs shown are facing in opposite directions relative to the plane of the page. The 20S core consists of four stacked heptameric rings ($\sim 110\ \text{Å diameter} \times 160\ \text{Å long}$), each containing either α (outer rings) or β (inner rings) subunits (blue). *Left:* Cutaway view of the 20S core, showing the inner chambers and regulated core gate. Proteolysis occurs within the central inner chamber of the core formed by the β rings. (b) Proteins are targeted for proteasomal degradation by polyubiquitinylation. Enzyme E1 is activated by the ATP-dependent attachment of a ubiquitin (Ub) molecule (step **1**) and then transfers this Ub molecule to a cysteine residue in E2 (step **2**). Ubiquitin ligase (E3) transfers the bound Ub molecule on E2 to the side-chain $-NH_2$ of a lysine residue in a target protein, forming an isopeptide bond (step **3**).

Additional Ub molecules are added to the Ub-modified target protein via isopeptide bonds to the previously added Ub by repeating steps **1** – **3**, forming a polyubiquitin chain (step **4**). The polyubiquitinylated target is recognized by Ub receptors in the proteasome's 19S RP (step **5**). It is followed by a series of coordinated steps accompanied by conformational changes in the 19S RP. The Ub groups are removed by the deubiquitinase enzyme (step **6**) as the target protein is moved into the proper position near the six protein (hexameric) ATPase. In step **7** ATP hydrolysis enables the hexameric ATPase subunits (red) to unfold the substrate (grabbing and pulling the polypeptide through its narrow central pore) and transfer the unfolded protein via the pore through the now opened core gate into the proteolysis chamber in the 20S core (step **7a**), in some cases coordinately with step **6**, and the protein is cleaved into short peptide digestion fragments (step **7b**) that are then released (step **7c**). (c) Space-filling model of ubiquitin showing in blue lysines (K) and amino terminus (M1) to which additional ubiquitins may be covalently added to form a polyubiquitin chain (additional discussion of polyubiquitin chains can be found later in this chapter.

[Part (a) Courtesy of Antje Aufderheide and Friedrich Foerster, data from P. Unverdorben et al., 2014, *Proc. Nat'l. Acad. Sci. USA* **111**(15):5544–5549, PDB ID 4cr2. Part (c) From R. Yau and M. Rape, 2016, *Nat. Cell Biol.* **18**:579–586.]

The catalytic core of the 20S proteasome comprises two inner rings of seven β subunits each that surround a ∼1.7-nm-diameter inner chamber. Each ring contains three proteolytic active sites facing toward the inner chamber. Two outer rings of seven α subunits each limit substrate access (see Figure 3-32a) via an entry pore called the *core gate*, whose opening is controlled by substrate binding to the 19S regulatory particle. Proteasomes can degrade most proteins thoroughly because the three active sites in each β subunit ring can cleave peptide bonds at hydrophobic residues, acidic residues, or basic residues. Polypeptide substrates must enter the chamber via the regulated ~1.3-nm-diameter core-gate aperture at the center of

the outer α subunit rings. The core gate is narrow and often allows entry of only unfolded proteins. In the 26S proteasome, the opening of the core gate is controlled by ATPases in the 19S RP. These ATPases are responsible for unfolding protein substrates and translocating those unfolded polypeptides into the inner chamber of the catalytic core (Figure 3-32b, *bottom right*). Once inside the inner chamber, the polypeptides are digested, yielding short peptides (2–24 residues long). These short peptide products exit the chamber and are further degraded rapidly by cytosolic peptidases; eventually they are converted to individual ("free") amino acids. One researcher has quipped that a proteasome is a "cellular chamber of doom" in which proteins suffer a "death by a thousand cuts."

Inhibitors of proteasome function have proved to be exceptionally useful in the laboratory and the clinic. Small molecules that inhibit proteasomes, such as MG132, are used to block degradation by proteasomes in the lab, thereby helping to reveal the role of the proteasome and, as we shall see below, of polyubiquitinylation in a wide variety of processes. Other small-molecule proteasome inhibitors have been used therapeutically. Because of the global importance of protein breakdown by proteasomes in cells, continuous, complete inhibition of proteasomes kills cells. However, the partial inhibition of proteasomes for short intervals is widely used as an approach to cancer chemotherapy, especially to treat multiple myeloma, a cancer involving the abnormal proliferation of cells that make

antibodies (immunoglobulins). The myeloma cells produce abnormally high levels of potentially toxic, aberrant immunoglobulin polypeptide chains, which are degraded by proteasomes. When proteasomes in these cancer cells are inhibited, the buildup of toxic, misfolded immunoglobulin polypeptides leads to cell death. In addition, to survive and grow, myeloma cells require the robust activity of a regulatory protein called *NF-κB* (see Chapter 16) as well as other *pro-survival* and *pro-proliferation* proteins. In turn, NF-κB can function fully and promote survival and proliferation only when its inhibitor, *I-κB*, is disengaged and degraded by proteasomes (see Chapter 16). Partial inhibition of proteasomal activity by a small-molecule inhibitor drug results in increased levels of I-κB and, consequently, reduced NF-κB activity (i.e., loss of its protective activity). Cancer cell proliferation slows, and the cells die by apoptosis (see Chapter 21). Thus multiple myeloma cells are more sensitive to proteasome inhibitors than normal cells. Consequently, *controlled* administration of proteasome inhibitors (e.g., bortezomib, carfilzomib, ixazomib), at levels that kill the cancer cells but not normal cells, has proved to be an effective therapy for multiple myeloma. ■

Ubiquitin Marks Cytosolic Proteins for Degradation in Proteasomes

Proteasomes must be able to distinguish between those proteins that need to be degraded because they are defective or scheduled to be removed and those that don't. Cells mark proteins that should be degraded by covalently attaching to them a linear chain of multiple copies of a 76-residue polypeptide called *ubiquitin (Ub)* that is highly conserved from yeast to humans. This *polyubiquitin tail* serves as a cellular "kiss of death," marking the protein for destruction in the proteasome. The ubiquitinylation process (see Figure 3-32b, steps **1**–**3**) involves three distinct steps:

1. Activation of *ubiquitin-activating enzyme (E1)* by the addition of a ubiquitin molecule, a reaction that requires ATP.
2. Transfer of this ubiquitin molecule to a cysteine residue in a *ubiquitin-conjugating enzyme (E2)*.
3. Formation of a covalent bond between the carboxyl group of the C-terminal glycine 76 of the ubiquitin bound to E2 and the amino group of the side chain of a lysine residue in the target protein, a reaction catalyzed by a *ubiquitin-protein ligase (E3)*. This type of bond is called an *isopeptide bond* because it covalently links a side-chain amino group, rather than the α-amino group, to the carboxyl group. Subsequent ligase reactions covalently attach a series of ubiquitin molecules to generate a polyubiquitin chain covalently attached to the target protein (step **4**). The ligase covalently attaches the C-terminal glycine of each additional ubiquitin molecule via an isopeptide bond to the side chain of lysine 48 of the previously added ubiquitin. (We will discuss ubiquitin linkages via other lysine side chains shortly.)

Generally once four or more ubiquitins are attached in a polyubiquitin chain, the 19S regulatory protein of the 26S proteasome (sometimes with the help of accessory proteins) recognizes the polyubiquitin-labeled protein using its Ub receptors (see Figure 3-32a). Using the hexameric ATPases, the regulatory protein unfolds the protein and transports it into the proteasome core for degradation. As a polyubiquitinylated substrate binds to the proteasome, is unfolded, and passed into the core of the proteasome, enzymes called *deubiquitinases (Dubs)* hydrolyze the bonds between the individual ubiquitins and between the targeted protein and ubiquitin, recycling the ubiquitins for additional rounds of protein modification (see Figure 3-32b). Analysis of the human genome sequence indicates that about 100 distinct Dubs are present, about 80 percent of which use cysteine in a catalytic triad similar to that in the serine proteases described earlier (the sulfhydryl in the cysteine side chain is used in place of the hydroxyl in the side chain of the serine). In some Dubs, zinc is a key participant in the catalytic reactions.

Specificity of Degradation

Targeting of specific proteins for proteasomal degradation is primarily achieved through the substrate specificity of E3 ligases (see Figure 3-32b, step **3**). As a testament to their importance, there are an estimated 600 or more ubiquitin ligase genes in the human genome. The many E3 ligases in mammalian cells ensure

that the wide variety of proteins to be polyubiquitinylated can be modified when necessary. Some E3 ligases are associated with chaperones that recognize unfolded or misfolded proteins; for example, the E3 ligase CHIP is a co-chaperone for Hsp70. These and other proteins (co-chaperones, escort factors, adapters) can guide E3 to dysfunctional proteins that cannot be readily refolded properly and, consequently, are essential to their delivery to proteasomes for degradation. In such cases, the chaperone-ubiquitinylation-proteasome system works in concert to achieve quality control of proteins. The proteasome-mediated quality control system also helps cells eliminate other deleterious proteins, including those that are mistranslated, not properly located within the cell, aggregated, mutated, chemically altered, or unable to assemble into their proper multipeptide quaternary structures. In addition to quality control, the ubiquitin-proteasome system can be used to regulate the activity of important cellular proteins. An example is the regulated degradation of proteins called *cyclins*, which control the cell cycle (see Chapter 19).

Other Functions of Ubiquitin and Ubiquitin-Related Molecules

There are several close relatives of ubiquitin that employ similar E1-, E2-, and E3-dependent mechanisms of activation and transfer to acceptor substrates. These ubiquitin-like modifiers include Sumo, which regulates nuclear import, and Atg8/LC3, a regulator of

autophagy (see Chapter 14). Furthermore, the attachment of ubiquitin to a target protein can be used for purposes other than to mark the protein for degradation, as we will see later in this section, and some of these functions involve polyubiquitin linkages other than those via Lys-48.

Like ubiquitinylation, deubiquitinylation is involved in processes other than proteasome-mediated protein degradation. Large-scale, mass-spectrometry-based *proteomic* methods described later in this chapter, together with sophisticated computational approaches, have suggested that Dubs, which are often bound in multiprotein complexes, are involved in an extraordinarily wide range of cell processes. These processes vary from cell division and cell cycle control (see Chapter 19) to membrane trafficking (see Chapter 14) to cell signaling pathways (see Chapters 15 and 16).

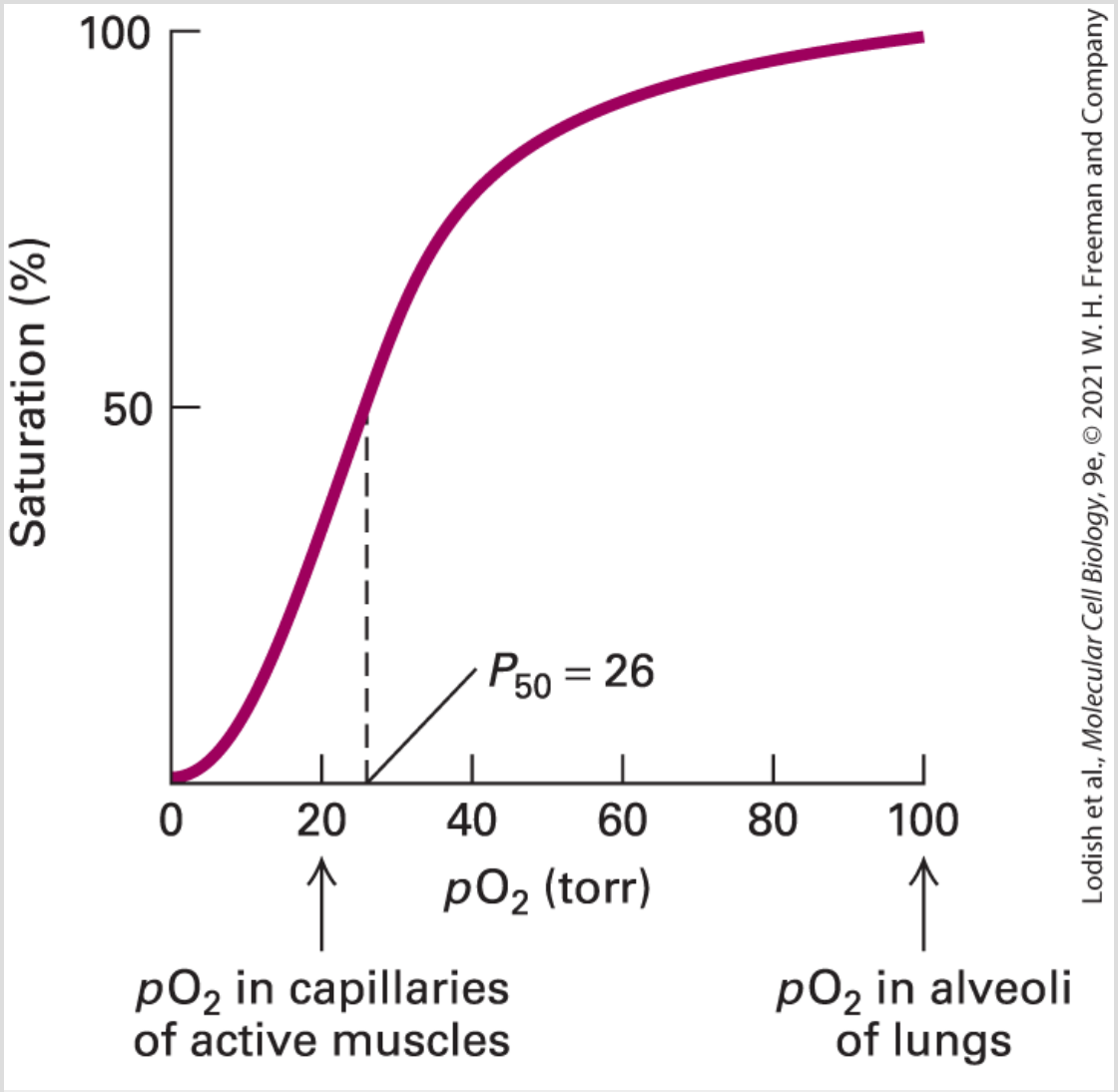
Noncovalent Binding Permits Allosteric, or Cooperative, Regulation of Proteins

In addition to regulating the amount of a protein, cells can also regulate the intrinsic activity of a protein. One of the most important mechanisms for regulating protein function is through allosteric interactions. Broadly speaking, **allostery** (from the Greek, "other shape") refers to any change in a protein's tertiary or

quaternary structure, or in both, induced by the noncovalent binding of a ligand. When a ligand binds to one site (A) in a protein and induces a conformational change that alters the activity of a different site (B), the ligand is called an *allosteric effector* of the protein, while site A is called an *allosteric-binding site*, and the protein is called an *allosteric protein*. By definition, allosteric proteins have multiple binding sites, at least one for the allosteric effector and at least one for other molecules with which the protein interacts. The allosteric change in activity can be positive or negative; that is, it can be an increase or a decrease in protein activity. Negative allostery often involves the end product of a multistep biochemical pathway binding to, and reducing the activity of, an enzyme that catalyzes an early, rate-controlling step in that pathway. In this way, excessive buildup of the product is prevented. This kind of regulation of a metabolic pathway is also called *end-product inhibition* or *feedback inhibition*. Allosteric regulation is particularly prevalent in multimeric enzymes and other proteins in which conformational changes in one subunit are transmitted to an adjacent subunit.

Cooperativity, a term that is often used synonymously with *allostery*, usually refers to the influence (positive or negative) that the binding of a ligand at one site has on the binding of another molecule of the *same* type of ligand at a different site. Hemoglobin presents a classic example of positive cooperative binding in that the binding of a single ligand, molecular oxygen (O_2), increases the affinity of

hemoglobin for the next oxygen molecule. Each of the four subunits in hemoglobin contains one heme molecule. The heme groups are the oxygen-binding components of hemoglobin. The binding of oxygen to the heme molecule in one of the four hemoglobin subunits induces a local conformational change whose effect spreads to the other subunits, lowering the K_d (increasing the affinity) for the binding of additional oxygen molecules to the remaining hemes and yielding a sigmoidal oxygen-binding curve (Figure 3-33). Because of the sigmoidal shape of the oxygen-saturation curve, it takes only a fourfold increase in oxygen concentration for the saturation of the oxygen-binding sites in hemoglobin to go from 10 to 90 percent. Conversely, if there were no cooperativity and the shape of the curve was typical of that for Michaelis–Menten (see Figure 3-25), or noncooperative, binding, it would take an 81-fold increase in oxygen concentration to accomplish the same increase in loading of its binding sites in hemoglobin. This cooperativity permits hemoglobin to take up oxygen very efficiently in the lungs, where the oxygen concentration is high, and unload it in tissues, where the concentration is low. Thus cooperativity amplifies the sensitivity of a system to changes in the concentration of its ligands, providing in many cases a selective evolutionary advantage.



EXPERIMENTAL FIGURE 3-33 Hemoglobin binds oxygen cooperatively. Each tetrameric hemoglobin molecule has four oxygen-binding sites; at saturation, all the sites are loaded with oxygen. The oxygen concentration in tissues is commonly measured as the partial pressure of oxygen (pO_2) in torr units (a standard measure of pressure equivalent to 1 mm of mercury under standard conditions). P_{50} is the pO_2 at which half the oxygen-binding sites are occupied; it is somewhat analogous to the K_m for an enzymatic reaction. The large change in the amount of oxygen bound over a small range of pO_2 values permits efficient unloading of oxygen in peripheral tissues such as muscle. The sigmoidal shape of a plot of saturation versus ligand concentration is indicative of cooperative binding, in which the binding of one oxygen molecule allosterically influences the binding of subsequent oxygen

molecules. In the absence of cooperative binding, a binding curve is a hyperbola, similar to the curves in Figure 3-25.

Noncovalent Binding of Calcium and GTP Are Widely Used as Allosteric Switches to Control Protein Activity

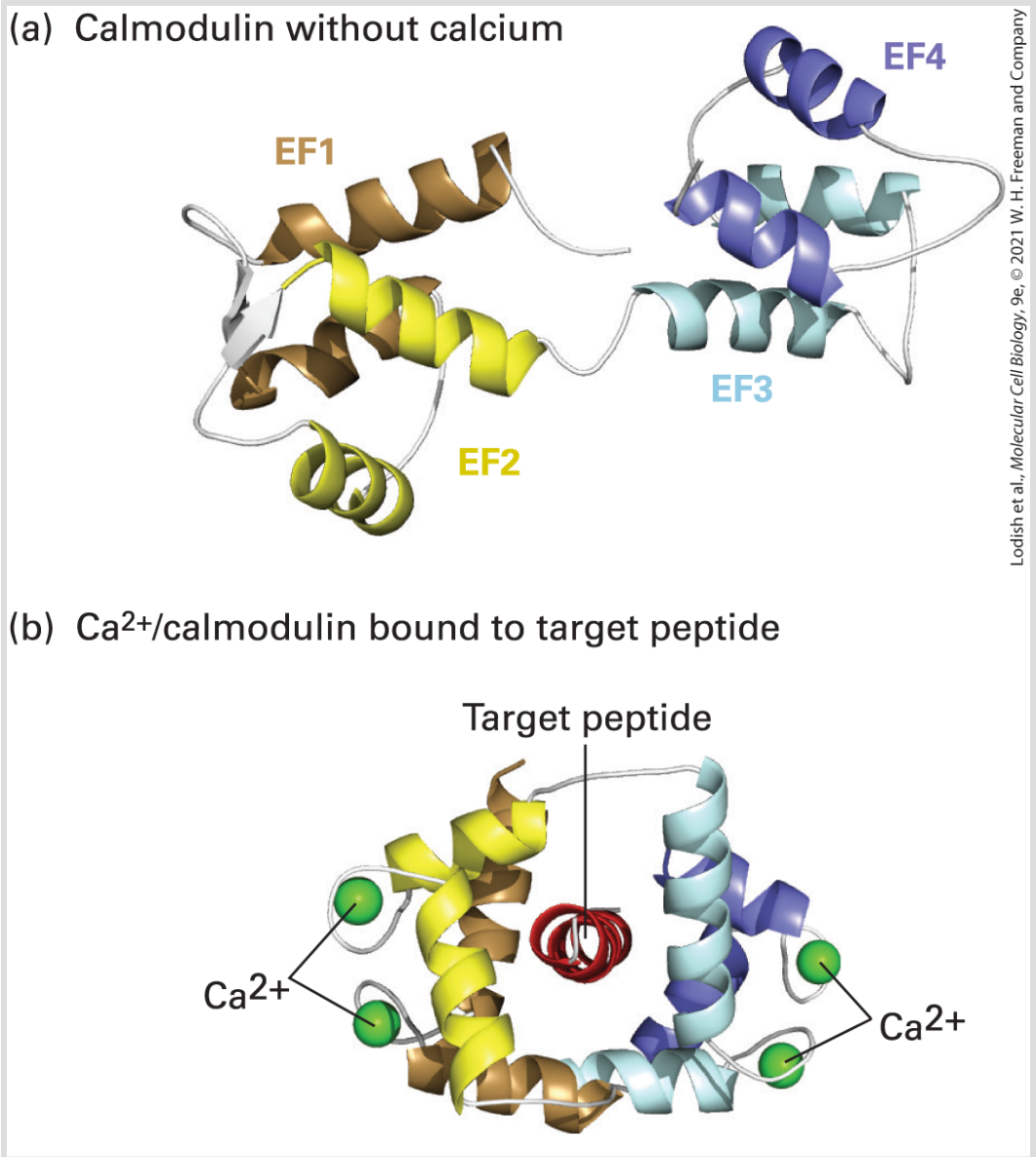
Unlike oxygen, which causes graded allosteric changes in the activity of hemoglobin, some other allosteric effectors act as switches, turning the activity of many different proteins on or off by binding to them noncovalently. Two important allosteric switches that we will encounter many times throughout this book, especially in the context of cell signaling pathways (see Chapters 15 and 16), are Ca^{2+} and GTP.

Ca^{2+}/Calmodulin-Mediated Switching

The concentration of Ca^{2+} that is free in the cytosol (not bound to molecules other than water) is kept very low ($\sim 10^{-7}$ M) by specialized membrane transport proteins that continually pump excess Ca^{2+} out of the cytosol (see Chapters 11 and 15). However, as we will learn in Chapters 11 and 15, the cytosolic Ca^{2+} concentration can increase tenfold to a hundredfold when Ca^{2+}-permeable channels in the cell-surface membranes open and allow extracellular Ca^{2+} to flow into the cell. This rise in cytosolic Ca^{2+} is

sensed by specialized Ca^{2+}-binding proteins, which alter cellular behavior by turning the activities of other proteins on or off. The importance of extracellular Ca^{2+} for cell activity was first documented by S. Ringer in 1883, when he discovered that isolated rat hearts suspended in an NaCl solution made with hard (Ca^{2+}-rich) London tap water contracted beautifully, whereas they beat poorly and stopped quickly if distilled (Ca^{2+}-depleted) water was used. The cytosolic Ca^{2+} concentration also can increase when Ca^{2+}-permeable channels in the endoplasmic reticulum open and allow Ca^{2+} stored in this organelle flow into the cytoplasm.

Many Ca^{2+}-binding proteins bind Ca^{2+} using the EF hand/helix-loop-helix structural motif discussed earlier (see Figure 3-7b). A well-studied EF hand protein, *calmodulin,* is found in all eukaryotic cells, where it may exist as an individual monomeric protein or as a subunit of a multimeric protein. This dumbbell-shaped molecule contains four Ca^{2+}-binding EF hands with K_ds of about 10^{-6} M. The binding of Ca^{2+} to calmodulin causes a conformational change that permits Ca^{2+}/calmodulin to bind to conserved sequences in various target proteins (Figure 3-34), thereby switching their activities on or off. Calmodulin and similar EF hand proteins thus function as *switch proteins,* acting in concert with changes in Ca^{2+} levels to modulate the activity of other proteins.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 3-34 Conformational changes induced by Ca^{2+} binding to calmodulin. Calmodulin is a widely distributed cytosolic protein that contains four Ca^{2+}-binding sites, one in each of its EF hand (helix-loop-helix) motifs (EF1–EF4; see also Figure 3-7). At cytosolic Ca^{2+} concentrations above about 5×10^{-7} M, binding of Ca^{2+} to calmodulin changes the protein's conformation from the dumbbell-shaped, Ca^{2+}-free form (a) to one in which hydrophobic side chains become more exposed to solvent. The resulting Ca^{2+}/calmodulin complex can wrap around exposed helices with specialized sequences in various target proteins (b), thereby altering their activities.

[Part (a) Data from H. Kuboniwa et al., 1995, *Nat. Struct. Biol.* **2**:768–776, PDB ID 1cfd. Part (b) Data from W. E. Meador, A. R. Means, and F. A. Quiocho, 1992, *Science* **257**: 1251–1255, PDB ID 1cdl.]

Switching Mediated by Guanine Nucleotide–Binding Proteins

Another group of intracellular switch proteins constitutes the *GTPase superfamily*. As the name suggests, these proteins are enzymes — GTPases — that can hydrolyze GTP (guanosine triphosphate) to GDP (guanosine diphosphate). They include the monomeric Ras protein (whose structure is shown in Figure 3-9, with bound GDP shown in blue) and the G_α subunit of the trimeric G proteins, both discussed at length in Chapters 15 and 16. Both Ras and G_α can bind to the plasma membrane, function in cell signaling, and play key roles in cell proliferation and differentiation. Other members of the GTPase superfamily function in protein synthesis, the transport of proteins between the nucleus and the cytoplasm, the formation of coated vesicles and their fusion with target membranes, vesicle-mediated membrane trafficking, and rearrangements of the actin cytoskeleton. Some GTPase proteins have a covalently attached lipid chain (see Figure 10-19) that mediates their binding to membranes. We examine the roles of various GTPase switch proteins in regulating intracellular signaling and other processes in several later chapters.

All the GTPase switch proteins exist in two forms, or conformations (Figure 3-35): (1) an active ("on") form with bound GTP, which can influence the activity of specific target proteins to which they bind, and (2) an inactive ("off") form with bound GDP. The switch is turned on — that is, the conformation of the protein changes from inactive to active — when a GTP molecule replaces a bound GDP in the inactive conformation. The switch is turned off when the relatively slow GTPase activity of the protein hydrolyzes bound GTP, converting it to GDP and leading the conformation to change to the inactive form. The amount of time any given GTPase switch remains in the active, GTP-bound form depends on the rate of its GTPase activity. Thus the GTPase activity acts as a timer to control this switch.

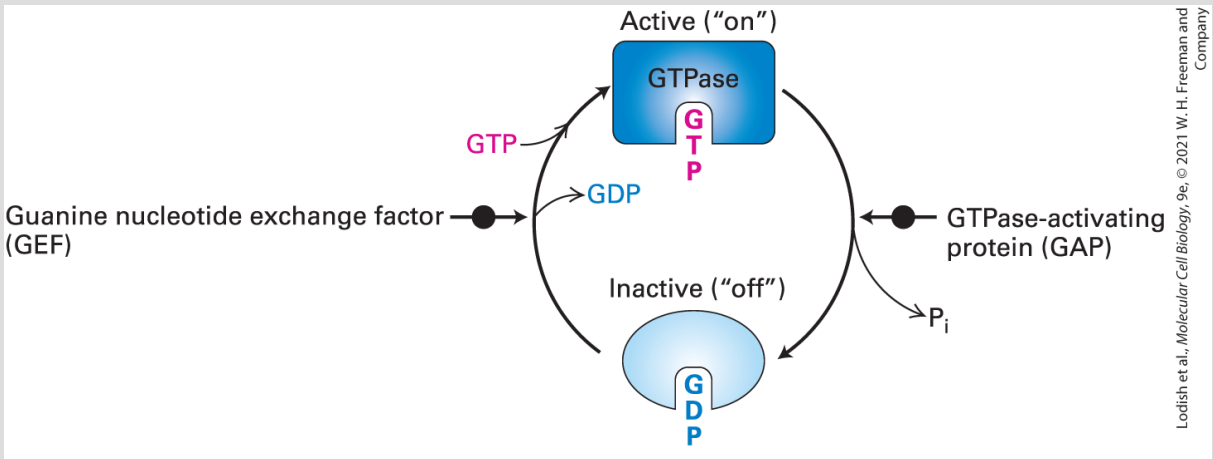


FIGURE 3-35 The GTPase switch. GTPases are enzymes that bind to GTP and hydrolyze it to GDP. When bound to GTP, the GTPase protein adopts its active, or "on," conformation and can interact with target proteins to regulate their activities. When the bound GTP is hydrolyzed to GDP by the intrinsic GTPase activity of the protein, the GTPase with GDP bound assumes an inactive, or "off," conformation. The GTPase switch can be turned back on when another protein, called a *GEF (guanine nucleotide exchange factor)*, mediates the replacement (exchange) of the bound GDP with a GTP molecule from the surrounding fluid.

GTPase-activating proteins, or GAPs, can influence the rates of GTP hydrolysis. The binding of the active form of the GTPase to its targets is a form of noncovalent regulation.

Cells contain a variety of proteins that can modulate the baseline (or intrinsic) rate of GTPase activity for any given GTPase switch and so can control how long the switch remains on. For example, GTPase-activating proteins, or GAPs, increase the rate of GTPase activity, thus reducing the time the GTPase is in the active form. Cells also have specific proteins whose function is to regulate the conversion of inactive GTPases to active ones — that is, to turn the switch on — by mediating the replacement of bound GDP with GTP (*GDP/GTP exchange*). These proteins are called *guanine nucleotide exchange factors*, or *GEFs*. GTPases with lipid anchors are also regulated by proteins called *guanine nucleotide dissociation inhibitors (GDIs)* that bind to the lipid chain and thus influence interactions with cellular membranes.

The GAPs, GEFs, and GDIs are themselves subject to regulation and, together with their GTPases, participate in complex regulatory networks that control a vast array of cellular activities. It is, therefore, not surprising that disruptions of these finely tuned regulatory networks by mutations or pathogens are associated with a wide variety of diseases. Examples of genetic diseases affecting these networks include Noonan syndrome (a developmental disorder), retinitis pigmentosa (a degenerative eye disease), and X-linked mental retardation. Examples of disruptions of these

networks by pathogens include bacterially induced food poisoning, dysenteries (inflammation of the intestines with diarrhea), Legionnaires' disease (a severe type of pneumonia that involves lung inflammation), and even the plague [also called the *Black Death*, which between 1347 and 1351 decimated the populations of China (~50 percent death rate) and Europe (~33 percent death rate)]. ■

Covalent Modification of Proteins Can Regulate their Activities

In addition to exploiting the noncovalent regulators described above, cells can use covalent modifications to regulate the intrinsic activity of a protein. Examples of such modifications, called **post-translational modifications (PTMs)**, include covalent addition of phosphate groups (phosphorylation), methyl groups (methylation), nitric oxide (nitroslyation), acetyl groups (acetylation), carbohydrate, fatty acyl or other hydrocarbon chains (lipidation), and even small proteins [e.g., ubiquitin (ubiquitinylation)]. The covalent modification can increase or decrease the activity of the modified protein by changing its surface characteristics (shape, charge, etc.), and thus its ability to bind to other molecules, its conformation, or both. The enzymes that covalently modify target proteins have been called *writers*, the proteins that recognize the modified proteins are called *readers*, and for those covalent modifications that are

reversible, the enzymes that remove the modifications are called *erasers*.

Phosphorylation and Dephosphorylation Covalently Regulate Protein Activity

One of the most common covalent mechanisms for regulating protein activity is **phosphorylation**, the reversible addition of a phosphate group from ATP to an hydroxyl group on the side chain of a serine, threonine, or tyrosine residue. Phosphorylated proteins are called *phosphoproteins*. Phosphorylation is catalyzed by enzymes called protein **kinases**, while the removal of phosphates, known as *dephosphorylation*, is catalyzed by **phosphatases**. The counteracting activities of kinases and phosphatases provide cells with a switch that can turn on or turn off the function of various proteins that are the substrates (or targets) of these enzymes (Figure 3-36). Protein kinases constitute one of the largest families of enzyme molecules; they are found in all eukaryotic cells and nearly all prokaryotic cells. Nearly 3 percent of all yeast proteins are protein kinases or phosphatases, indicating the importance of phosphorylation and dephosphorylation reactions even in these simple cells. Analysis of the human genome indicates there are approximately 500 human protein kinases composing the human *kinome* (see Figure 3-12b). For

simplicity, we will limit our discussion here to eukaryotic protein kinases.

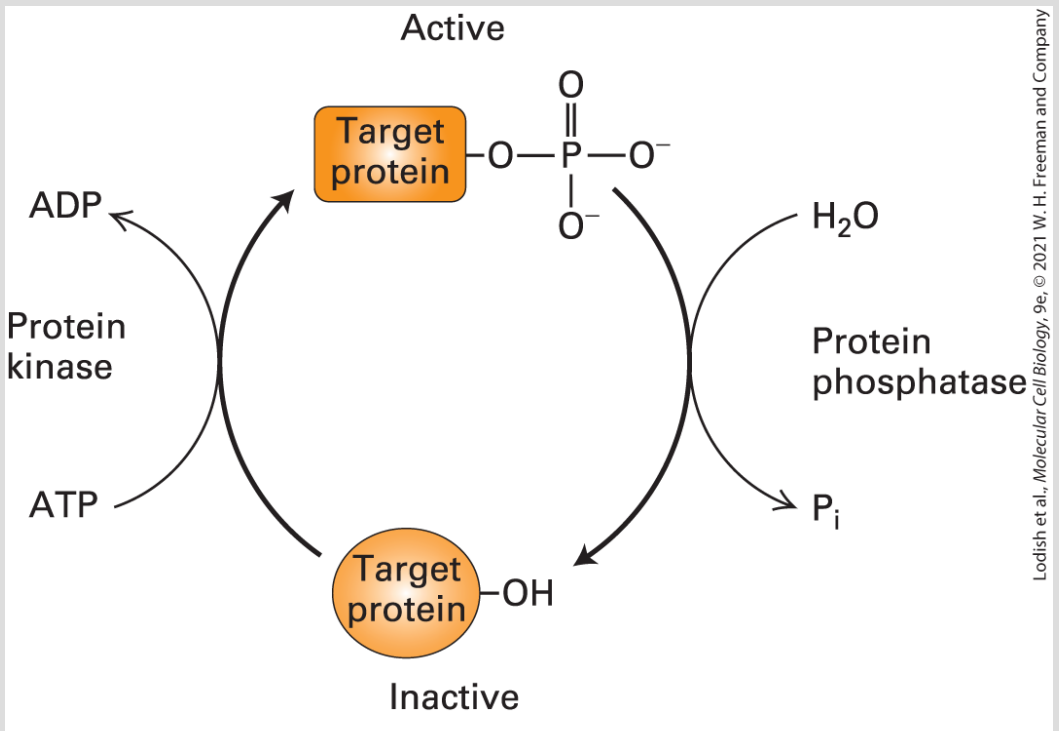


FIGURE 3-36 Regulation of protein activity by phosphorylation and dephosphorylation. The cyclic phosphorylation and dephosphorylation of a protein is a common cellular mechanism for regulating protein activity. In this example, the target protein is active (*top*) when phosphorylated and inactive (*bottom*) when dephosphorylated; some proteins have the opposite response to phosphorylation.

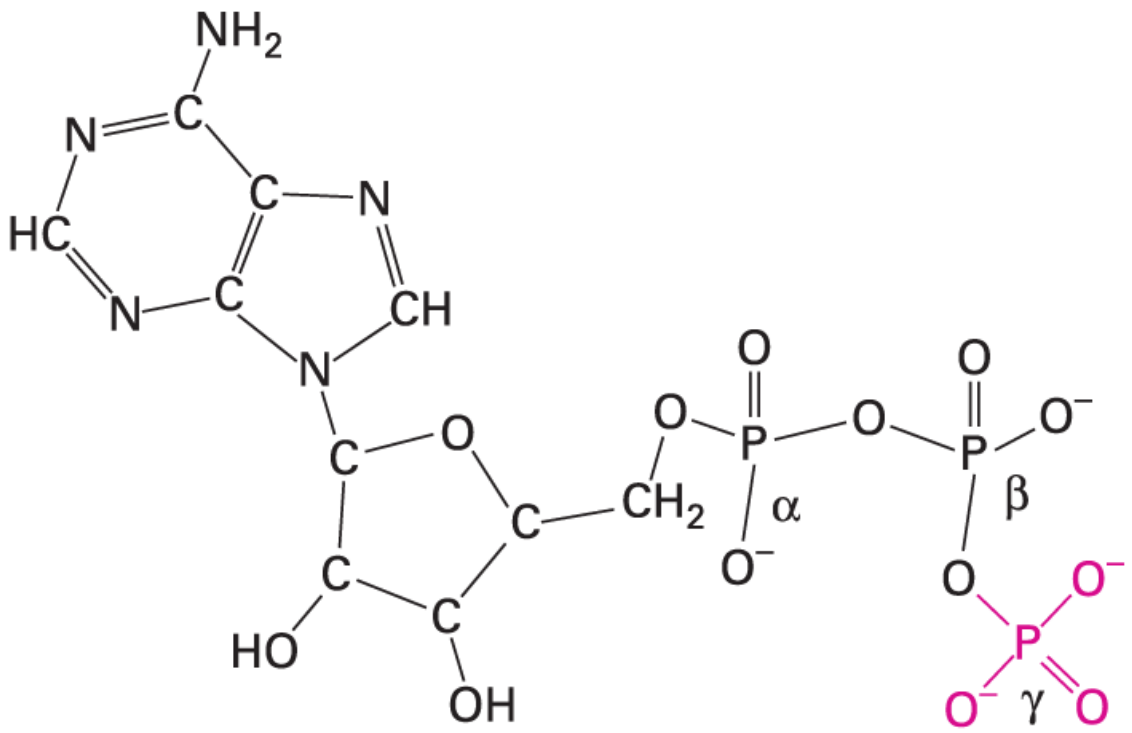
Phosphorylation or dephosphorylation, writing or erasing, can influence the location of a protein within cells (e.g., its attachment to the inner surface of the plasma membrane); its structure and intrinsic (e.g., enzymatic) activity; its ability to bind to other molecules, including metabolites, DNA, or other proteins; its ability to undergo further covalent modification; or its stability (rate of

degradation). In addition, several conserved protein domains, such as the SH2 domain (see Figure 16-9), bind specifically to phosphorylated peptides (readers). Thus phosphorylation can mediate the formation of protein complexes that can generate or extinguish a wide variety of cellular activities, discussed in many subsequent chapters.

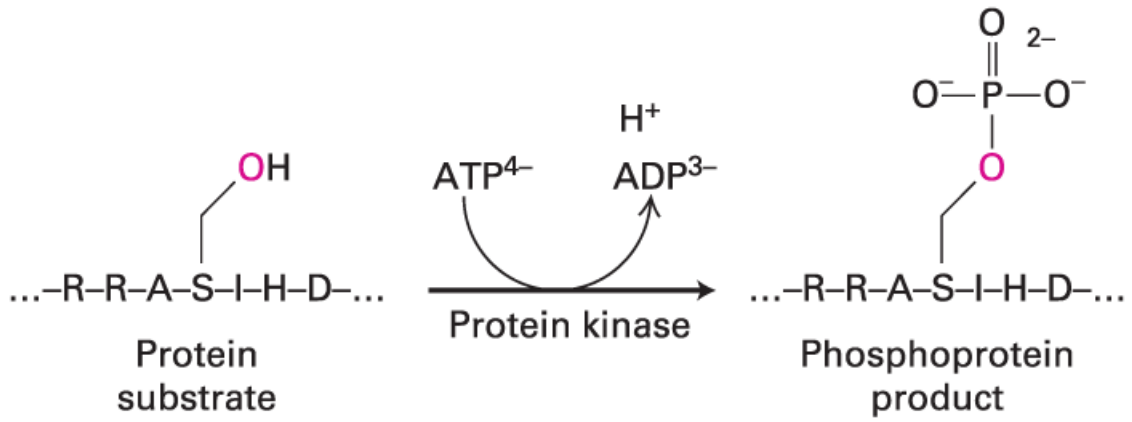
All classes of proteins — including structural proteins, scaffolds, enzymes, membrane channels, and signaling molecules — have members regulated by kinase and phosphatase modifications. Different protein kinases and phosphatases are specific for different target proteins, often recognizing different linear sequences in which the residue to be phosphorylated is embedded, and so can regulate distinct cellular pathways, as discussed in later chapters. Some kinases have many targets, so that a single kinase can serve to integrate the activities of many targets simultaneously. Frequently, the target of a kinase or phosphatase is yet another kinase or phosphatase, creating a cascade effect. There are many examples of such kinase cascades, which permit amplification of a signal and many levels of fine-tuning (see Chapters 15 and 16).

The Structure and Function of Protein Kinase A Is Typical of Many Kinases

The phosphate group transferred by a protein kinase to a protein substrate is provided by ATP whose three phosphates are designated α, β, and γ (red):



The γ-phosphate is transferred to a serine, threonine, or tyrosine of the polypeptide substrate. Specifically, the hydroxyl group that is part of the side chain of these three amino acids is replaced by a phosphoryl group:

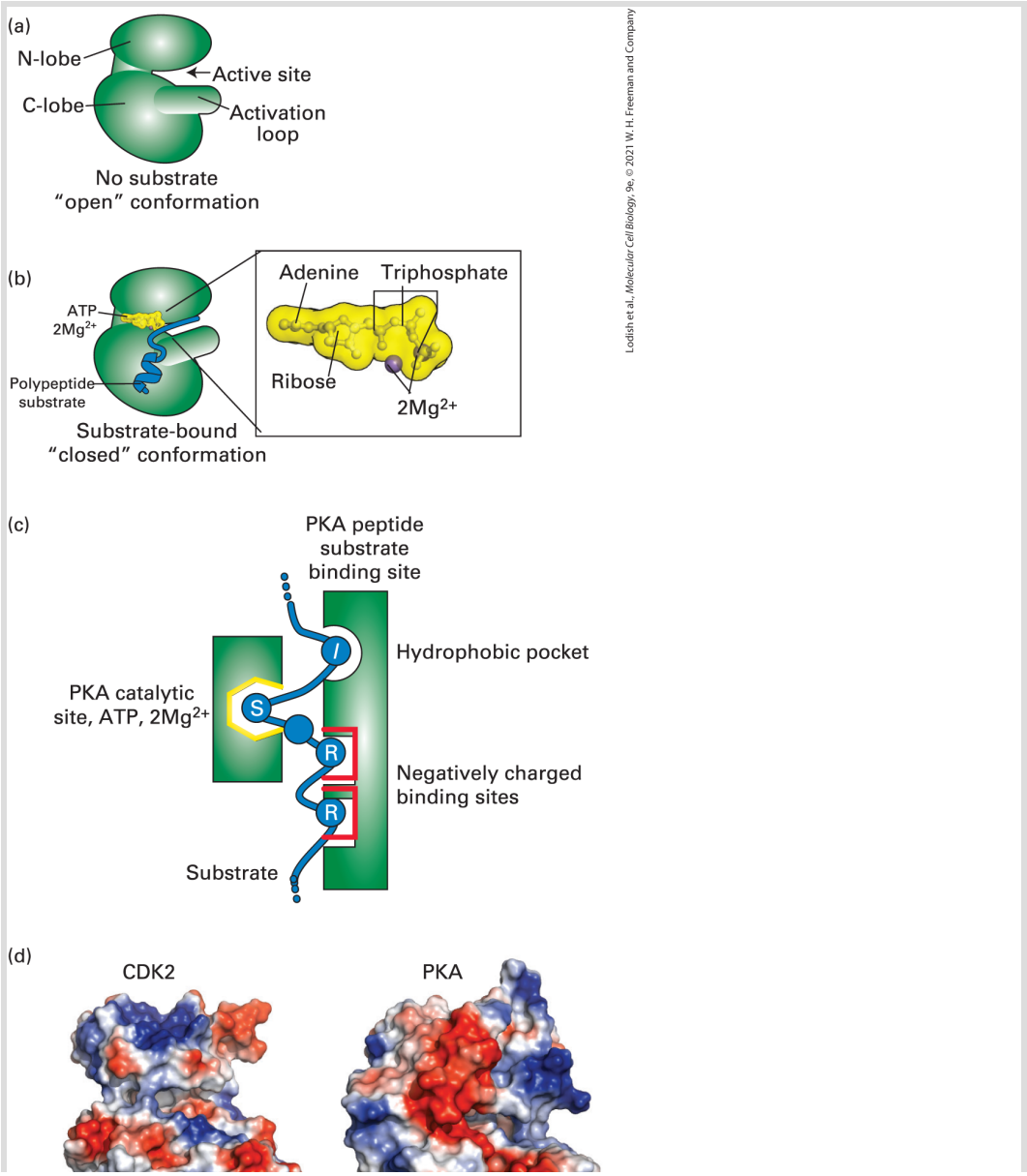


The reaction products are ADP, H^+, and a phosphorylated protein — that is, a protein containing a phosphate ester (charges on the molecules are shown as superscripts). In this example, the amino acid sequence is a substrate for protein kinase A (PKA), a protein kinase that we will encounter a number of times throughout the book.

All protein kinases contain a *kinase domain*, an enzymatic module of around 250 residues that catalyzes the phosphate transfer reaction. Many protein kinases also contain additional protein domains or binding partners that assist in selecting the protein substrate, participate in regulating the kinase's enzymatic activity, help form multiprotein complexes with other cellular proteins (scaffolds), and localize the enzyme to particular sites in the cell.

In 1991, the first x-ray crystal structure of a catalytic kinase domain from the enzyme protein kinase A was determined. Its bilobed structure resembles a kidney bean (Figure 3-37a). The lobe containing the N-terminus of the domain is called the *N-lobe* and

that containing the C-terminus is called the *C-lobe*. A region in the C-lobe called the *activation loop* participates in the regulation of the kinase's enzymatic activity and will be discussed later. Most protein-kinase catalytic domains have similar structures and use similar mechanisms to catalyze phosphorylation of target proteins. Here we will focus on the prototypical kinase, PKA, one of the most thoroughly studied kinases.



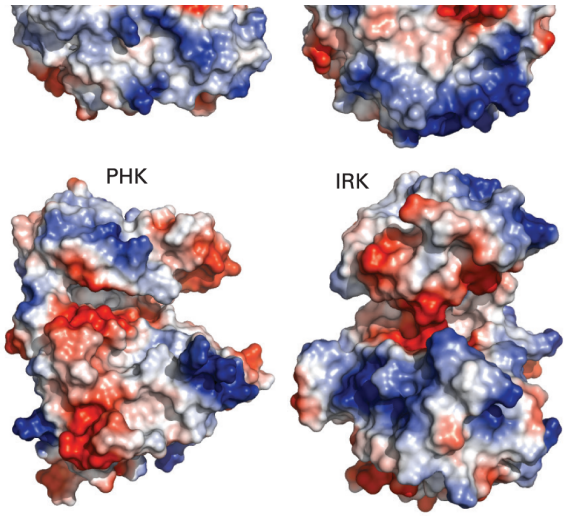


FIGURE 3-37 The kinase domain of protein kinases: structure and substrate binding. A kinase domain comprises two globular lobes (N-lobe and C-lobe). The active site is located in a cleft between the two lobes, and an activation (A-) loop in the C-lobe contributes to the active site and projects outward [(a) and (b)]. (a) When catalytically active and in the absence of substrates, the active form of the kinase domain is in an "open" conformation that facilitates substrate binding. (b) When bound to substrates (ATP, $2\mathrm{Mg}^{2+}$, polypeptide substrate) in an enzyme substrate (ES) complex, the active site narrows to its "closed" conformation to facilitate the phosphorylation of the polypeptide substrate. The inset shows a magnified view of the ATP and two Mg^{2+} ions in the ES complex. Divalent metal ions help neutralize the negative charges on the phosphates of ATP. (c) Most protein kinases exhibit substrate specificity by binding to a specific sequence of residues (a sequence motif) on either side of the hydroxyl-containing residue to be phosphorylated. Binding pockets in the substrate-binding site (*right*) of PKA establish its specificity for the sequence R-R-X-(Ser or Thr)-hydrophobic (e.g., isoleucine, I, is shown). Negatively charged binding sites (rectangles) that recognize the side chains of Arg or Lys in the substrate and a hydrophobic pocket (round) that recognizes hydrophobic side chains in the substrate are illustrated. (d) Surface representations of protein kinase domains from four different proteins: CDK2; PKA, phosphorylase kinase (PHK); and insulin receptor kinase (IRK). Each has a distinctive, detailed structure and exhibits a distinct distribution of surface charges (positively charged are blue, negatively charged are red).

[Part (b) Substrate structures data from PDB ID 1ATP (https://www.rcsb.org/3d-view/1ATP/1). Parts (c) and (d) Data from Figures 3 and 1 from J. A. Ubersax and J. E. Ferrell, 2007, *Nat. Rev. Mol. Cell Biol.* **8**:530–541.]

The kinase active site is located in and around the cleft between the two lobes (see Figure 3-37a). The active site contains substrate-binding sites that bind and properly orient the ATP molecule (and two associated magnesium divalent cations, Mg^{2+}) within the cleft and the polypeptide substrate on the ledge of and adjacent to the cleft. This binding appropriately positions the hydroxyl group of the phospho-accepting serine, threonine, or tyrosine side chain relative to the ATP and the key catalytic residues in the kinase (Figure 3-37b). In the absence of substrates, the active site cleft is "open" (see Figure 3-37a). The binding of the substrates places the γ-phosphate of the ATP in close proximity to the hydroxyl group of the acceptor amino acid on the polypeptide substrate and, in an example of induced fit, induces a conformational change, primarily in the N-lobe, to a "closed" conformation that narrows the cleft (see Figure 3-37b) and establishes a configuration that optimizes catalysis. The active site also contains the catalytic site that facilitates in the transfer of the γ-phosphate group from the ATP to the polypeptide substrate.

Only a small subset of protein sequences, those containing the proper substrate motif, are able to bind productively to the active site and be phosphorylated. Different protein kinase enzymes differ in their substrate specificities. The substrate specificity of any kinase enzyme often depends on the amino acid sequence

immediately surrounding the acceptor residue, although specificity can also be influenced by other regions in the substrate protein. For example, PKA typically phosphorylates its substrate proteins on serine and threonine residues only if they are found within the sequence motif RRX**S**Φ or RRX**T**Φ, where X is any amino acid and Φ denotes a hydrophobic amino acid (e.g., see Figure 3-37c) and the phosphorylated residue is shown in bold.

A glimpse at the active site of PKA's kinase domain explains the molecular basis for this substrate specificity (Figure 3-37c). Three negatively charged, acidic side chains in the substrate-binding site directly bind to the two positively charged Arg (R) side chains in the substrate motif, and a hydrophobic pocket in PKA recognizes the hydrophobic amino acid side chain in the motif immediately C-terminal to the acceptor serine or threonine residue. As a consequence, only proteins with this sequence motif can bind productively to PKA and be phosphorylated. The substrate specificities of other kinases are similarly established by the molecular complementarity of their substrate-binding sites (including the distributions of charges and locations of binding pockets on their surfaces (Figure 3-37d) and the substrate polypeptides. Substrate specificity also can be influenced by other regions of both the kinase and substrate polypeptide.

In addition to the substrate-binding site, there are two stretches of amino acids in kinases called the *G-loop* and *A-loop* that are critically

important for substrate binding and catalysis. Binding of ATP-2Mg^{2+} or the protein enhances binding of the other substrate, in an example of cooperative binding, and, in addition, induces conformational changes in the enzyme, an example of induced fit. Once the ATP and polypeptide substrate are appropriately aligned in the active site (Figure 3-37b), the catalytic site mediates the rapid transfer of the γ-phosphate of ATP to the hydroxyl group of the acceptor amino acid on the protein substrate (see the protein kinase reaction, this page above) at rates 10^9-fold to 10^{11}-fold greater than that of the uncatalyzed phosphorylation reaction. The catalytic mechanism is completed with the sequential release of the phosphoprotein product with one bound Mg^{2+} followed by what in some kinases is the slowest step in the reaction pathway: the release of ADP with one bound Mg^{2+}. The slowest step in a multistep, enzyme-mediated catalysis is usually called the **rate-limiting step** because that step has the most profound influence on the overall rate of the reaction.

Protein Kinase Activity Is Often Regulated by Phosphorylation of the Kinase

If the activities of protein substrates are influenced by their phosphorylation by protein kinases, then there must be mechanisms that regulate the protein kinases themselves, and

indeed there are. Here we will consider one common mechanism employed to switch many protein kinases from an inactive state to an active state. This mechanism is the covalent attachment of a phosphate to a particular residue in the kinase's A-loop. Thus many kinases must be phosphorylated by another kinase — a type of enzymatic cascade — before they can be active and phosphorylate their own substrates. Phosphorylation of the kinase induces a change in its conformation that brings its substrate-binding and catalytic residues into the proper open orientation to permit efficient substrate binding and subsequent phosphate transfer. The G- and A-loops mentioned above have sequences that are conserved in many kinases and play particularly important roles in controlling the enzymatic activities of kinases.

The conserved sequence of the G-loop is G-X-G-X-X-G-X-V where X is any residue (the G in G-loop refers to the conserved glycine residues). Residues in this sequence interact with the β and γ phosphates of ATP for productive ATP binding and catalysis. A-loop residues regulate substrate binding and directly influence catalysis. When phosphorylated, one of the A-loop residues triggers activation; hence it is called the *activation (A) loop*. The residues in both loops must be properly oriented in the protein if substrates are to bind productively and phosphate transfer is to be catalyzed efficiently. Figure 3-38 shows that for PKA in the absence of phosphorylation of Thr-197 in the A-loop (*left*), neither the A-loop (thick red line) nor the G-loop (thick orange line) are in the proper

conformation and, as a consequence, the kinase is inactive. Phosphorylation of Thr-197 (yellow circle, *center*) induces a conformational change in the enzyme, generating an active, open, structure to which substrates can bind productively (Figure 3-38a *right* and 3-38b). While the conformations of the A-loop in the active structures of many different protein kinases are similar, the conformations of the A-loops in the inactive structures vary considerably. In subsequent chapters we will see many examples of how cells activate or inactivate protein kinases to control important cellular functions. We also will see, in addition to phosphorylation of the A-loop and conformational changes in the G-loop, there are a variety of scaffold proteins and inhibitory proteins that can bind to protein kinases and influence their activities.

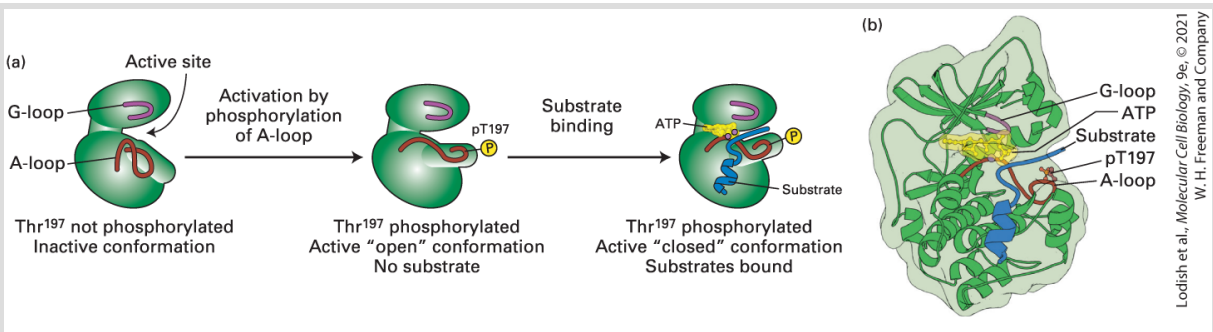


FIGURE 3-38 PKA's G-loop, catalytic loop, and A-loop mediate substrate binding and catalysis after activation by A-loop phosphorylation. (a) The protein kinase domain of PKA is transformed from an inactive conformation (*left*) to the active, open conformation (*center*) by phosphorylation (yellow circle) of Thr-197 in the A-loop (thick red line). Phosphorylation results in conformational changes in both the A-loop and the G-loop (thick orange line). Subsequent substrate binding (ATP, yellow; $2\ Mg^{2+}$, brown; peptide substrate, blue) leads to formation of the active, closed conformation (*right*) in which substrate phosphorylation is catalyzed. (b) A surface-and-ribbon (green) representation of phosphorylated PKA in the

active, closed conformation with ATP (yellow) and a polypeptide substrate (blue) bound. The phosphorylated tyrosine (pT197) is yellow.

[Data from E. A. Madhusudan et al., 2002, *Nat. Struct. Biol*. **9**:273–277; S. S. Taylor and A. P. Kornev, 2011, *Trends Biochem. Sci*. **36**:65–77; and J. M. Steichen et al., 2010, *J. Biol. Chem*. **285**:3825–3832.]

Ubiquitinylation and Deubiquitinylation Covalently Regulate Protein Activity

We have already seen how polyubiquitinylation can be used to tag proteins for degradation by proteasomes. The covalent addition of ubiquitin and ubiquitin-like proteins (of which there are more than a dozen in humans) to a target protein also can be used to regulate the activity of the target in a manner analogous to phosphorylation. In addition, deubiquitinases (Dubs) can remove ubiquitins in a manner analogous to the action of phosphatases. **Ubiquitinylation** can involve attachment of a single ubiquitin to a protein (*monoubiquitinylation*), addition of multiple single ubiquitin molecules to different sites on one target protein (***multiubiquitinylation***), or addition of a polymeric chain of ubiquitins to a protein (*polyubiquitinylation*), as we previously described for tagging target proteins for proteasomal degradation. An additional source of variation is that different amino groups in the ubiquitin molecule can be used to form an isopeptide bond with

the C-terminal Gly-76 in another ubiquitin to form a polyubiquitin chain. All seven lysine residues in ubiquitin (Lys-6, Lys-11, Lys-27, Lys-29, Lys-33, Lys-48, and Lys-63) and its N-terminal amino group (see Figure 3-32c) can participate in such linkages. These linkages are catalyzed by a variety of ubiquitin ligases, each specific for both the target to be ubiquitinylated and the lysine side chains that participate in the isopeptide linkages (Lys-63 or Lys-48, etc.) (Figure 3-39a). Because any one ubiquitin molecule has multiple lysines to which additional ubiquitins can be attached (see Figure 3-32c), polyubiquitin chains can be linear (Figure 3-39a) or branched (Figure 3-39b).

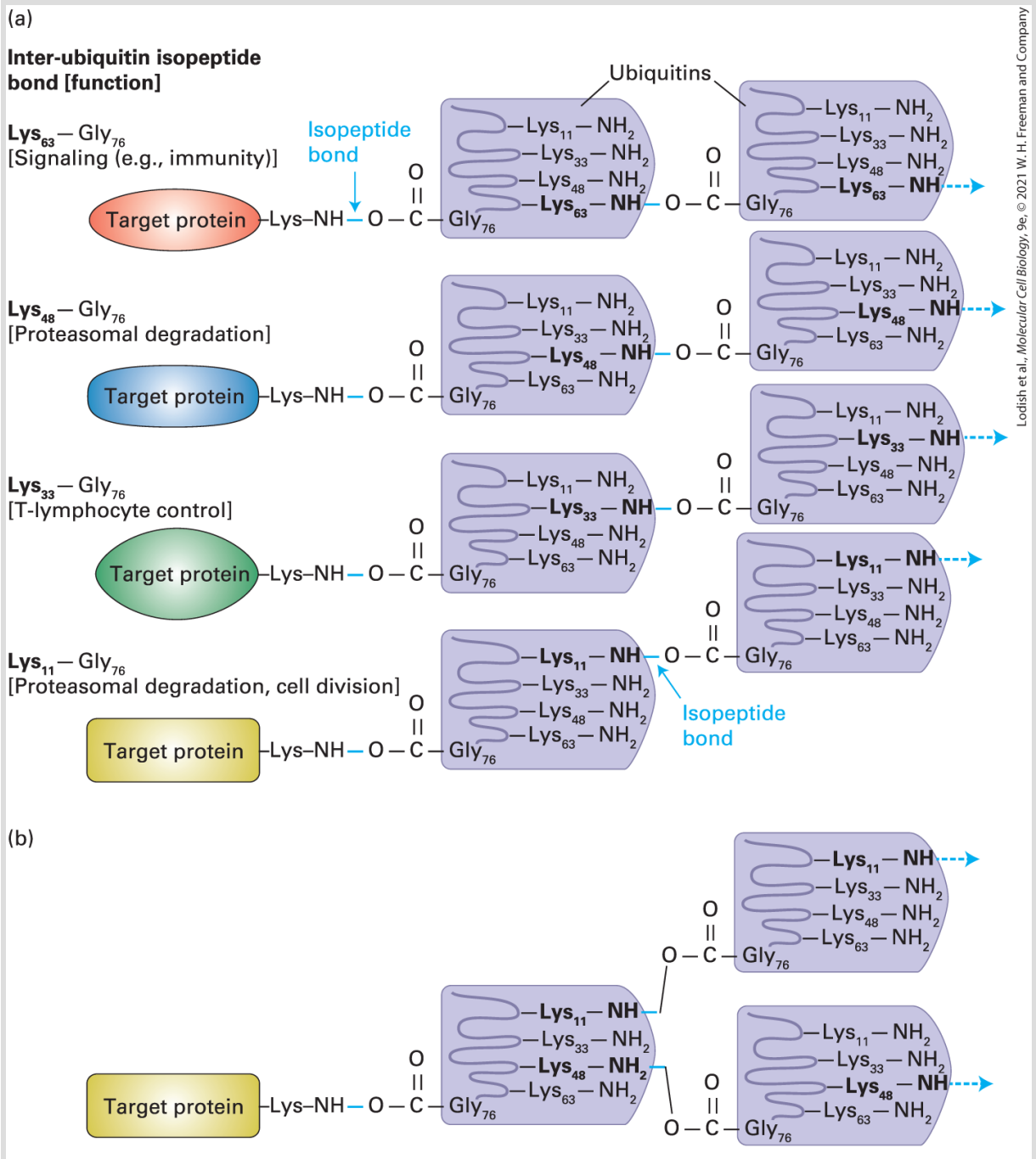


FIGURE 3-39 Determination of polyubiquitin function by the lysine used in isopeptide bonds linking ubiquitins. (a) Different ubiquitin ligases catalyze polyubiquitinylation of distinct target proteins (colored ovals) using distinct lysine side chains of ubiquitin molecules (purple) to generate the inter-ubiquitin isopeptide linkages (blue) with Gly-76 of the adjacent ubiquitin. Dotted blue arrows represent additional ubiquitins in the chain that are not shown. The lysine used for the isopeptide bonds determines the function of the

polyubiquitinylation. For example, polyubiquitins with Lys-48:Gly-76 isopeptide bonds direct the target to proteasomes for degradation. Those that use Lys-63, Lys-33, and Lys-11 influence signaling, T-lymphocyte control, and cell division, respectively. Isopeptide bonds involving ubiquitin's Lys-6, Lys-27, and Lys-29 and bonds involving its N-terminal amino group (not shown) can also be used to generate polyubiquitin chains. (b) Branched polyubiquitin chains can be generated by the covalent attachment of two ubiquitin molecules to two different amino groups on a single ubiquitin, in this case Lys_{48} and Lys_{11}.

The multiple forms of ubiquitinylation result in the generation of a wide variety of recognition surfaces that can participate in many protein-protein interactions with the hundreds of proteins (>200 in humans) that contain more than a dozen distinct ubiquitin-binding domains (UBD). In addition, any given polyubiquitin chain has the potential to bind simultaneously to more than one UBD-containing protein, leading to the formation of multiprotein complexes. Some deubiquitinases can remove an intact polyubiquitin chain from a modified protein ("anchored" chain) and thus generate a polyubiquitin chain not covalently linked to another protein ("unanchored" chain). Even these unanchored chains may serve a regulatory role.

With this great structural diversity, it is not surprising that cells use ubiquitinylation and deubiquitinylation to control many different functions.

Such functions include cellular internalization of molecules via endocytosis (see Chapter 14), repair of damaged DNA, metabolism, messenger RNA synthesis, defense against pathogens, cell division

and cell cycle progression, cell signaling pathways, trafficking of proteins within a cell, and apoptosis. The lysine used to form the inter-ubiquitin isopeptide bonds can vary depending on the cellular system that is regulated (Figure 3-39). For example, we have already seen that polyubiquination via Lys-48 is used to target proteins for proteasomal degradation, and there is evidence that polyubiquitinylation via other Lys residues (e.g., Lys-11 and Lys-33, but not Lys-63) can also target proteins to the proteasome. Polyubiquitinylation with Lys-63 linkages is used in many cellular identification and signaling systems, such as those that recognize the presence of intracellular viruses and bacteria and induce a protective immune response, as well as those that direct these pathogens to lysosomes for degradation. Lys-11-linked polyubiquitin chains regulate cell division. Lys-33-linked chains help suppress the activity of receptors on specialized white blood cells, called *T lymphocytes* (see Chapter 24) and thereby control the activity and function of those lymphocytes. Furthermore, additional post-translational modifications (phosphorylation, acetylation, etc.) of the ubiquitin molecules that are attached to target proteins can confer even greater capacity to manipulate the activity of the target protein.

Proteolytic Cleavage Irreversibly Activates or Inactivates Some Proteins

Unlike phosphorylation and ubiquitinylation, which are reversible, the activation or inactivation of protein function by proteolytic cleavage is an irreversible mechanism for regulating protein activity. For example, many polypeptide hormones, such as insulin, are synthesized as longer precursors, and prior to secretion from cells some of their peptide bonds must be hydrolyzed for them to fold properly. In some cases, a single, long precursor *prohormone* polypeptide is cleaved into several distinct active hormones. To prevent the pancreatic serine proteases from inappropriately digesting proteins before they reach the small intestine, they are synthesized as *zymogens*, inactive precursor enzymes. Cleavage of a peptide bond near the N-terminus of trypsinogen (the zymogen of trypsin) by a highly specific protease in the small intestine generates a new N-terminal residue (Ile-16) whose amino group can form an ionic bond with the carboxylic acid side chain of an internal aspartic acid. This binding causes a conformational change that opens the substrate-binding site, activating the enzyme. The active trypsin can then activate trypsinogen, chymotrypsinogen, and other zymogens. Similar but more elaborate protease cascades (with one protease activating inactive precursors of others) that can amplify an initial signal play important roles in several systems, such as the blood-clotting cascade and the complement system (see Chapter 24). The importance of carefully regulating such systems is clear — inappropriate clotting, for example, could fatally clog the circulatory system, while insufficient clotting could lead to uncontrolled bleeding.

An unusual and rare type of proteolytic processing, termed *protein self-splicing*, takes place in bacteria and some eukaryotes. This process is analogous to editing film: an internal segment of a polypeptide is removed and the ends of the polypeptide are rejoined (ligated). Unlike other forms of proteolytic processing, protein self-splicing is an autocatalytic process that proceeds by itself without the participation of other enzymes. The excised peptide appears to eliminate itself from the protein by a mechanism similar to that used in the processing of some RNA molecules (see Chapter 9). In vertebrate cells, the processing of some proteins includes self-cleavage, but the subsequent ligation step is absent. One such protein is Hedgehog, a membrane-bound signaling molecule that is critical to a number of developmental processes (see Chapter 16).

Higher Order Regulation Includes Control of Protein Location

All the regulatory mechanisms heretofore described affect a protein locally at its site of action, altering the protein's concentration or turning its activity on or off. Normal functioning of a cell, however, also requires the segregation of proteins to particular compartments, such as the mitochondria, nucleus, or lysosomes. In regard to enzymes, compartmentation not only provides an opportunity for controlling the delivery of substrate or the exit of product but also permits competing reactions to take place

simultaneously in different parts of a cell. We describe the mechanisms that cells use to direct various proteins to different compartments in Chapters 13 and 14. We now turn from the fundamental properties of proteins to the methods cell biologists use to study them.

KEY CONCEPTS OF SECTION 3.4

Regulating Protein Function

- Proteins may be regulated at the level of protein synthesis, protein degradation, or the intrinsic activity of proteins through noncovalent or covalent interactions.
- The life span of an intracellular protein is largely determined by its susceptibility to proteolytic degradation.
- Many proteins are marked for destruction with a covalently attached polyubiquitin tag assembled by ubiquitin ligases and then degraded within proteasomes, large cylindrical complexes with multiple protease active sites in their interior chambers (see Figure 3-32).
- Ubiquitinylation of proteins is reversible due to the activity of deubiquitinylating enzymes.
- In allostery, the noncovalent binding of one ligand molecule, the allosteric effector, induces a conformational change that alters a protein's activity or affinity for other ligands. The allosteric effector can be identical in structure to or different from the other ligands, whose binding it affects. The allosteric effector can be an activator or an inhibitor.
- In multimeric proteins, such as hemoglobin, that bind multiple identical ligand molecules (e.g., oxygen), the binding of one ligand molecule may increase or decrease the protein's affinity for subsequent ligand molecules. This type of allostery is known as *cooperativity* (see Figure 3-33).
- Several allosteric mechanisms act as switches, turning protein activity on and off in a reversible fashion.
- Noncovalent and covalent binding of ions and molecules to target proteins can regulate the activities of those target proteins.
- Two classes of intracellular switch proteins regulate a wide variety of cellular processes: (1) Ca^{2+}-binding proteins (e.g., calmodulin) and (2) members of the

GTPase superfamily (e.g., Ras), which cycle between active GTP-bound and inactive GDP-bound forms (see Figure 3-35). GTPases participate in complex regulatory networks that include proteins (GAP, GEF, GDI) that regulate the cycling of the GTPase between its active and inactive forms.

- The phosphorylation and dephosphorylation of hydroxyl groups on serine, threonine, or tyrosine side chains by protein kinases and phosphatases provide reversible on/off regulation of numerous proteins (see Figure 3-36).
- Most protein kinases use a common bilobed catalytic domain and a common catalytic mechanism to phosphorylate their substrate proteins.
- Ubiquitinylation is used to control a wide variety of cellular functions in addition to proteasome-mediated degradation, such as changes in the location or activity of proteins (see Figure 3-39). The specific covalent attachment of ubiquitin to proteins (mono-, multi-, and polyubiquitinylation involving a variety of linkages between the ubiquitin monomers) can vary depending on the cellular system that is regulated.
- Many types of covalent and noncovalent regulation are reversible, but some forms of regulation, such as proteolytic cleavage, are irreversible.
- Higher order regulation includes the intracellular location, or compartmentation, of proteins.

3.5 Purifying, Detecting, and Characterizing Proteins

A protein often must be purified before its structure and the mechanism of its action can be studied in detail. However, because proteins vary in size, shape, oligomerization state, charge, and water solubility, no single method can be used to isolate all proteins. To isolate one particular protein from the estimated 10,000 different proteins in a particular type of cell is a daunting task that requires methods both for separating proteins and for detecting the presence of specific proteins.

Any molecule, whether protein, carbohydrate, or nucleic acid, can be separated, or *resolved*, from other molecules on the basis of their differences in one or more physical or chemical characteristics. The larger and more numerous the differences between two proteins, the easier and more efficient their separation. As a practical matter, the more abundant a particular protein is in a biological sample, the easier it is to separate it from the other molecules in the sample. The three most widely used characteristics for separating proteins are *size*, defined as either length or mass; net electrical charge; and *affinity* for binding to specific ligands. In this section, we briefly outline several important techniques for separating proteins; these separation techniques are also useful for the separation of nucleic acids and other biomolecules. (Specialized methods for removing membrane proteins from membranes are described in Chapter 10

after the unique properties of these proteins are discussed.) We then consider the use of radioactive compounds for tracking biological activity. Finally, we consider several techniques for characterizing a protein's mass, sequence, and three-dimensional structure.

Centrifugation Can Separate Particles and Molecules That Differ in Mass or Density

The first step in a typical protein-purification scheme is centrifugation. The principle behind centrifugation is that two types of particles in suspension (cells, cell fragments, organelles, or molecules) with different masses or densities will settle to the bottom of a test tube at different rates. Remember, mass is the weight of a sample (measured in daltons or molecular weight units), whereas density is the ratio of its mass to volume (often expressed as grams per liter because of the methods used to measure density). Proteins vary greatly in mass, but not in density. Unless a protein has an attached lipid or carbohydrate, its density will not vary by more than 15 percent from $1.37\ \mathrm{g/cm^3}$, the average protein density. Heavier or denser molecules settle, or *sediment*, more quickly than lighter or less dense molecules.

A centrifuge speeds sedimentation by subjecting particles in suspension to centrifugal forces as great as 1 million times the force

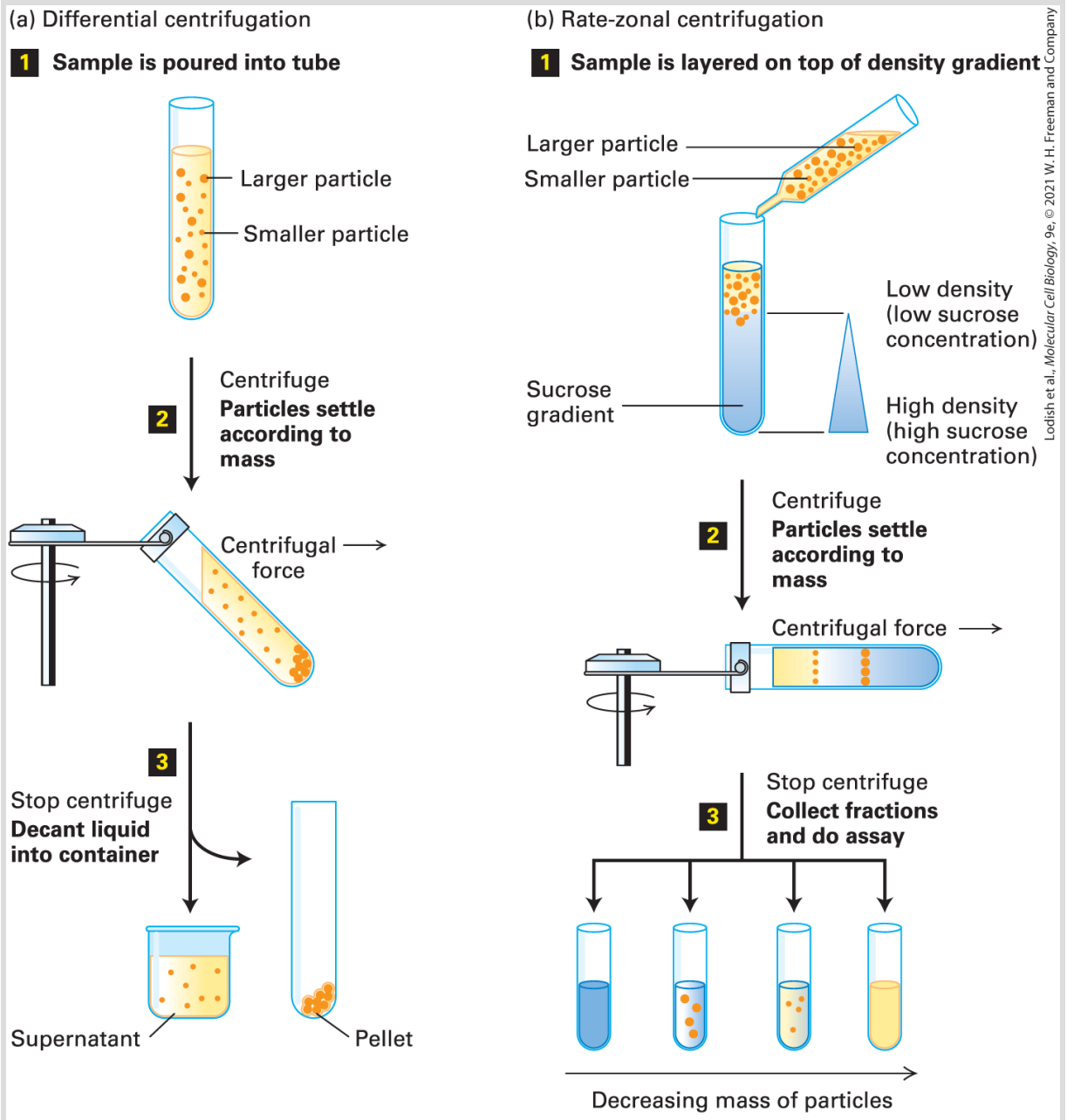
of gravity, *g*, which can sediment particles as small as 10 kDa. Modern ultracentrifuges achieve these forces by reaching speeds of 150,000 revolutions per minute (rpm) or greater. However, small particles with masses of 5 kDa or less will not sediment uniformly even at such remarkably high rotation rates. The extraordinary technical achievements of modern ultracentrifuges can be appreciated by considering that they can rotate a several-pound rotor (about the size of an American football) that holds the samples in tubes at rates as high as 2500 revolutions per second!

Centrifugation is used for two basic purposes: (1) as a preparative technique to separate one type of material from others with the goal of obtaining enough of the material to perform subsequent experiments and (2) as an analytical technique to measure physical properties (e.g., molecular weight, density, shape, and equilibrium-binding constants) of macromolecules. The sedimentation constant, *s*, of a protein is a measure of its sedimentation rate. The sedimentation constant is commonly expressed in Svedberg units (S), where a typical large protein complex is about 3–5S, a proteasome is 26S, and a eukaryotic ribosome is 80S.

Differential Centrifugation

The most common initial step in protein purification from cells or tissues is the separation of water-soluble proteins from insoluble cellular material by *differential centrifugation*. A starting mixture,

commonly a cell homogenate (mechanically broken cells), is poured into a tube and spun at a rotor speed, and for a period of time that forces cell organelles such as nuclei as well as large unbroken cells or large cell fragments to collect as a pellet at the bottom; the soluble proteins remain in the liquid in the tube above the pellet, which is called the *supernatant* (Figure 3-40a). The supernatant then is poured off, and either it or the pellet can be subjected to other purification methods to separate the many different proteins that they contain.



EXPERIMENTAL FIGURE 3-40 Centrifugation techniques separate particles that differ in mass or density. (a) In differential centrifugation, a cell homogenate or other mixture is spun long enough to sediment the larger particles (e.g., cell organelles, cells), which collect as a pellet at the bottom of the tube (step 2). The smaller particles (e.g., soluble proteins, nucleic acids) remain in the liquid supernatant, which can be transferred to another tube (step 3). (b) In rate-zonal centrifugation, a mixture is spun (step 1) just long enough to separate molecules that differ in mass but may be similar in shape and density (e.g., globular proteins, RNA molecules) into discrete zones within a density gradient commonly

formed by a concentrated sucrose solution. Fractions are removed from the bottom of the tube and subjected to testing (assayed).

Rate-Zonal Centrifugation

On the basis of differences in their masses, water-soluble proteins can be separated by centrifugation through a solution of increasing density, called a *density gradient*. A concentrated sucrose solution is commonly used to form a density gradient in a centrifuge tube (with higher concentrations of sucrose, and thus a higher solution density, toward the bottom of the tube, Figure 3-40b). When a protein mixture is placed on top of a sucrose density gradient in a tube and subjected to centrifugation, each protein in the mixture migrates down the tube at a rate controlled by the protein's physical properties. All the proteins start from the thin layer of the sample that was placed at the top of the tube and separate into bands (actually, disks) of proteins of different masses as they travel at different rates through the gradient. In this separation technique, called *rate-zonal centrifugation*, samples are centrifuged just long enough to separate the molecules of interest into discrete bands, also called *zones* (Figure 3-40b). If a sample is centrifuged for too short a time, the different protein molecules will not separate sufficiently. If a sample is centrifuged much longer than necessary, all the proteins will end up mixed together at the bottom of the tube.

Although the sedimentation rate is strongly influenced by particle mass, rate-zonal centrifugation is seldom effective in determining precise molecular weights because variations in shape also affect the sedimentation rate. The exact effects of shape are hard to assess, especially for proteins or other molecules, such as single-stranded nucleic acid molecules, that can assume many complex shapes. Nevertheless, rate-zonal centrifugation has proved to be a practical method for separating many different types of polymers and particles. A second density-gradient technique, called *equilibrium density-gradient centrifugation*, is used mainly to separate DNA, lipoproteins that carry lipids through the circulatory system, or organelles (see Figure 4-37).

Electrophoresis Separates Molecules on the Basis of Their Charge-to-Mass Ratio

Electrophoresis, a technique for separating molecules in a mixture under the influence of an applied electric field, is one of the most frequently used techniques to study proteins and nucleic acids. Dissolved molecules in an electric field move, or migrate, at a speed determined by their charge-to-mass (charge:mass) ratio and the physical properties of the medium through which they migrate. For example, if two molecules have the same mass and shape, the one

with the greater net charge will move faster toward an electrode of the opposite polarity.

SDS-Polyacrylamide Gel Electrophoresis

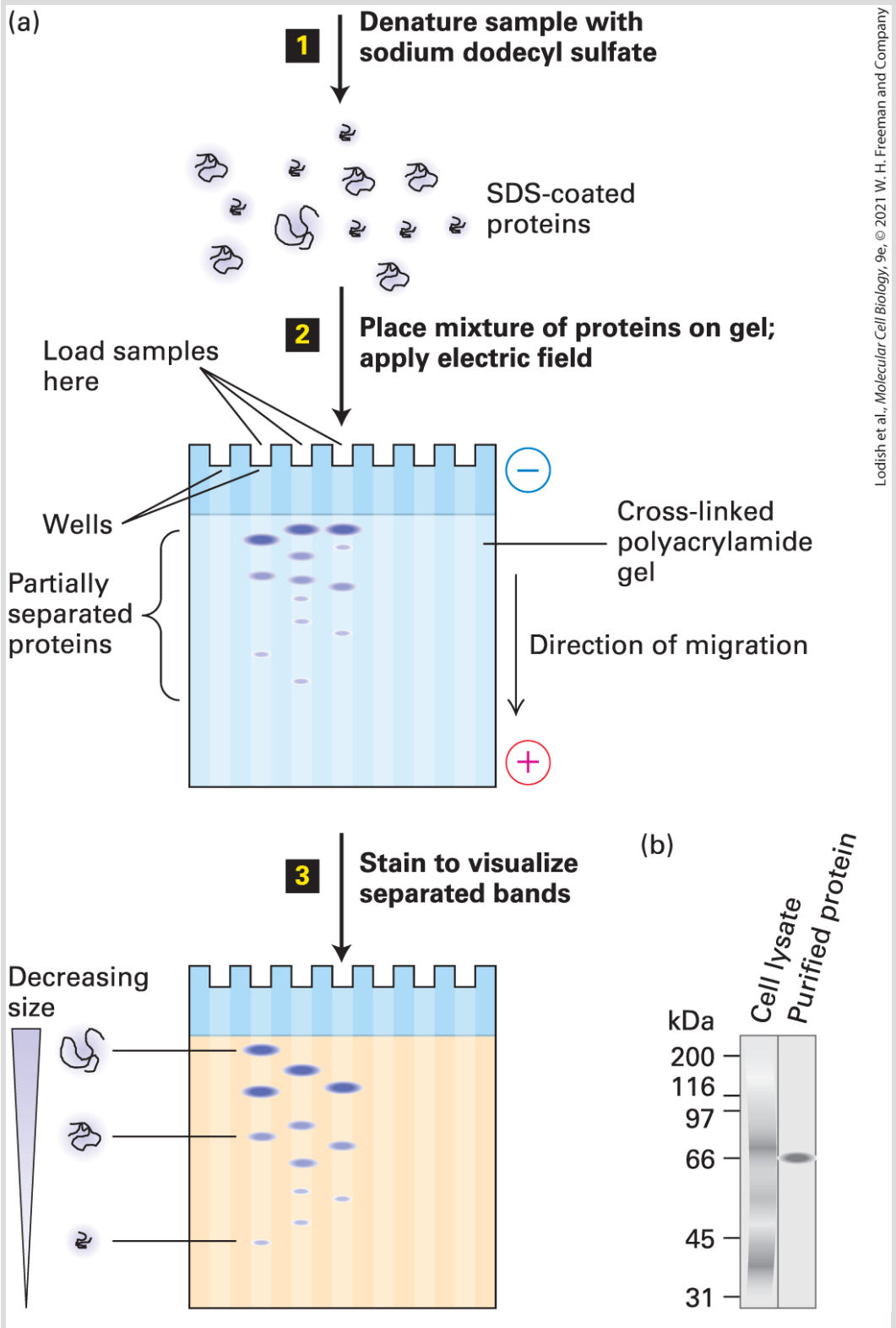
Because many proteins or nucleic acids that differ in size and shape have nearly identical charge:mass ratios, electrophoresis of these macromolecules in a liquid solution results in little or no separation of molecules of different lengths. However, successful separation of proteins and nucleic acids can be accomplished by electrophoresis in various gels (semisolid suspensions in water similar to the congealed gelatin found in desserts). These gels are commonly cast into flat, relatively thin slabs between a pair of glass plates. When a mixture of proteins is placed in a gel and an electric current is applied, the gel acts as a sieve, allowing smaller species to maneuver more rapidly through its pores than larger species do. The shape of a molecule can also influence its rate of migration (long, asymmetric molecules migrate more slowly than spherical ones of the same mass).

Electrophoretic separation of proteins is most commonly performed in polyacrylamide gels. These gels are made by polymerizing a solution of acrylamide monomers into polyacrylamide chains and simultaneously cross-linking the chains into a semisolid matrix. The pore size of a gel can be varied by adjusting the concentrations of polyacrylamide and the cross-linking reagent. The rate at which a protein moves through a gel is influenced by the gel's pore size and

the strength of the electric field. By suitable adjustment of these parameters, proteins of widely varying sizes can be resolved (separated from one another) by this technique, known as *p*oly*a*crylamide *g*el *e*lectrophoresis (PAGE).

In the most powerful technique for resolving protein mixtures, proteins are exposed to the ionic detergent SDS (sodium dodecyl sulfate) before and during gel electrophoresis (Figure 3-41). SDS denatures proteins, in part because it binds to hydrophobic side chains, destabilizing the hydrophobic interactions in the core of a protein that contribute to its stable conformation. SDS treatment is usually combined with heating, often in the presence of reducing agents that break disulfide bonds. Under these conditions, most multimeric proteins dissociate into their subunits. Typically, the amount of SDS that binds to the protein is proportional to the length of the polypeptide chain and relatively independent of the sequence. Two proteins of similar size will bind the same absolute quantity of SDS, whereas a protein twice that size will bind twice the amount of SDS. Denaturation of a complex protein mixture with SDS in combination with heat usually forces each polypeptide chain into an extended conformation and imparts on each of the proteins in the mixture a constant charge:mass ratio because the dodecyl sulfate, which is negatively charged, is the major contributor of charge. As the SDS-bound proteins move through the polyacrylamide gel, they are separated according to size by the sieving action of the gel. SDS treatment thus eliminates the effect of differences in native

conformation; therefore, chain length, which is proportional to mass, is the principal determinant of the migration rate of proteins in *SDS-polyacrylamide electrophoresis (SDS-PAGE)*. Even chains that differ in molecular weight by less than 10 percent can be resolved by this technique. Moreover, the molecular weight of a protein can be estimated by comparing the distance that it migrates through a gel with the distances that proteins of known molecular weight (called molecular weight "standards") migrate in the same gel (there is a roughly linear relationship between migration distance and the log of the molecular weight). Proteins within the gels can be extracted for further analysis (e.g., identification by the methods described below).



EXPERIMENTAL FIGURE 3-41 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins primarily on the basis of their masses. (a) Initial

treatment with SDS, a negatively charged detergent, dissociates multimeric proteins and denatures all the polypeptide chains (step **1**). During electrophoresis, the SDS-protein complexes migrate through the polyacrylamide gel (step **2**). Small complexes are able to move through the pores faster than larger ones. Thus the proteins separate into bands according to their sizes as they migrate. The separated protein bands are visualized by staining with a dye (step **3**). (b) Example of SDS-PAGE separation of all the proteins in a whole-cell lysate (detergent-solubilized cells). *Left*: The many separate stained proteins appear almost as a continuum. *Right*: A single protein purified from the lysate by a single step of antibody-affinity chromatography. The proteins were visualized by staining with a silver-based dye.

[Part (b) Data from B. Liu and M. Krieger, 2002, *J. Biol. Chem*. **277**:34125–34135.]

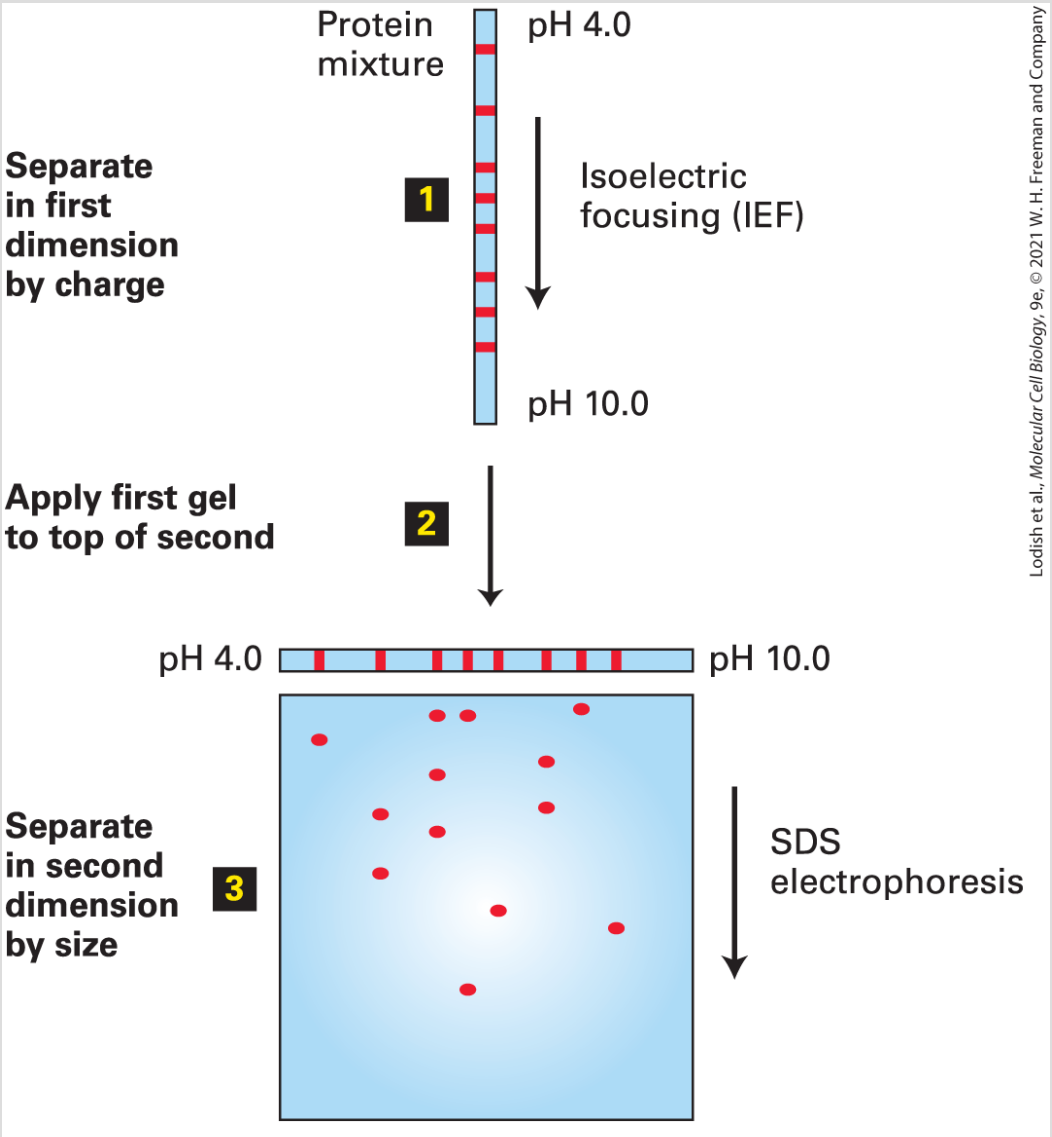
If two or more polypeptides are cross-linked by disulfide bonds, the protein's migration rate in SDS-PAGE will depend on whether or not the protein has been reduced to break those bonds prior to electrophoresis. The cross-linked proteins will appear larger than the individual, reduced subunits. By examining samples with and without reduction, one can identify such proteins and their component polypeptides.

Two-Dimensional Gel Electrophoresis

Electrophoresis of a complex biological mixture of proteins, for example, purified organelles or even whole cells or tissues, by SDS-PAGE can separate proteins having relatively large differences in mass, but cannot readily resolve proteins having similar masses (e.g., a 41-kDa protein versus a 42-kDa protein). To separate proteins of similar masses, another physical characteristic must be exploited.

Most commonly, this characteristic is electric charge, which is determined by the pH of the sample and by the relative number of the protein's positively and negatively charged groups, which is in turn dependent on the pK_a's of the ionizable groups (see Chapter 2) on the proteins (usually the amino and carboxyl termini and side chains such as those in lysine and aspartic acid). Two unrelated proteins having similar masses are unlikely to have identical net charges because their sequences, and thus the number of acidic and basic residues, are different.

In *two-dimensional gel electrophoresis*, proteins are separated sequentially, first by their charges using a very sensitive method called *isoelectric focusing (IEF)* and then by their masses using SDS electrophoresis. Isoelectric focusing (IEF) separates proteins on the basis of their **isoelectric points (pI**, Figure 3-42) and can clearly separate (resolve) proteins that differ by only one charge unit. Two-dimensional gel electrophoresis is sensitive enough to separate phosphorylated and nonphosphorylated versions of the same protein. Today sophisticated mass spectrometry methods, described below, are often used in place of two-dimensional gel electrophoresis, both to separate and to identify the protein components of a complex sample as well as to compare changes in the amounts of those components in different biological specimens.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

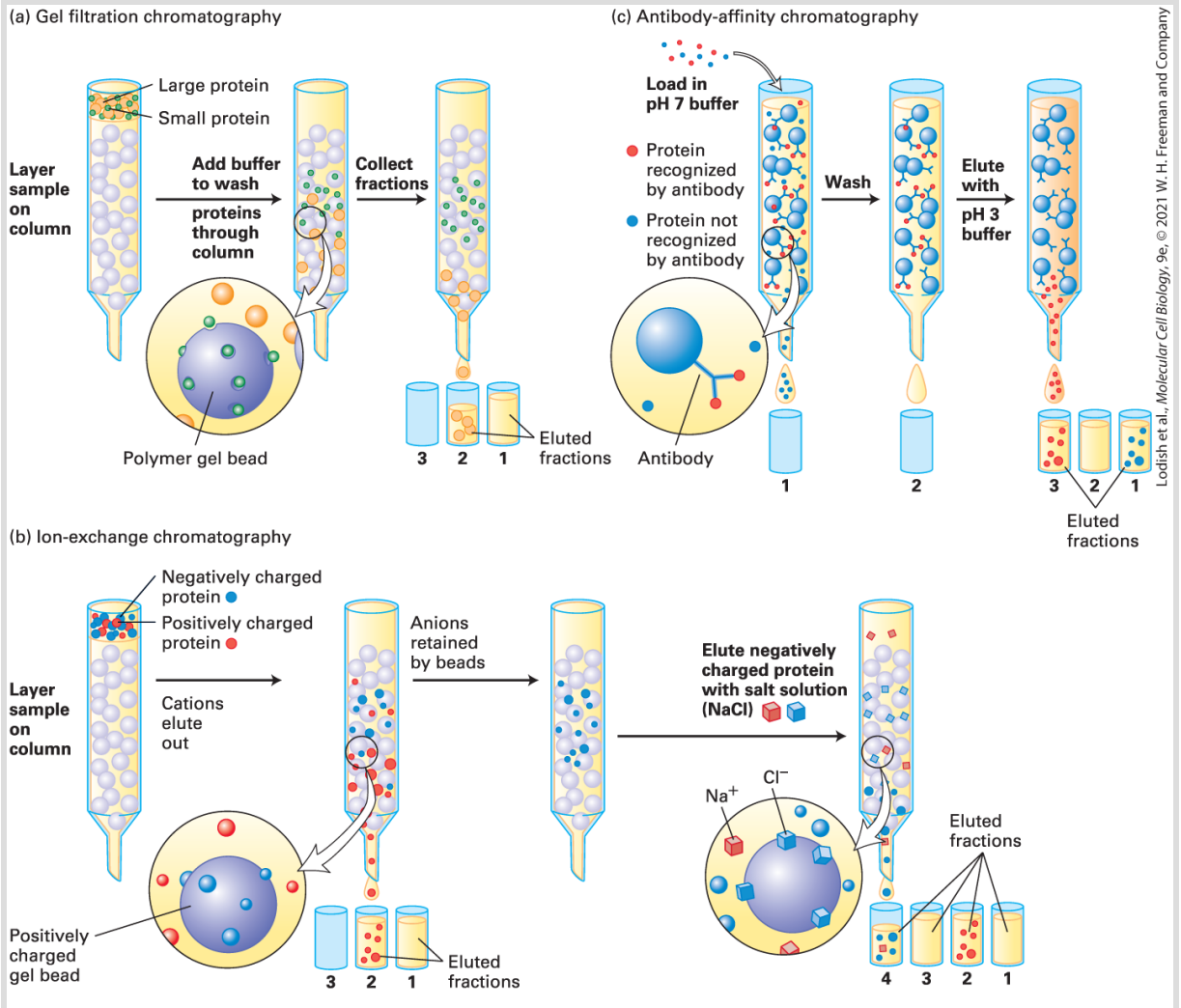
EXPERIMENTAL FIGURE 3-42 Two-dimensional gel electrophoresis separates proteins on the basis of charge and mass. In this technique, proteins are first separated into bands on the basis of their charges by isoelectric focusing (IEF, step 1). For the first step using IEF, a cell or tissue extract is fully denatured by high concentrations (8 M) of urea (and sometimes SDS) and then layered on a strip of gel that contains urea, which removes any bound SDS. The gel contains a continuous pH gradient along its length ranging from pH 3 to pH 10 formed by ampholytes, which are polyanionic, and polycationic small molecules that are cast into the gel. When an electric field is applied to the gel, the ampholytes will migrate. Ampholytes with an excess of negative charges will migrate toward the anode (upward in step 1), where they establish an acidic pH (many protons),

while ampholytes with an excess of positive charges will migrate toward the cathode, where they establish an alkaline pH. The careful choice of the mixture of ampholytes and careful preparation of the gel allow the construction of stable pH gradients ranging from pH 3 to pH 10. A charged protein placed at one end of such a gel will migrate through the gradient under the influence of the electric field until it reaches its isoelectric point (pI), the pH at which the net charge of the protein is zero. With no net charge, the protein will migrate no further, forming a band or thin strip in the gel strip. To achieve the second dimension of separation, the resulting gel strip is applied to an SDS-polyacrylamide gel (step 2) to permit separation of the proteins on the basis of their molecular weights. The IEF gel strip is placed lengthwise on one outside edge of a square or rectangular slab of polyacrylamide gel, this time saturated with SDS to confer on each separated protein a more or less constant charge:mass ratio. When an electric field is imposed, the proteins will migrate from the IEF gel into the SDS gel and then separate according to their masses to form spots in two dimensions defined by their pI and molecular weight (step 3). The polypeptide spots can be visualized either by staining with dyes or, if the proteins are radioactive, by autoradiography.

Liquid Chromatography Resolves Proteins by Mass, Charge, or Affinity

A third common technique for separating mixtures of proteins or fragments of proteins, as well as other molecules, is based on the principle that molecules in solution can differentially interact with (bind to and dissociate from) a particular solid surface, depending on the physical and chemical properties of the molecule and the surface. If the solution is allowed to flow across the surface, then molecules that interact frequently with the surface will spend more

time bound to (or retained at) the surface, and thus flow past the surface more slowly, than molecules that interact infrequently with it. In this technique, called **liquid chromatography (LC)**, the sample is placed on top of a tightly packed column of spherical beads held within a glass, metal, or plastic cylinder (Figure 3-43). The sample then flows down the column, driven by gravitational or hydrostatic forces alone or sometimes with the assistance of a pump. In some LC systems, the composition of the fluid flowing out of the column is monitored continuously (e.g., by spectroscopy). Small aliquots of fluid flowing out of the column, called *fractions*, are collected sequentially and can be analyzed subsequently for their contents and chemical activities (e.g., enzymatic activity). The nature of the beads in the column determines whether the separation of proteins depends on differences in their mass, charge, or other binding properties (e.g., affinity for substances attached to the beads).



EXPERIMENTAL FIGURE 3-43 Three commonly used liquid chromatographic techniques separate proteins on the basis of mass, charge, or affinity for a specific binding partner. (a) Gel filtration chromatography separates proteins that differ in size. A mixture of proteins is carefully placed, or loaded, on the top of a cylinder packed with porous beads. Smaller proteins travel through the column more slowly than larger proteins. Thus the different proteins emerging in the eluate flowing out of the bottom of the column at different times (different elution volumes) can be collected in separate tubes, called *fractions*. (b) Ion-exchange chromatography separates proteins that differ in net charge in columns packed with beads that carry either a positive charge (shown here) or a negative charge. Proteins having the same net charge as the beads are repelled and flow through the column, whereas proteins having the opposite charge bind to the beads more or less tightly, depending on their structures. Bound proteins — in this case, negatively charged proteins — are subsequently eluted by passing a salt gradient (usually of NaCl or KCl) through the column.

As the ions bind to the beads, they displace the proteins; more tightly bound proteins require higher salt concentrations in order to be released. (c) In antibody-affinity chromatography, a mixture of proteins is passed through a column packed with beads to which a specific antibody is covalently attached. Only proteins with high affinity for the antibody are retained by the column; all the nonbinding proteins flow through. After the column is washed, the bound protein is eluted with an acidic solution or some other solution that disrupts the antigen-antibody complexes; the released protein then flows out of the column and is collected.

Gel Filtration Chromatography

Proteins that differ in mass can be separated on a column of porous beads made from polyacrylamide, dextran (a bacterial polysaccharide), or agarose (a seaweed derivative) — a technique called *gel filtration chromatography*. Although proteins flow around the beads, they spend some time within the large depressions that cover a bead's surface. Because smaller proteins can penetrate these depressions more readily than larger proteins can, they travel through a gel filtration column more slowly than do larger proteins (Figure 3-43a). (In contrast, proteins migrate *through* the pores in an electrophoretic gel; thus smaller proteins move faster than larger ones.) The total volume of liquid required to elute (or separate and remove) a protein from a gel filtration column depends on the protein's mass: the smaller its mass, the longer it is trapped on the beads, the longer it takes to traverse the column, and the greater the elution volume. If proteins of known mass are used as standards to calibrate the column, the elution volume can be used to estimate the mass of a protein in a mixture. A protein's shape, whether is it

elongated or compact, as well as its mass can influence the elution volume.

Ion-Exchange Chromatography

In ion-exchange chromatography, proteins are separated on the basis of differences in their charges. This technique makes use of specially modified beads whose surfaces are covered by amino groups or carboxyl groups and thus carry either a positive charge (NH_3^+) or a negative charge (COO^-) at neutral pH.

The proteins in a mixture carry various net charges at any given pH. When a solution of mixed proteins flows through a column of positively charged beads, only proteins with a net negative charge (acidic proteins) adhere to the beads; neutral and positively charged (basic) proteins flow unimpeded through the column (Figure 3-43b). The acidic proteins are then eluted selectively from the column by passing a solution of increasing concentrations of salt (a salt gradient) through the column. At low salt concentrations, protein molecules and beads are attracted by their opposite charges. At higher salt concentrations, negatively charged salt ions bind to the positively charged beads, displacing the negatively charged proteins. In a gradient of increasing salt concentrations, weakly bound proteins — those with a relatively low charge — are eluted first, and highly charged proteins are eluted last. Similarly, a negatively

charged column can be used to retain and fractionate basic (positively charged) proteins.

Affinity Chromatography

The ability of proteins to bind specifically to other molecules is the basis of affinity chromatography. In this technique, ligands or other molecules (affinity reagents) that bind to the protein of interest are covalently attached to the beads used to form the column. Ligands can be enzyme substrates, inhibitors or their analogs, or other small molecules that bind to specific proteins. In a widely used form of this technique — *antibody-affinity*, or *immunoaffinity, chromatography* — the molecule attached to the beads is an antibody specific for the desired protein (Figure 3-43c). (We discuss antibodies as tools for studying proteins next; see also Chapter 24, which describes how antibodies are made.)

In principle, an affinity column will retain only those proteins that bind the molecule attached to the beads; the remaining proteins, regardless of their charges or masses, will pass through the column because they do not bind. In this way, a single type of protein can be isolated from a very complex biological mixture, including whole cells or tissues, in a single step. However, if a retained protein is in turn bound to other molecules, forming a complex, then the entire complex is retained on the column. The proteins bound to the affinity column are then eluted by adding an excess of a soluble

form of the ligand, by exposure of bound materials to detergents, or by changing the salt concentration or pH such that the binding to the molecule on the column is disrupted. The ability of this technique to separate particular proteins depends on the selection of appropriate binding partners that bind more tightly to the protein of interest than to other proteins.

One variation of affinity chromatography is to covalently modify a protein of interest with some compound that in turn can bind to an affinity reagent on a bead. For example, a small molecule such as biotin (vitamin B) can be chemically cross-linked to a protein and then that modified protein can be affinity purified by a column whose beads are linked to the bacterial protein streptavidin that binds biotin exceedingly tightly. Alternatively, a gene can be engineered to express as a single polypeptide chain a chimeric, or fusion, protein comprising a protein of interest together with a peptide or protein that can bind tightly to an affinity reagent attached to a bead. Fusion proteins containing the enzyme glutathione-S-transferase (GST) can be affinity purified using beads linked to the small molecule glutathione. Fusion proteins containing a short peptide of polyhistidine residues can be purified using beads linked to divalent metals such as Ni^{2+} or Zn^{2+}. Fusion proteins containing a short peptide can be purified using beads linked to an antibody that binds to that peptide.

Highly Specific Enzyme and Antibody Assays Can Detect Individual Proteins

The purification of a protein, or any other molecule, requires a specific *assay* that can detect the presence of that molecule as it is separated from other molecules (e.g., in column or density-gradient fractions or gel bands or spots). Such an assay capitalizes on some highly distinctive characteristics of a protein: the ability to bind a particular ligand, to catalyze a particular reaction, or to be recognized by a specific antibody. The assay must also be simple and fast to minimize errors and the possibility that the protein of interest will become denatured or degraded while the assay is being performed. The goal of any purification scheme is to isolate sufficient amounts of a given protein for study; thus a useful assay must also be sensitive enough that only a small proportion of the available material is consumed by it. Many common protein assays require just 10^{-9} to 10^{-12} g of material.

Chromogenic Enzyme Reactions

Many assays are tailored to detect some functional aspect of a protein. For example, assays of enzymatic activity are based on the ability to detect the loss of substrate or the formation of product. Some enzymatic activity assays use chromogenic substrates, which change color in the course of the reaction. (Some substrates are naturally chromogenic; those that are not can be linked to a

chromogenic molecule.) Because of the specificity of an enzyme for its substrate, only samples that contain the enzyme will change color in the presence of a chromogenic substrate; the rate of the change provides a measure of the quantity of enzyme present. Enzymes that catalyze chromogenic reactions can also be fused or chemically linked to an antibody and used to report the presence or location of an antigen to which the antibody binds (see the next page).

Antibody Assays

As noted earlier, antibodies have the distinctive characteristic of binding tightly and specifically to antigens. As a consequence, preparations of antibodies that recognize a protein antigen of interest can be generated and used to detect the presence of that protein, either in a complex mixture of other proteins (finding a needle in a haystack, as it were) or in a partially or completely purified preparation of a particular protein. The presence of the antigen can be detected by labeling the antibody with an enzyme, a fluorescent molecule, or a radioactive isotope, which can be detected using an enzyme assay, fluorescence microscopy or spectroscopy, or a radiation detector, respectively. For example, luciferase, an enzyme present in fireflies and some bacteria, can be linked to an antibody. In the presence of ATP and its substrate, luciferin, luciferase catalyzes a light-emitting reaction. In either case, after the antibody binds to the protein of interest (the antigen) and unbound antibody is washed away, substrates of the linked

enzyme are added and the appearance of color or emitted light is monitored. The intensity is proportional to the amount of enzyme-linked antibody, and thus antigen, in the sample. Alternatively, after a first (or "primary") antibody that is not otherwise labeled binds to its target protein, a second ("secondary"), labeled antibody that can recognize the first antibody is used to bind to the complex of the first antibody and its target. This combination of primary and secondary antibodies (sometimes called an antibody "sandwich") permits very high sensitivity in the detection of a target protein because the labeled secondary antibody is often a mixture of antibodies that bind to multiple sites on the first antibody and thus results in a stronger signal than would labeling of the primary antibody alone. It is important to remember that an antibody recognizes and binds to only its epitope on a target antigen. If that epitope is altered — for example, by partial unfolding or post-translational modifications — or is blocked when the antigen protein is bound to some other molecule, the ability of the antibody to bind may be reduced or completely lost. Thus the absence of antibody binding does not necessarily mean that the antigen is not present in a sample, only that the epitope portion of that antigen is not present or accessible for antibody binding.

To generate antibodies to a protein (discussed in detail in Chapter 24), the intact protein, or a fragment of the protein, is injected into an animal (usually a rabbit, mouse, or goat). Sometimes a short synthetic peptide of 10–15 residues based on the sequence of the protein of interest is injected as the antigen to induce antibody

formation. Such a synthetic peptide, when coupled to a large protein carrier, can induce an animal to produce antibodies that bind specifically to that part (the epitope) of the full-sized, natural protein. Biosynthetically or chemically attaching the epitope to an unrelated protein is called *epitope tagging*. As we'll see throughout this book, antibodies generated using either synthetic peptide epitopes or intact proteins are extremely versatile reagents for isolating, detecting, and characterizing proteins.

Detecting Proteins by Attaching Green Fluorescent Protein

An alternative to epitope tagging that is particularly useful in detecting specific proteins within live cells makes use of *green fluorescent protein (GFP)*, a naturally fluorescent protein found in jellyfish (see Figure 4-16). A chimeric protein containing both the protein of interest and GFP, linked together in one polypeptide chain, is expressed in cells by introducing into the cells a gene encoding the combined protein. The amounts and intracellular distribution of the chimeric protein can then be determined readily. This chimeric protein approach is described in Chapter 4.

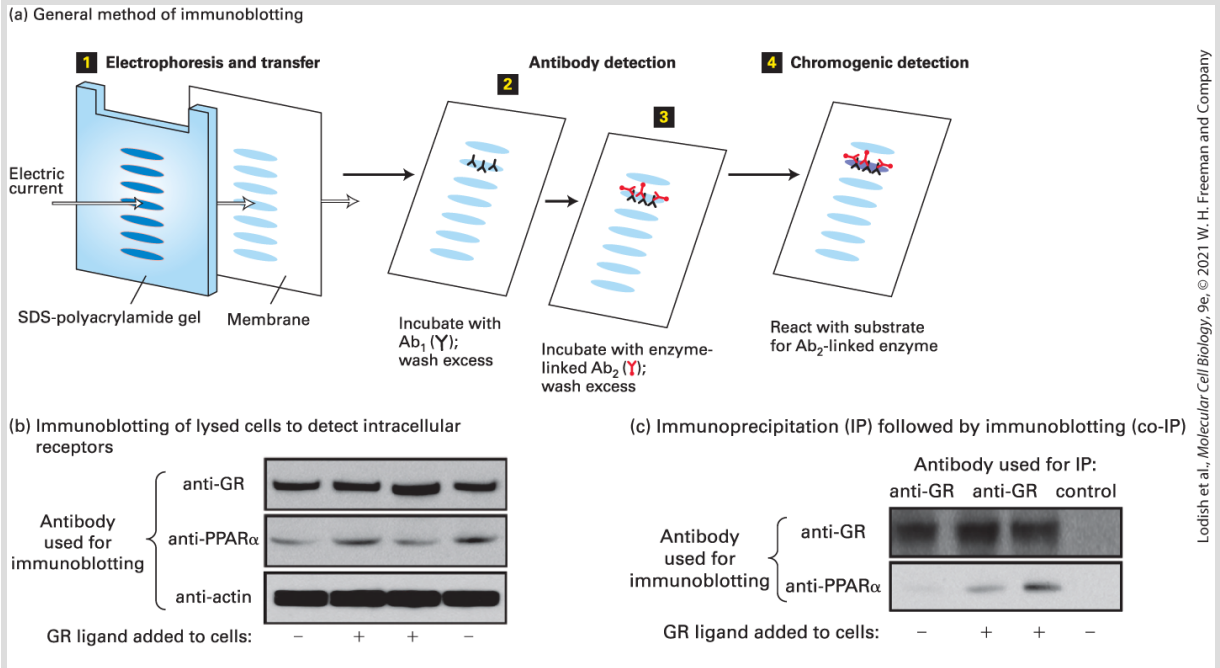
Detecting Proteins in Gels

Proteins embedded within a gel usually are not visible. The two general approaches to detecting proteins in gels are either to label

or stain the proteins while they are still within the gel or to electrophoretically transfer the proteins to a membrane made of nitrocellulose or polyvinylidene difluoride and then detect them. Proteins within gels are usually stained with an organic dye or a silver-based stain, both of which are detectable with normal visible light, or with a fluorescent dye, which requires specialized detection equipment. Coomassie blue, the most commonly used organic dye, is typically used to detect about 1000 ng of protein, with a lower limit of detection of about 4–10 ng. Silver staining and fluorescence staining are more sensitive (with a lower limit of ~1 ng). Coomassie and other stains can also be used to visualize proteins after transfer to membranes; however, the most common method of visualizing proteins in membranes is immunoblotting.

Immunoblotting, also called *Western blotting*, combines the resolving power of gel electrophoresis with the specificity of antibodies. This multistep procedure is commonly used to separate proteins and then identify a specific protein of interest. As shown in Figure 3-44, two different antibodies are used, one that is specific for the protein of interest (primary antibody) and a secondary antibody that binds to the first and is linked to an enzyme or other molecule that permits detection of the first antibody (and thus the protein of interest to which it binds). The enzyme to which the second antibody is attached can either generate a visible colored product or, by a process called *chemiluminescence*, produce light that can readily be recorded by film or a sensitive electronic detector. An example of

the results of an immunoblotting experiment can be seen in Figure 15-10. If an antibody to the protein of interest is not available, but the gene encoding the protein is available and can be used to express the protein, then recombinant DNA methods (see Chapter 5) can incorporate a small peptide epitope into the normal sequence of the protein (epitope tagging) that can be detected by a commercially available antibody to that epitope.



EXPERIMENTAL FIGURE 3-44 Immunoblotting (IB, or Western blotting) and co-immunoprecipitation (co-IP) can detect specific proteins and their binding partners. (a) Immunoblotting method. Step 1: After a protein mixture has been electrophoresed through an SDS gel, the separated bands (or spots, for two-dimensional gel electrophoresis) are transferred (blotted) from the gel onto a porous membrane. Individual proteins (blue ovals) are not visible at this stage. Step 2: The membrane is flooded with a solution of an antibody (Ab_1) specific for the protein of interest and allowed to incubate. Ab_1 binds to the protein of interest (second from the top), but not to any other proteins attached to the membrane, forming a layer of antibody molecules coincident with the protein (whose position still cannot be seen at this point). The membrane is washed to remove unbound

Ab_1. Step **3**: The membrane is incubated with a second antibody (Ab_2) that specifically recognizes and binds to the first (Ab_1). This second antibody is covalently linked to an enzyme that catalyzes a chromogenic reaction or releases light (e.g., chemiluminescence), a radioactive isotope, or some other substance whose presence can be detected with great sensitivity. Step **4**: Finally, the location and amount of bound Ab_2 are detected (e.g., by its color for a chromogenic reaction or by detectors or film that measure the light released by chemiluminescence). The mass of the protein of interest is determined by its electrophoretic mobility and its quantity by the band's intensity. (b) Immunoblotting was used to detect intracellular receptors and the influence of exposure to a ligand for one of the receptors. In this experiment, cells that are precursors to red blood cells were maintained in vitro in petri dishes and then treated with either no ligand ($-$, leftmost and rightmost lanes) or a ligand that binds to GR, the glucocorticoid receptor ($+$, duplicate samples in the two center lanes). The cells were then lysed in detergent, and immunoblotting was performed on the total cell lysates using three different antibodies: to the GR receptor (anti-GR), to a receptor called *PPARα* (anti-PPARα), or to actin, an abundant intracellular protein whose abundance was not expected to be sensitive to treatment with the ligand. [In these experiments, to compare the intensities of the bands in different lanes of the gel (representing the abundance of the corresponding protein in each sample), equal amounts of cell lysate must be loaded into each lane. The "loading control" in this experiment is the amount of actin seen in each lane. Because the intensities of the bands detected using the anti-actin antibody (bottom row) were equal, we can conclude that equal amounts of the different cell lysates were loaded into each lane of the gel. Thus it is valid to compare the band intensities between lanes.] The approximately equal intensities of the bands for both GR and PPARα with or without prior incubation of the cells with the GR ligand showed that the ligand did not substantially alter the amounts of either of these proteins in the cells. Portions of the same cell lysates used for the immunoblotting in part (b) were also used for the immunoprecipitation/immunoblotting shown in part (c). (c) Immunoprecipitation (IP) followed by immunoblotting (together called *co-IP*) was used to determine if the GR ligand can induce formation of a stable complex that contains both GR and PPARα. Portions of the cell lysates were immunoprecipitated with an antibody to GR (left and center lanes) or a control antibody (right lane) that cannot bind to either GR or PPARα. The immunoprecipitates were separated from the rest of the lysates by centrifugation and then analyzed by immunoblotting with either anti-GR (top box) or anti-PPARα (bottom box). As expected, the top box shows that the GR protein was detected in the immunoprecipitates generated using the α-GR when the same anti-GR antibody was used for the immunoblotting, but not in the immunoprecipitates generated with the control antibody

(no band observed). Strikingly, when one examines the immunoprecipitates by immunoblotting with the anti-PPARα antibody (bottom box), a substantial amount of PPARα is seen when the GR ligand is present (center lanes), whereas little co-precipitates in the absence of the GR ligand (left lane) or in the control immunoprecipitate (right lane). These results indicate that the GR ligand induces formation of a complex containing both the glucocorticoid receptor and the PPARα proteins. These results do not establish whether or not the GR and PPARα proteins bind directly to each other when the GR ligand is present or if there are additional molecules in the complex that act as intermediates holding the GR and PPARα tightly together when the ligand is present.

[Reprinted by permission from Nature Publishing Group, from H. Y. Lee et al., 2015, "PPAR-α and Glucocorticoid Receptor Synergize to Promote Erythroid Progenitor Self-Renewal," *Nature* **522**:474–477; permission conveyed through Copyright Clearance Center, Inc.]

Immunoprecipitation

Immunoprecipitation, often abbreviated as **IP**, exploits the specificity of antibodies to separate a protein of interest from other molecules in a complex mixture — for example, all proteins extracted from a sample of cells or a sample of blood. An antibody to the protein of interest is added to a sample, and the antibody is given time to bind to epitopes on the target protein. An agent that binds to the antibody is then added to cause the antibody and its bound target to precipitate out of solution into particles that can be isolated by centrifugation. Detailed examples of this technique are described in Figure 3-44 and Chapter 15. The precipitate is then solubilized under denaturing conditions — for example, in a detergent-containing buffer — to separate the antibody from the protein, and the immunoprecipitated target protein can then be analyzed. If the immunoprecipitated target is tightly bound to one or

more other molecules, those bound molecules may be precipitated along with the protein of interest (*co-immunoprecipitation*, sometimes abbreviated as *co-IP*). The co-IP method is used frequently to identify and characterize quaternary structures and supramolecular complexes.

Radioisotopes Are Indispensable Tools for Detecting Biological Molecules

A sensitive method for tracking a protein or other biological molecule is by detecting the radioactivity emitted from a radiolabel introduced into the molecule. In a radiolabeled molecule, at least one atom is present in a radioactive form, called a **radioisotope**.

Radioisotopes Useful in Biological Research

Hundreds of biological molecules (e.g., amino acids, nucleosides, and numerous other small-molecule metabolites) labeled with various radioisotopes are commercially available. These preparations vary considerably in their *specific activity*, which is the amount of radioactivity per unit of material, measured in disintegrations per minute (dpm) per millimole (mmol). The specific activity of a labeled compound depends on the radioisotope's *half-life*, the time required for half the atoms to undergo radioactive decay, which releases the detectable radiation. In general, the

shorter the half-life of a radioisotope, the higher its specific activity (Table 3-1). The specific activity of a labeled compound must be high enough for accurate detection of its emitted radiation.

TABLE 3-1 • Radioisotopes Commonly Used in Biological Research

Isotope	Half-Life
Phosphorus-32	14.3 days
Iodine-125	60.4 days
Sulfur-35	87.5 days
Tritium (hydrogen-3)	12.4 years
Carbon-14	5730.4 years

A common approach to radiolabeling macromolecules (proteins, RNA, DNA) in cells is to add a radiolabeled biosynthetic precursor to the extracellular medium [e.g., ^{3}H- or ^{35}S-labeled amino acids, ^{32}P-labeled phosphate (precursor for ^{32}P-labeled ATP), or ^{3}H-labeled nucleic acid precursors such as deoxythymidine (also simply called thymidine)]. The precursor enters the cells via transporters (see Chapter 11) and is incorporated into newly synthesized macromolecules by the cells (see Chapter 5). For example, methionine and cysteine labeled with sulfur-35 (^{35}S) are widely used to label cellular proteins because preparations of these amino acids with high specific activities ($>10^{15}$ dpm/mmol) are available. Kinases within cells (or used in vitro) can transfer a ^{32}P-labeled

phosphate from ^{32}P-labeled ATP to label phosphoproteins. In intact cells, ^{32}P-labeled phosphate added to the medium is taken up via phosphate transporters and then incorporated into various phosphate containing molecules, including ATP. Likewise, commercial preparations of ^{3}H-labeled nucleic acid precursors have much higher specific activities than those of the corresponding ^{14}C-labeled preparations. In most experiments, the former are preferable because they allow RNA or DNA to be adequately labeled a shorter time after incorporation or require a smaller cell sample. Various phosphate-containing compounds in which the phosphorus atom is the radioisotope phosphorus-32 are readily available. Because of their high specific activity, ^{32}P-labeled nucleotides are routinely used to label nucleic acids in cell-free systems.

Labeled compounds in which a radioisotope replaces atoms normally present in the molecule have virtually the same chemical properties as the corresponding unlabeled compounds. Enzymes, for instance, generally cannot distinguish between substrates labeled in this way and their unlabeled substrates. The presence of such radioactive atoms is indicated with the isotope in brackets (no hyphen) as a prefix (e.g., [^{3}H]leucine). In contrast, labeling of almost any biomolecule (e.g., protein or nucleic acid) with the radioisotope iodine-125 (^{125}I) requires the covalent addition of ^{125}I to a molecule that normally does not have iodine as part of its structure. Because this labeling procedure modifies the chemical structure, the biological activity of the labeled molecule may differ somewhat

from that of the unlabeled form. The presence of such radioactive atoms is indicated with the isotope as a prefix followed a hyphen (no bracket) (e.g., ^{125}I-trypsin). Standard methods for labeling proteins with ^{125}I result in covalent attachment of the ^{125}I primarily to the aromatic rings of tyrosine side chains (mono- and diiodotyrosine). Nonradioactive isotopes are finding increasing use in cell biology, especially in nuclear magnetic resonance studies and in mass spectroscopy applications, as will be explained below.

Labeling Experiments and Detection of Radiolabeled Molecules

Whether labeled compounds are detected by **autoradiography** — exposure of the sample on a two-dimensional detector (photographic emulsion or electronic detector) — or their radioactivity is measured in an appropriate *counter*, the amount of a radiolabeled compound in a sample can be determined with great precision.

In one use of autoradiography, a tissue, cell, or cell constituent is labeled with a radioactive molecule, unassociated radioactive material is washed away, and the structure of the sample is stabilized either by chemically cross-linking the macromolecules in the sample ("fixation") or by freezing it. The sample is then overlaid with a photographic emulsion that is sensitive to radiation. Development of the emulsion yields small silver grains whose

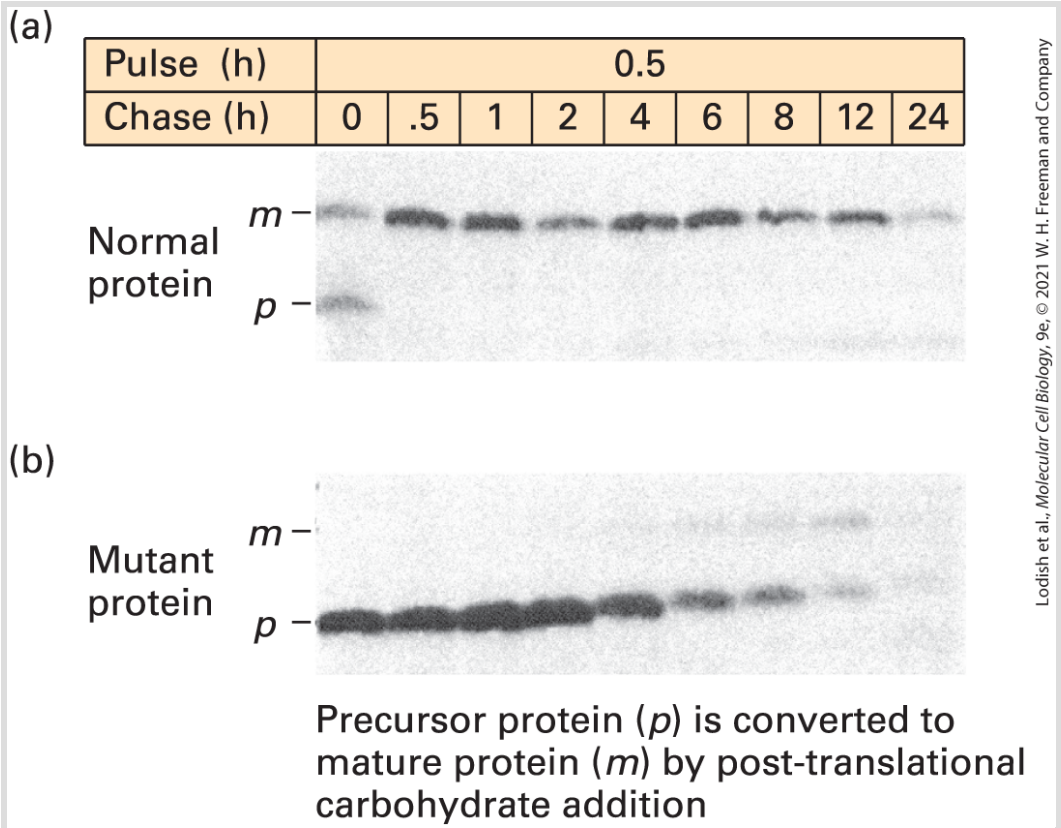
distribution corresponds to that of the radioactive material and is usually detected by microscopy. Autoradiographic studies of whole cells were crucial in determining the intracellular sites where various macromolecules are synthesized and the subsequent movements of those macromolecules within cells. Various techniques employing fluorescence microscopy, which we describe in Chapter 4, have largely supplanted autoradiography for studies of this type. However, autoradiography is sometimes used in various assays for detecting specific, isolated DNA or RNA sequences at specific tissue locations (see Chapter 6) in a technique referred to as in situ hybridization.

Quantitative measurements of the amount of radioactivity in a labeled material are performed with several different instruments. A Geiger counter measures ions produced in a gas by the β particles or γ rays emitted from a radioisotope. These instruments are mostly handheld devices used to monitor radioactivity in the laboratory to protect investigators from excess exposure. In a scintillation counter, a radiolabeled sample is mixed with a liquid containing a fluorescent compound that emits a flash of light when it absorbs the energy of the β particles or γ rays released in the decay of the radioisotope; a phototube in the instrument detects and counts these light flashes. Phosphorimagers detect radioactivity using a two-dimensional array detector, storing digital data on the number of disintegrations per minute per small pixel of surface area. These instruments, which can be thought of as a kind of reusable electronic film, are commonly used to quantify radioactive

molecules separated by gel electrophoresis and are replacing photographic emulsions for this purpose.

Combinations of labeling and biochemical techniques and of visual and quantitative detection methods are often employed in labeling experiments. For instance, to identify the major proteins synthesized by a particular cell type, a sample of the cells is incubated with a radiolabeled amino acid (e.g., [^{35}S] methionine) for a few minutes, during which time the labeled amino acid enters the cells and mixes with the cellular pool of unlabeled amino acids, and some of it is biosynthetically incorporated into all newly synthesized proteins. Subsequently, unincorporated radiolabeled amino acid is washed away from the cells. The cells are harvested, and the mixture of cellular proteins is extracted from the cells (e.g., by a detergent solution) and then separated by any of the methods commonly used to resolve complex protein mixtures into individual components. Gel electrophoresis in combination with autoradiography or phosphorimager analysis is often the method of choice. The radioactive bands in the gel correspond to newly synthesized proteins, which have incorporated the radiolabeled amino acid. To detect a specific protein of interest, rather than the entire ensemble of biosynthetically radiolabeled proteins, a specific protein can be isolated by immunoprecipitation. The precipitate is then solubilized, for example, in an SDS-containing buffer, and the sample is analyzed by SDS-PAGE followed by autoradiography to detect the protein that is radioactively labeled. In this type of

experiment, a fluorescent compound that is activated by the radiation ("scintillator") may be infused into the gel on completion of the electrophoretic separation so that the light emitted can be used to detect the presence of the labeled protein, using either film or a two-dimensional electronic detector. An example is shown in the experiment described in Figure 3-45. This method is particularly useful for weak β emitters such as ^{3}H.



EXPERIMENTAL FIGURE 3-45 Pulse-chase experiments can track the pathway of protein modification within cells. (a) To follow the fate of a specific newly synthesized protein in cells, cells were incubated with [^{35}S]methionine for 0.5 hours (the pulse) to label all newly synthesized proteins, and any radioactive amino acid not incorporated into the cells was then washed away. The cells were further incubated (the chase) for varying times up to 24 hours, and samples from each time point of the chase were subjected to

immunoprecipitation to isolate one specific protein (here the low-density lipoprotein receptor). SDS-PAGE of the immunoprecipitates followed by autoradiography permitted visualization of the target protein, which is initially synthesized as a small precursor (*p*) and then rapidly modified to a larger mature form (*m*) by addition of carbohydrates. About half of the labeled protein was converted from *p* to *m* during the pulse; the rest was converted after 0.5 hours of chase. The protein remained stable for 6–8 hours before it began to be degraded within cells (as indicated by reduced band intensity). (b) The same experiment was performed in cells in which a mutant form of the protein is made. The mutant *p* form cannot be properly converted to the *m* form, and it is more quickly degraded than the normal protein.

[©1986 Kozarsky et al. *J .Cell Biol.* **102**:1567–1575. doi:10.1083/jcb.102.5.1567]

Pulse-chase experiments are particularly useful for tracing changes in the intracellular location of proteins or the modification of a protein or metabolite over time. In this experimental protocol, a cell sample is exposed to a radiolabeled compound that can be incorporated into or otherwise attached to a cellular molecule of interest — the *pulse* — for a brief period. The pulse ends when the unincorporated radiolabeled molecules are washed away and the cells are exposed to a vast excess of the identical, but unlabeled, compound to dilute the radioactivity of any remaining, but unincorporated, radiolabeled compound. This procedure prevents any incorporation of significant amounts of radiolabel after the pulse period and initiates the *chase* period (see Figure 3-45). Samples taken periodically during the chase period are assayed to determine the location or chemical form of the radiolabel as a function of time. Pulse-chase experiments in which the radiolabeled protein is detected by autoradiography after immunoprecipitation and SDS-

PAGE are often used to follow the rate of synthesis, modification, and degradation of proteins. In these experiments, radiolabeled amino acid precursors are added during the pulse, and the amounts and characteristics of the radiolabeled target protein are detected during the chase. One can thus observe postsynthetic modifications of the protein, such as the covalent addition of sugars (see Chapters 13 and 14) or proteolytic cleavage, that change its electrophoretic mobility, as well as the rate of degradation of the protein, which is detected as the loss of signal with increasing time of chase. A classic use of the pulse-chase technique with autoradiography was in studies that elucidated the pathway traversed by secreted proteins from their site of synthesis in the endoplasmic reticulum to the cell surface (see Chapter 14).

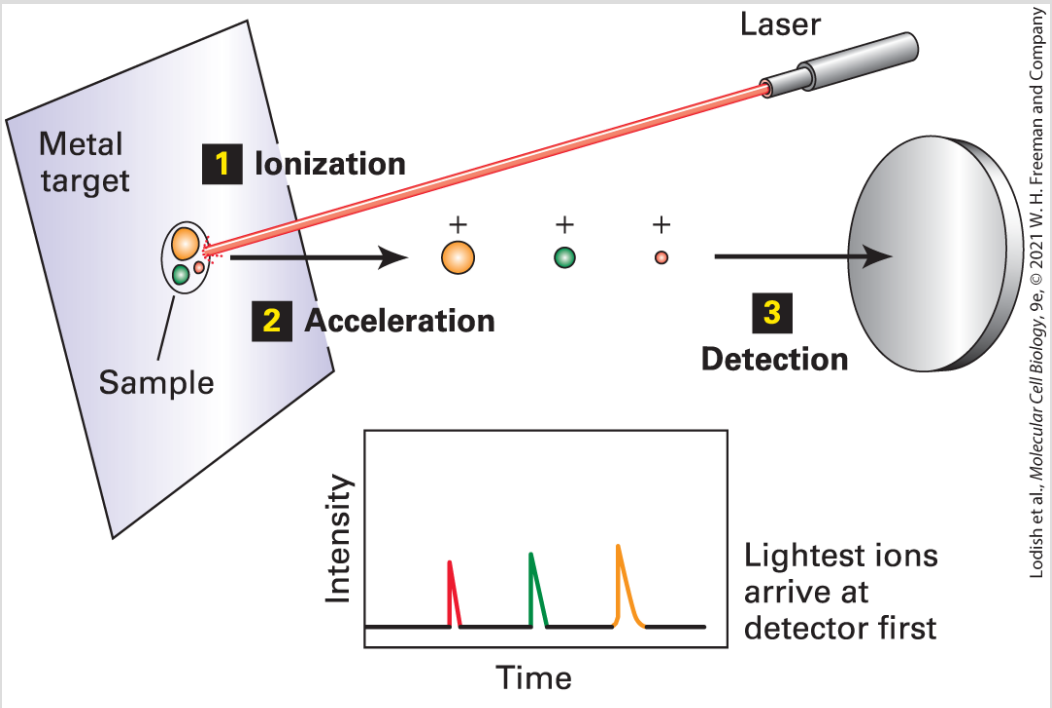
Mass Spectrometry Can Determine the Mass and Sequence of Proteins

Mass spectrometry (MS) is a powerful technique for characterizing proteins, especially for determining the mass of a protein or fragments of a protein. With such information in hand, it is also possible to determine part or all of the protein's sequence. This method permits the accurate direct determination of the ratio of the mass (m) of a charged molecule (a molecular ion) to its charge (z), or m/z. Additional techniques are then used to deduce the absolute mass of the molecular ion.

All mass spectrometers have four key features. The first is an ion source, from which charge, usually in the form of protons, is transferred to the peptide or protein molecules under study, causing them to become ionized. The conversion of polypeptide molecules to ions occurs in the presence of a high electric field, which directs the charged molecular ions into the second key component, the mass analyzer. The mass analyzer, which is always in a high vacuum chamber, physically separates the ions on the basis of their differing mass-to-charge (*m/z*) ratios. The separated ions are subsequently directed to strike a detector, the third key component, which provides a measure of the *m/z* ratio of each ion in the sample and the relative abundances of those ions. The fourth essential component is a computerized data system that is used to calibrate the instrument; to acquire, store, and process the resulting data; and often to direct the instrument to automatically collect additional specific types of data from the sample, based on the initial observations. This type of automated feedback is used for the tandem MS (MS/MS) peptide-sequencing methods described below.

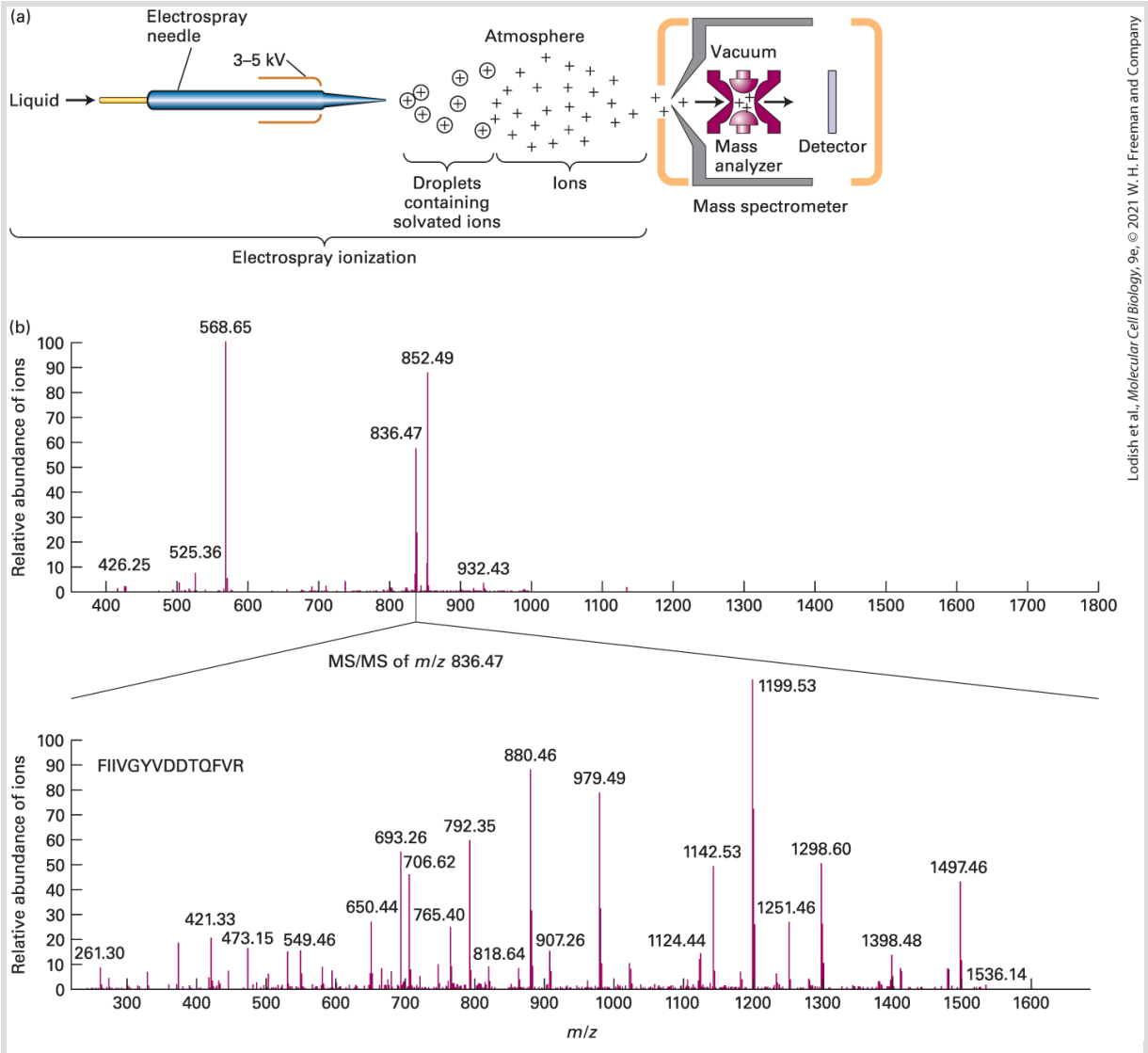
The two most frequently used methods of generating ions of proteins and protein fragments are (1) matrix-assisted laser desorption/ionization (MALDI) and (2) electrospray (ES). In MALDI (Figure 3-46), the peptide or protein sample is mixed with the matrix — a low-molecular-weight organic acid that absorbs UV light — and the mixture is dried onto a metal target. Energy from a UV laser is absorbed by the matrix and converted into heat. As a consequence,

the sample vaporizes and the protein sample is ionized by protonation, producing singly charged molecular ions from the constituent molecules. In ES (Figure 3-47a), a sample of peptides or proteins in solution is converted into a fine mist of tiny droplets by spraying through a narrow capillary at atmospheric pressure. The droplets are formed in the presence of a high electric field, which renders them highly charged. The solvent evaporates from the droplets in their short flight (mm) to the entrance of the mass spectrometer's mass analyzer, forming multiply charged ions from the peptides and proteins. The gaseous ions enter the mass analyzer, where they are accelerated by electric fields and separated on the basis of their m/z.



EXPERIMENTAL FIGURE 3-46 Molecular mass can be determined by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. In a MALDI-TOF mass spectrometer, pulses of light from a laser

ionize a protein or peptide mixture that is absorbed on a metal target (step 1). An electric field in the mass analyzer accelerates the ions in the sample toward the detector (steps 2 and 3). The time it takes an ion to reach the detector is proportional to the square root of the mass-to-charge (m/z) ratio. Among ions having the same charge, the smaller ions move faster (shorter time to the detector). The molecular weight of each ion from the sample is calculated using the time of flight of a standard.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 3-47 Molecular mass of proteins and peptides can be determined by electrospray ionization ion-trap mass spectrometry. (a) Electrospray (ES) ionization converts proteins and peptides in a solution into highly charged gaseous ions by passing the

solution through a needle (forming the droplets) that has a high voltage across it (charging the droplets). Evaporation of the solvent produces gaseous ions that enter a mass spectrometer. The ions are analyzed by an ion-trap mass analyzer that separates the ions by their *m/z* ratios and then directs the ions to the detector. (b) *Top:* Mass spectrum of a mixture of three major and several minor peptides from the mouse H-2 class I histocompatibility antigen Q10 α chain. The data are presented as the relative abundance of the ions striking the detector (*y* axis) as a function of the mass-to-charge (*m*/*z*) ratio (*x* axis). *Bottom:* In an MS/MS instrument such as the ion trap shown in part (a), a specific peptide ion can be selected for fragmentation into smaller ions that are then analyzed and detected. The MS/MS spectrum (also called the *product-ion spectrum*) provides detailed structural information about the parent ion, including sequence information for peptides. Here the ion with an *m*/*z* of 836.47 was selected and fragmented and the *m*/*z* mass spectrum of the product ions measured. Note there is no longer an ion with an *m*/*z* of 836.47 present because it was fragmented. From the varying sizes of the product ions, the understanding that peptide bonds are often broken in such experiments, the known *m*/*z* values for individual amino acid fragments, and database information, the sequence of the peptide, FIIVGYVDDTQFVR, can be deduced.

[Part (b) Unpublished data from S. Carr.]

The two most frequently used types of mass analyzers are time-of-flight (TOF) instruments and ion traps. TOF instruments exploit the fact that the time it takes an ion to pass through the length of the mass analyzer before reaching the detector is proportional to the square root of *m/z* (smaller ions move faster than larger ones with the same charge; see Figure 3-46). In ion-trap analyzers, tunable electric fields are used to capture (or "trap") ions with a specific *m/z* and to sequentially pass the trapped ions out of the mass analyzer onto the detector in order of their *m/z* ratios (see Figure 3-47a). By varying the electric fields, researchers can examine ions with a wide range of *m/z* values one by one, producing a mass spectrum, which

is a graph of *m/z* (*x* axis) versus relative abundance, determined by the intensity of the signal measured by the detector (*y* axis) (Figure 3-47b, *top*).

In tandem, or MS/MS, instruments, any given parent ion in the original mass spectrum (see Figure 3-47b, *top*) can be chosen (mass-selected) for further analysis. The chosen ions are transferred into a second chamber in which they are broken into smaller fragment ions by collision with an inert gas, and then the *m/z* and relative abundances of the resulting fragment ions are measured in a second MS analyzer (Figure 3-47b, *bottom*; see also Figure 3-50). These multiple mass-analysis and fragmentation steps all take place within the same machine in about 0.1 seconds per selected parent ion. The fragmentation and subsequent mass analysis permit the sequences of short peptides (<25 amino acids) to be determined because collisional fragmentation occurs primarily at peptide bonds, so the differences in masses between the multiple ion fragments generated correspond to the in-chain masses of the individual amino acids, permitting deduction of the sequence in conjunction with database sequence information (see Figure 3-47b, *bottom*).

Mass spectrometry is highly sensitive, able to detect as little as $0.001–0.010 \times 10^{-18}$ mol (0.001–0.010 attomoles) of a peptide or 1×10^{-18} mol (1 attomole) of a protein of 200,000 MW. The accuracy of the measurements of mass are dependent on the specific mass analyzer used, but error rates are typically about +/− 5 ppm for

peptides and 0.05–0.1 percent for proteins. So sensitive are its measurements of mass that MS can readily distinguish between two peptides that are chemically identical yet differ only in that one of the peptides contains only atoms with the most common, *light* isotopes (e.g., ^{1}H, ^{12}C, ^{14}N) whereas the other contains one or more atoms with the corresponding *heavy*, stable isotopes (e.g., the nonradioactive isotopes ^{2}H, ^{13}C, ^{15}N).

As described in Section 3.6, one of the most frequent uses of MS is to analyze complex mixtures of proteins such as those from cells, tissues, or biofluids such as plasma. Most commonly, protein samples are digested by proteases, and the peptide digestion products are subjected to analysis by MS. An especially powerful application of MS is to take a complex mixture of proteins from a biological specimen and digest it with the enzyme trypsin, which hydrolyzes peptide bonds at lysine and arginine residues, or other proteases. Peptides in the resulting complex mixture are then separated using liquid chromatography. The fluid flowing out of the liquid chromatography column is transferred into an ES tandem mass spectrometer. This technique, which is called *LC-MS/MS* and is described in more detail below, permits the nearly continuous analysis of a very complex mixture of proteins.

The abundances of ions determined by mass spectrometry in any given sample are relative, not absolute, values. Therefore, if one wants to use MS to compare the absolute amounts of a particular

protein in two different samples (e.g., from a normal versus a mutant organism), it is necessary to have an internal standard in the samples — a molecule whose amounts do not differ between the two samples. One then determines the amounts of the protein (or peptide fragments of the protein) of interest relative to the amount of the standard in each sample. This approach permits quantitatively accurate comparisons of protein levels between samples. An alternative approach involves simultaneously comparing in a single MS analysis the amounts of proteins from two or more different cell or tissue samples that have been mixed together. The different samples might differ genetically (wild-type versus mutant cells, clinical samples from different patients) or might differ due to the application of different environmental conditions (e.g., drug or hormone treatments). This mixing approach is possible provided the *m/z* ratios of proteins or peptides in each sample can be distinguished by MS from the corresponding and chemically identical proteins or peptides in the other samples. To make it possible to distinguish identical proteins from different samples, chemical, enzymatic, or biosynthetic techniques are used to label the different samples with different stable isotopes (e.g., $^{1}H/^{2}H$, $^{12}C/^{13}C$, $^{14}N/^{15}N$). For example, cells or organisms can first be grown in the presence of amino acids containing either heavy or light isotope atoms so that these amino acids are biosynthetically incorporated into all the proteins of that sample. Cells are typically incubated with the heavy or light amino acids for five or more cell divisions to ensure that all proteins are thoroughly labeled. Proteins

from the two samples are then mixed together and digested into peptides, and the peptides analyzed simultaneously by mass spectrometry. Proteins and peptides derived from the heavy sample can be distinguished in the mass spectrometer from those from the other, light, sample because of their higher masses, but are otherwise chemically identical. This method is called *s*table *i*sotope *l*abeling with *a*mino *a*cids in *c*ell *c*ulture (SILAC).

Chemical methods with names such as iTRAQ and TMT permit investigators to attach chemically identical but isotopically distinct labels to the peptides in each of 4–16 distinct biological samples. After labeling, the samples are mixed together and then analyzed simultaneously by MS. The isotopically distinct tags permit one to determine the sample from which any peptide arose, even though all of the tagged samples were mixed together prior to MS analysis. As a consequence, the relative amounts of all detectable proteins in each of the different samples can be determined quickly, efficiently, and accurately. In this way, the effects of a genetic, environmental, or chemical (e.g., drug) perturbation of cells on the cellular proteins can be assessed. Later in this chapter, we will consider the importance of analyzing such perturbations on the abundances of many different proteins simultaneously.

Protein Primary Structure Can Be Determined by Chemical Methods and

from Gene Sequences

The classic method for determining the amino acid sequence of a protein is Edman degradation. In this procedure, the free amino group of the N-terminal amino acid of a polypeptide is labeled, and the labeled amino acid is then cleaved from the polypeptide and identified by high-pressure liquid chromatography. The polypeptide is left one residue shorter, with a new amino acid at the N-terminus. The cycle is repeated on the ever-shortening polypeptide until all the residues have been identified.

Before about 1985, biologists commonly used Edman degradation for determining protein sequences. Now, however, complete protein sequences usually are determined by analysis of genome and messenger RNA sequences. The complete genomes of many organisms have already been sequenced, and the database of genome sequences from humans and numerous model organisms is expanding rapidly. As discussed in Chapter 6, the sequences of proteins can be deduced from DNA sequences that are predicted to encode proteins.

A powerful approach for determining the primary structure of an isolated protein combines MS and the use of sequence databases. First, the peptide mass fingerprint of the protein is obtained by MS. A *peptide mass fingerprint* is the list of the molecular weights of peptides that are generated from the protein by digestion with a specific protease, such as trypsin. The molecular weights of the

parent protein and its proteolytic fragments are then used to search genome databases for any similar-sized protein with identical or similar peptide mass fingerprints. Mass spectrometry can also be used to directly sequence peptides using MS/MS, as described above.

Protein Conformation Is Determined by Sophisticated Physical Methods

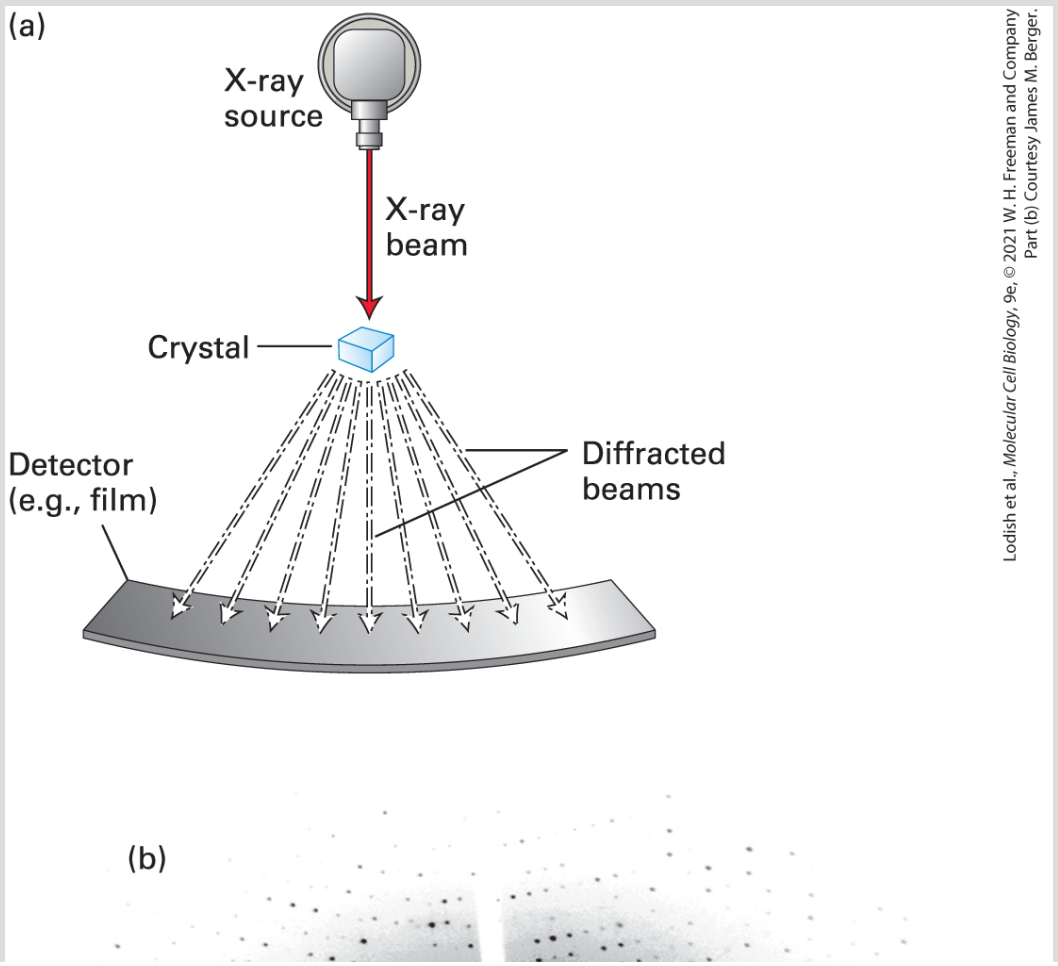
In this chapter, we have emphasized that protein function is dependent on protein structure. Thus, to figure out exactly how a protein works, its three-dimensional structure must be determined. Determining a protein's conformation requires sophisticated physical methods and complex analyses of the experimental data. Here we briefly describe three methods used to generate three-dimensional models of proteins.

X-ray Crystallography

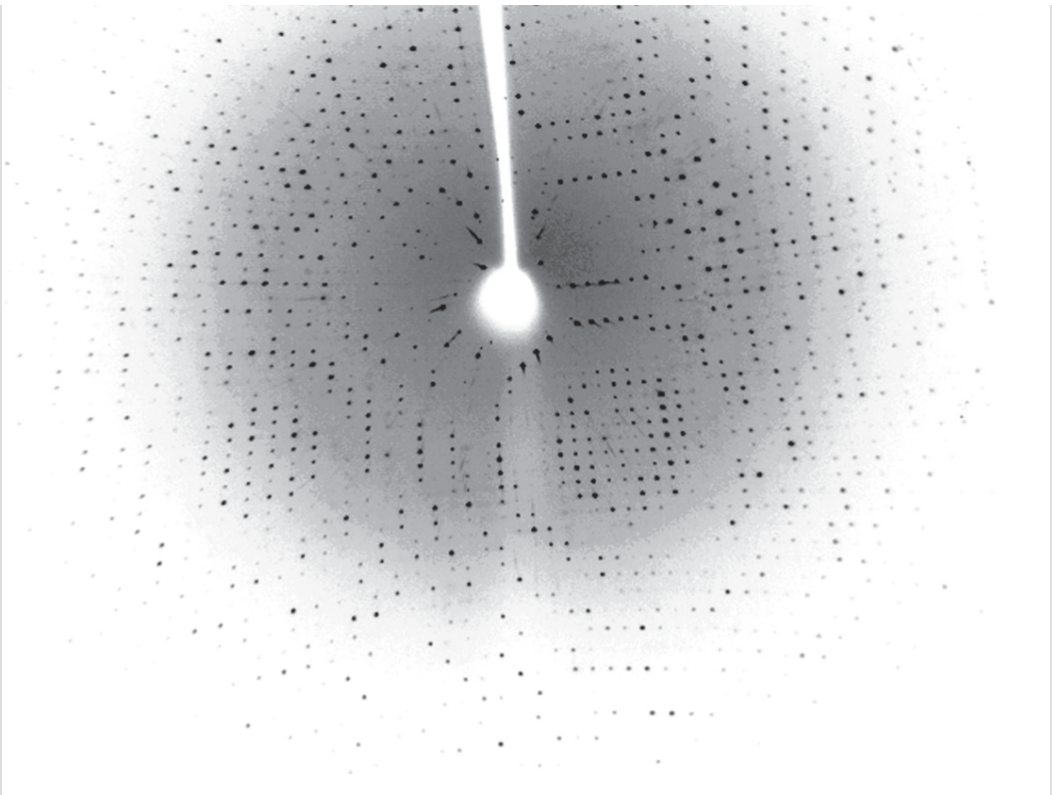
The use of **x-ray crystallography** to determine the three-dimensional structures of proteins was pioneered by Max Perutz and John Kendrew in the 1950s. In this technique, beams of x-rays are passed through a protein crystal, in which millions of protein molecules are precisely aligned with one another in a rigid crystalline array. The wavelengths of x-rays are about 0.1–0.2 nm, short enough to determine the positions of individual atoms in the

protein. The electrons in the atoms of the crystal scatter the x-rays, which produce a diffraction pattern of discrete spots when they are intercepted by photographic film or an electronic detector (Figure 3-48). Such patterns are extremely complex — composed of as many as 25,000 diffraction spots, or reflections, whose measured intensities vary depending on the distribution of the electrons in the sample, which is, in turn, determined by the atomic structure and three-dimensional conformation of the protein. Elaborate calculations and modifications of the protein (e.g., the binding of heavy metals) must be made to interpret the diffraction pattern and calculate the distribution of electrons (called the *electron density map*). A portion of an electron density map of a protein can be seen in Figure 2-9. With the three-dimensional electron density map in hand, one then fits a molecular model of the protein to match the electron density, and it is these models that one sees in many of the various diagrams of proteins throughout this book (e.g., see Figure 3-9). The process is analogous to reconstructing the precise shape of a rock from the ripples that it creates when thrown into a pond. The precision with which one can determine the structure of a molecule from x-ray crystallography depends on the quality of the crystals used for the analysis and on how flexible part or all of the molecule is within the crystal lattice. A commonly used measure of the precision of a structure is its *resolution* given in Angstrom units (Å). Very high resolution protein structures have a resolution of <1 Å (Figure 2-9), and the positions of individual atoms can be observed. In structures whose resolutions are between 1 and 3 Å, the positions of many side

chains are well determined but individual atom positions must be inferred. Although sometimes the structures of parts of the protein cannot be clearly defined, using x-ray crystallography, researchers are systematically determining the structures of representative types of most proteins. To date, more than 140,000 detailed three-dimensional structures, including more than 35,000 distinct protein sequences, have been established using x-ray crystallography. These structures can be found in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (http://www.rcsb.org/pdb/home/home.do), each with its own "PDB" entry.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Part (b) Courtesy James M. Berger.

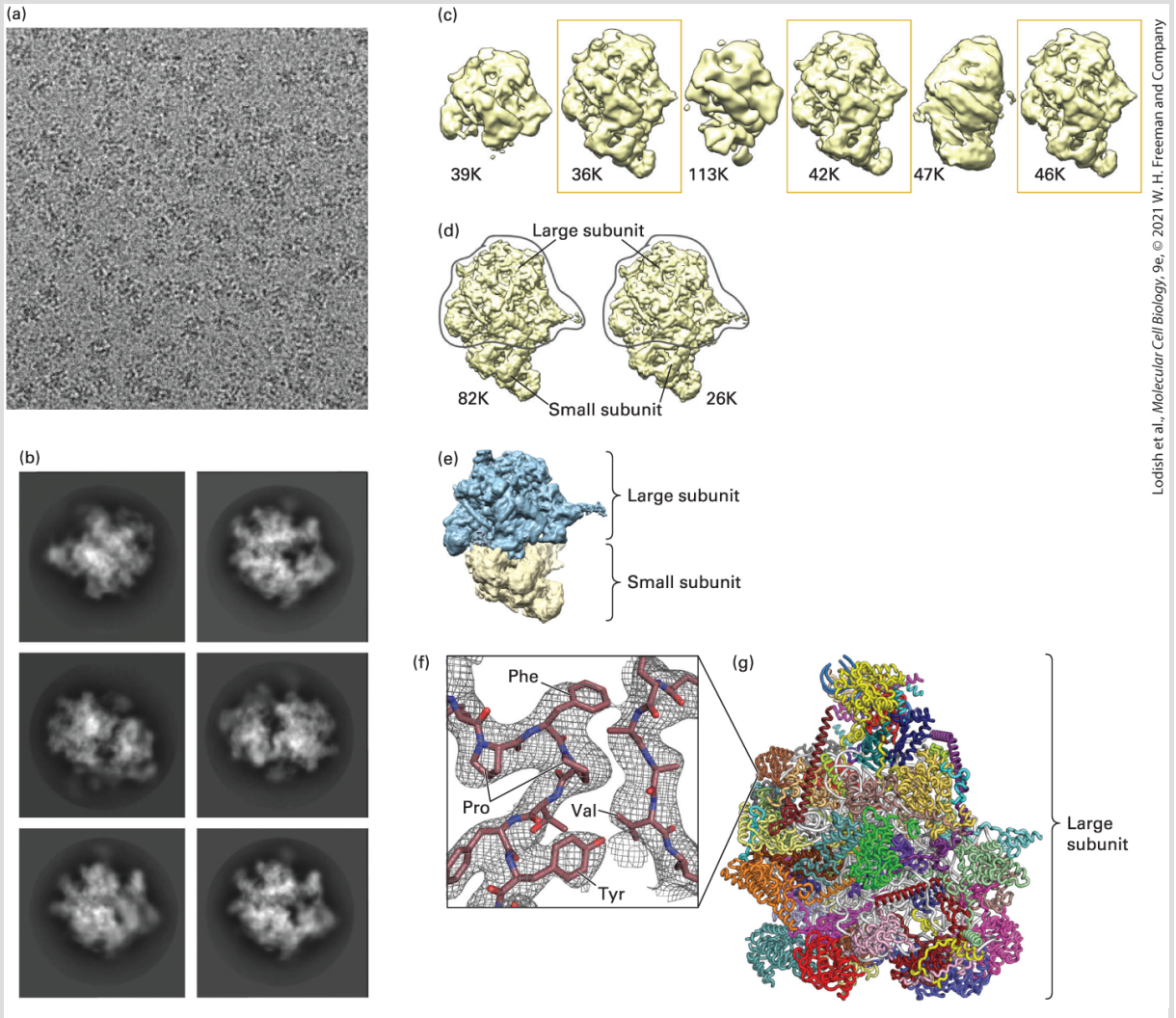


EXPERIMENTAL FIGURE 3-48 X-ray crystallography provides diffraction data from which the three-dimensional structure of a protein can be determined. (a) Basic components of an x-ray crystallographic determination. When a narrow beam of x-rays strikes a crystal, part of the beam passes straight through and the rest is scattered (diffracted) in various directions. The intensity of the diffracted waves, which form periodic arrangements of diffraction spots, is recorded on an x-ray film or with a solid-state electronic detector. (b) X-ray diffraction pattern for a protein crystal collected on a solid-state detector. From complex analyses of patterns of spots like this one, the locations of the atoms in a protein can be determined.

Cryoelectron Microscopy

Although some proteins readily crystallize, obtaining crystals of others — particularly large multisubunit proteins and membrane-associated proteins — requires a time-consuming, often robot-

assisted trial-and-error effort to find just the right conditions, if they can be found at all. (Growing crystals suitable for structural studies is as much an art as a science.) There are several alternative ways to determine the structures of such difficult-to-crystallize proteins. One is cryoelectron microscopy (Figure 3-49). In this technique, a dilute protein sample in an aqueous solution is applied in a thin layer to an electron microscope sample holder (a "grid") and rapidly frozen in liquid helium to preserve its structure. It is then examined in the frozen, hydrated state in a cryoelectron microscope. Images of the protein are recorded on a very sensitive camera using a low dose of electrons to prevent radiation-induced damage to the structure. Since the individual proteins are in different orientations in the frozen sample, sophisticated computer algorithms analyze the images to sort them into groups with the same orientation. The average image of each orientation is calculated from images of the thousands of different molecules in each group, and then the computer assembles the average images, each of which show views of the protein from different orientations, to reconstruct the protein's structure in three dimensions. Recent advances in this technology have produced structures in which the polypeptide backbone and amino acid side chains can be discerned with resolutions reported to be as high as $\sim$2 Å and for many structures greater than 4 Å. These structures help provide insight into the mechanisms underlying the protein's function. The use of cryoelectron microscopy and other types of electron microscopy for visualizing cell structures is discussed in Chapter 4.



EXPERIMENTAL FIGURE 3-49 Cryoelectron microscopy analysis of the structure of the human mitochondrial ribosome. The mitochondrion is a complex, multifunctional intracellular organelle best known for its ability to synthesize the energy carrier ATP (see Chapter 12). Human mitochondria can synthesize proteins encoded by mitochondrial DNA using large (1.7 MDa), multiprotein (at least 78) and multi-RNA complexes called *mitochondrial ribosomes* that differ somewhat from cytoplasmic ribosomes. (a) Cryoelectron micrograph of isolated human mitochondrial ribosomes. The low contrast between the ribosomes and the buffer solution makes it difficult to clearly see individual, frozen ribosome particles, which are oriented randomly in the image. (b) Automated image processing of 323,292 individual particles permits their grouping into classes based on orientation and averaging of the images within each class to generate clearer images of the ribosome. (c) Additional computational analysis generates distinct structures, each based on tens of thousands of individual particles (the number of particles analyzed for each

structure in thousands [K] is shown beneath each). The structures enclosed in boxes were selected for additional analysis, which produced the two very similar models shown in (d) containing virtually identical large subunits. (e) Color-coded, low-resolution model of the electron density of the large (blue) and small (yellow) subunits. The conformational heterogeneity of the small subunit prevented the determination of its high-resolution structure from the data shown here. (f) High-magnification view of the experimentally determined electron density (meshwork) from a very small portion of one protein within the large subunit. This view illustrates how the electron density is used to build the superimposed molecular model of polypeptide chains. The side chains of proline (Pro), phenylalanine (Phe), valine (Val), and tyrosine (Tyr) residues are easily seen and demonstrate the power of cryoelectron microscopy to determine protein structures at very high resolutions. (g) Model of the 48 protein subunits (different colors) in the large subunit determined at 3.4 Å resolution.

[Republished with permission of American Association for the Advancement of Science, from A. Brown et al., 2014, "Structure of the Large Ribosomal Subunit from Human Mitochondria." *Science* **346**(6210): 718–722; permission conveyed through Copyright Clearance Center, Inc.]

NMR Spectroscopy

The three-dimensional structures of small proteins containing as many as 200 amino acids can be studied routinely with nuclear magnetic resonance (NMR) spectroscopy, and specialized approaches can be used to extend the size range to somewhat larger proteins. In this technique, a concentrated protein solution is placed in a magnetic field, and the effects of different radio frequencies on the nuclear spin states of different atoms are measured. The spin state of any atom is influenced by neighboring atoms in adjacent residues, with closely spaced residues having a greater effect than distant residues. From the magnitude of the effect, the distances

between residues can be calculated by a triangulation-like process; these distances are then used to generate a model of the three-dimensional structure of the protein. An important distinction between x-ray crystallography and NMR spectroscopy is that the former method directly determines the locations of the atoms, while the latter directly determines the distances between the atoms, from which the structure is deduced.

Although NMR does not require the crystallization of a protein — a definite advantage — this technique is usually limited to proteins smaller than about 80 kDa (although new techniques permit analysis of the dynamics in much larger proteins). The precision of an NMR structure can be as high as ~2 Å. NMR analysis can provide information about the ability of a protein to adopt a set of closely related, but not exactly identical, conformations and to move between those conformations (protein dynamics). This is a common feature of proteins, which are not absolutely rigid structures, but can "breathe" or exhibit slight variations in the relative positions of their constituent atoms. In some cases, these variations can have functional significance; for example, they may influence how proteins bind to one another. NMR structural analysis has been particularly useful in studying isolated protein domains, which can often be obtained as stable structures and tend to be small enough for this technique. To date, there are more than 13,000 NMR-determined structures available in the Protein Data Bank.

Another powerful approach to studying protein dynamics and protein-protein interactions is hydrogen/deuterium exchange mass spectrometry (HXMS). When a protein is placed in a deuterated water (D_2O) solution, the rate at which deuterium is exchanged for hydrogen in the amides in the peptide bonds depends on the accessibility of an amide to the solvent. Those amides exposed on the protein's surface are highly accessible and exhibit rapid proton/deuterium exchange. Those amides buried in the center of the protein or in a protein-to-protein interface, as well as those participating in hydrogen bonds with other parts of the protein, exhibit slower proton/deuterium exchange rates. A change in protein conformation or binding to other molecules has the potential to alter the rate of hydrogen/deuterium exchange of one or more amides of a protein. MS analysis permits a hypersensitive assay of such conformational changes, allowing the identification of those parts of the protein that directly bind to other molecules or undergo such conformational changes.

KEY CONCEPTS OF SECTION 3.5

Purifying, Detecting, and Characterizing Proteins

- Proteins can be separated from other cell components and from one another on the basis of differences in their physical and chemical properties.
- Centrifugation separates proteins on the basis of their rates of sedimentation, which are influenced by their masses and shapes (see Figure 3-40).
- Electrophoresis separates proteins on the basis of their rates of movement in an applied electric field. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) can resolve polypeptide chains differing in molecular weight by 10 percent or less (see Figure 3-

41). Two-dimensional gel electrophoresis provides additional resolution by separating proteins first by charge (first dimension) and then by mass (second dimension).

- Liquid chromatography separates proteins on the basis of their rates of movement through a column packed with spherical beads. Proteins differing in mass are resolved on gel filtration columns: those differing in charge using ion-exchange columns and those differing in ligand-binding properties using affinity columns (see Figure 3-43).
- Various assays are used to detect and quantify proteins. Some assays use a light-producing reaction to generate a readily detected signal. Other assays produce an amplified colored signal with enzymes and chromogenic substrates.
- Antibodies are powerful reagents used to detect, quantify, and isolate proteins.
- Immunoblotting, also called *Western blotting*, is a frequently used method to study specific proteins that exploits the high specificity and sensitivity of protein detection by antibodies and the high-resolution separation of proteins by SDS-PAGE (see Figure 3-44).
- Immunoprecipitation, often abbreviated as *IP*, permits the separation of a protein of interest from other proteins in a complex mixture using antibodies specific for the protein of interest. The antibodies are used to precipitate their target protein out of solution for subsequent analysis. Molecules tightly bound to the target protein can precipitate with it (co-immunoprecipitation).
- Isotopes, both radioactive and nonradioactive, play a key role in the study of proteins and other biomolecules. They can be incorporated into molecules without changing the chemical composition of the molecule or as add-on tags. They can be used to help detect the synthesis, location, processing, and stability of proteins.
- Autoradiography is a technique for detecting radioactively labeled molecules in cells, tissues, or electrophoretic gels using two-dimensional detectors (photographic emulsion or electronic detectors).
- Pulse-chase experiments can determine the intracellular fate of proteins and other metabolites (see Figure 3-45).
- Mass spectrometry is a very sensitive and highly precise method of detecting, identifying, and characterizing proteins and peptides.
- Three-dimensional structures of proteins are obtained by x-ray crystallography, cryoelectron microscopy, and NMR spectroscopy. X-ray crystallography provides the most detailed structures but requires protein crystallization. Cryoelectron microscopy is most useful for large protein complexes, which are difficult to crystallize. Only relatively small proteins are amenable to NMR three-dimensional structural analysis.

3.6 Proteomics

For most of the twentieth century, the study of proteins was restricted primarily to the analysis of individual proteins. For example, one would study an enzyme by determining its enzymatic activity (its substrates, products, rate of reaction, requirement for cofactors, pH dependence, etc.), its structure, and its mechanism of action. In some cases, the relationships between a few enzymes that participate in a metabolic pathway might also be studied. On a broader scale, the localization and activity of an enzyme would be examined in the context of a cell or tissue. The effects of mutations, diseases, or drugs on the expression and activity of the enzyme might also be the subject of investigation. This multipronged approach provided deep insight into the function and mechanisms of action of individual proteins or small numbers of interacting proteins. However, studying proteins one at a time does not provide a global picture of what is happening for the collection of all proteins — the **proteome** — of an organelle, cell, tissue, or entire organism.

Proteomics Is the Study of All or a Large Subset of Proteins in a Biological System

The advent of genomics (sequencing of genomic DNA and its associated technologies, e.g., simultaneous analysis of the levels of all mRNAs in cells and tissues) clearly showed that a global, or systems, approach to biology could provide unique and highly valuable insights. Many scientists recognized that a global analysis of the proteins in biological systems had the potential for equally valuable contributions to our understanding. Thus a new field was born — **proteomics**. Proteomics is the systematic study of the amounts, modifications, interactions, localization, and functions of all or subsets of proteins at the whole-organism, tissue, cellular, and subcellular levels.

A number of broad questions are addressed in proteomic studies:

- In a given sample (whole organism, tissue, cell, subcellular compartment), what fraction of the whole proteome is expressed (i.e., which proteins are present)?
- Of those proteins present in the sample, what are their relative abundances? The relative amounts of particular proteins provide a measure of the influence those proteins have on the behavior or function of a cell and how a cell might interact with its environment. Although the levels of mRNAs clearly influence the amounts of the corresponding protein made (see Chapter 8), there is relatively poor correlation between the relative amounts of proteins and corresponding mRNA abundance. This is because protein levels are also influenced by translation efficiency and regulation, post-translational modifications, and rates of protein degradation (see Chapter 9). Consequently, one

cannot always predict the level of a protein in a cell from its mRNA abundance. Thus even though there are rapid, efficient, and cost effective methods to measure mRNA levels of many genes in samples, including single cells (see Chapter 6), direct measurement of protein levels is necessary to understand the mechanisms that control protein abundance and the roles of proteins in cellular functions.

- What are the relative amounts of the different splice forms and chemically modified forms (e.g., phosphorylated, methylated, fatty acylated) of the proteins?
- Which proteins are present in large multiprotein complexes, and which proteins are in each complex? What are the functions of these complexes, and how do they interact?
- When the state (e.g., growth rate, stage of cell cycle, differentiation, stress level) of a cell changes, do the proteins in the cell, or those secreted from the cell, change in a characteristic (*fingerprint*-like) pattern? Which proteins change, and how (relative amounts, modifications, splice forms, etc.)? [Answering these questions requires a form of *protein expression profiling* that complements the *transcriptional (mRNA) profiling* discussed in Chapter 8.]
- Can such fingerprint-like changes be used for diagnostic purposes? For example, do certain cancers or heart disease cause characteristic changes in blood proteins or do certain patterns in the levels of proteins predict an increased risk for disease? Can the proteomic fingerprint help determine if a given cancer is resistant or sensitive to a particular chemotherapeutic drug? [Proteomic fingerprints can also be the

starting point for studies of the mechanisms underlying the change of state. Proteins (and other biomolecules) that show changes that are diagnostic of a particular state are called *biomarkers*.]

- Which proteins in a cell are in close proximity to one another? Which are joined in multiprotein complexes, stably reside close to one another as independent proteins (e.g., in the matrix or inner membrane space of the mitochondrion; see Chapter 12), or transiently come together due to some signaling process, such as when a hormone induces changes in the cell's state?
- Can changes in the proteome help define targets for drugs or suggest mechanisms by which a drug might induce toxic side effects? If so, it might be possible to engineer modified versions of the drug with fewer side effects. Proteomics methods can also be used to identify cellular proteins that bind to small-molecule drugs and thus identify potential protein targets of the drugs.
- Proteins undergo post-translational changes including chemical modifications, alterations in stability, interactions with other proteins, and changes in intracellular localization. What post-translational changes in proteins arise from changes (mutations, copy number variants) to genes in the genome? Some of these post-translational changes in proteins (e.g., phosphorylation or dephosphorylation) can only be identified by direct examination of the proteins themselves.

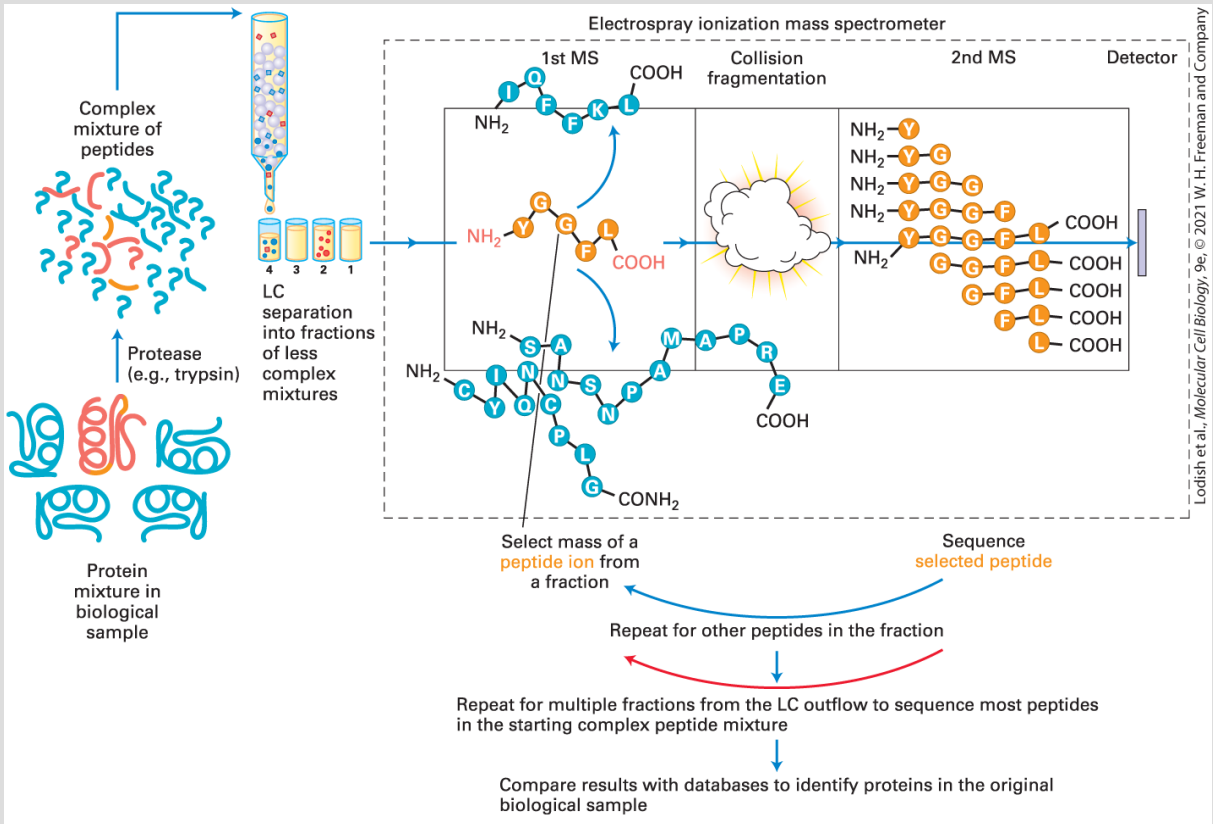
These are just a few of the questions that can be addressed using proteomics. The methods used to answer these questions are as

diverse as the questions themselves, and their numbers are growing rapidly.

Advanced Techniques in Mass Spectrometry Are Critical to Proteomic Analysis

Advances in proteomics technologies (e.g., mass spectrometry) profoundly affect the types of questions that can be practically studied. A widely used method to identify the protein components in a complex biological sample is *high-throughput LC-MS/MS*. Figure 3-50 outlines the general LC-MS/MS approach. In this approach, a complex mixture of proteins is digested with a protease, and the myriad resulting peptides are fractionated by LC into multiple, less complex fractions. After exiting the LC column, the fractions are continuously subjected to electrospray ionization and injection into a tandem mass spectrometer. The first fraction from the LC column enters the mass spectrometer and the multiple peptides (orange and teal) in that fraction are trapped as ions in the first MS analyzer. One of these (orange) is selected based on its mass for fragmentation and then its fragments are characterized in the second MS analyzer to determine its sequence. This process of selection, fragmentation, and sequencing is repeated sequentially for the other peptides in the first fraction that remains trapped in the first MS analyzer. Each of the subsequent fractions that enter the mass spectrometer are then

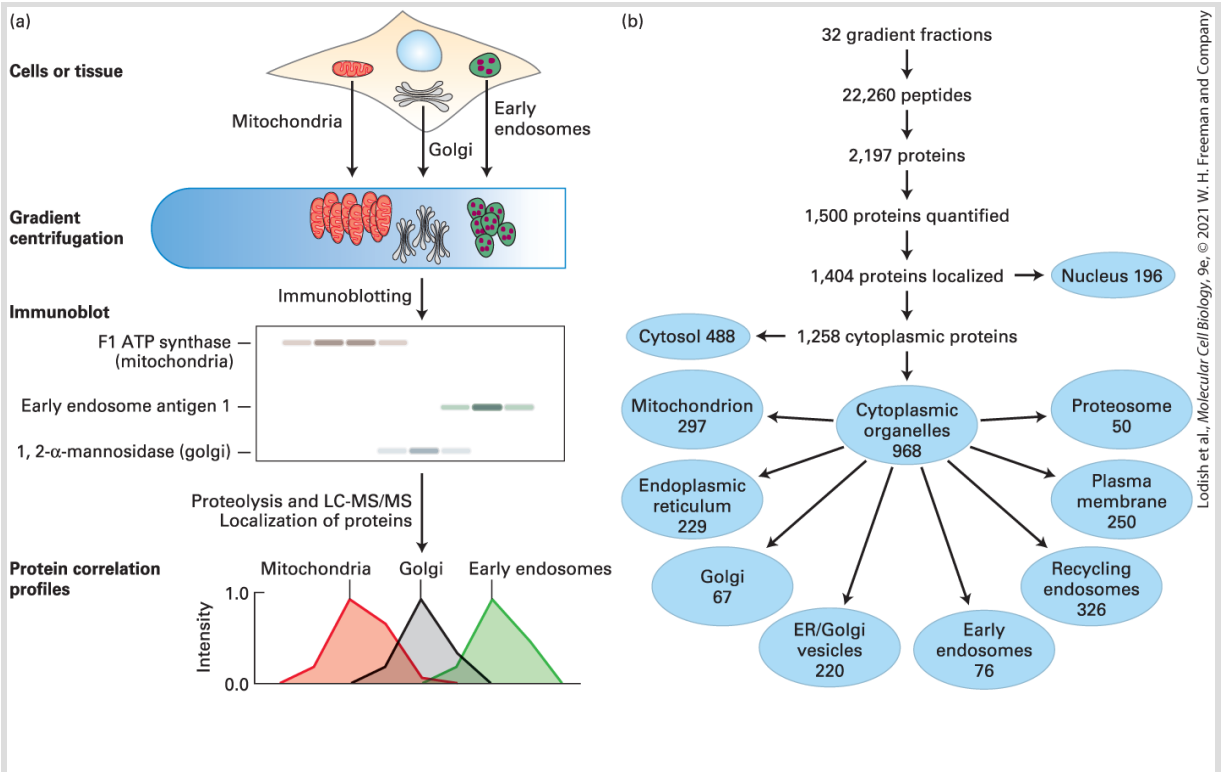
sequentially subjected to the same series of multiple cycles of MS/MS (peptide selection, fragmentation, sequencing) until the sequences of many of the peptides in the original mixture of peptides have been determined. Then, computations methods are used to identify, from the large collection of sequenced short peptides and protein databases, virtually all the proteins in the original biological sample. Using just a few 10's of μg of protein per sample, a substantial fraction of the actively expressed proteome in cells or tissues can be detected and quantified by this method, amounting to 10,000–12,000 proteins. At present, as many as 5000–7000 distinct proteins can be detected from a sample of as little as 1–2 μg of protein. This much sample can be obtained from 10,000–150,000 cells, depending on the types of cells used and losses of material during sample preparation. Efforts are under way to increase the sensitivity of the method so that eventually one might be able to analyze the entire proteome of an individual cell.



EXPERIMENTAL FIGURE 3-50 LC-MS/MS is used to identify the proteins in a complex biological sample. A complex mixture of proteins in a biological sample (e.g., an isolated preparation of Golgi organelles) is digested with a protease. The mixture of resulting peptides is fractionated by liquid chromatography (LC) into multiple, less complex, fractions, which are slowly but continuously injected by electrospray ionization into a tandem mass spectrometer. From the peptides in each fraction trapped in the first MS analyzer (orange and teal), one (orange) is selected for fragmentation and sequencing, and the remaining peptides in that fraction are sequentially fragmented and sequenced. The subsequent fractions to enter the mass spectrometer are then sequentially subjected to the same process of multiple cycles of MS/MS until the masses and sequences of many of the peptides in the original complex mixture have been determined and used to identify by computation the proteins in the original biological sample through comparison with protein databases.

An example of the use of LC-MS/MS to identify many of the proteins in each organelle is seen in Figure 3-51. Cells from murine (mouse)

liver tissue were mechanically broken to release the organelles, and the organelles were partially separated by density-gradient centrifugation. The locations of the organelles in the gradient were determined using immunoblotting with antibodies that recognized previously identified, organelle-specific proteins. Fractions from the gradient were then subjected to LC-MS/MS to identify the proteins in each fraction, and the distributions in the gradient of many individual proteins were compared with the distributions of the organelles. This strategy permitted the assignment of many individual proteins to one or more organelles (organelle proteome profiling). More recently, a combination of organelle purification, MS, biochemical localization, and computational methods has been used to show that at least a thousand distinct proteins are localized in the mitochondria of humans and mice.



EXPERIMENTAL FIGURE 3-51 Density-gradient centrifugation and LC-MS/MS can be used to identify many of the proteins in organelles. (a) The cells in liver tissue were mechanically broken to release the organelles, and the organelles were partially separated by density-gradient centrifugation. The locations of the organelles — which were spread out through the gradient and somewhat overlapped with one another — were determined using immunoblotting with antibodies that recognized previously identified, organelle-specific proteins. Fractions from the gradient were subjected to proteolysis and LC-MS/MS to identify the peptides, and hence the proteins, in each fraction. Comparisons with the locations of the organelles in the gradient (called *protein correlation profiling*) permitted assignment of many individual proteins to one or more organelles (organelle proteome identification). (b) The hierarchical breakdown of data derived from the procedures in part (a). Note that not all proteins identified could be assigned to organelles and that some proteins were assigned to more than one organelle.

[Data from L. J. Foster et al., 2006, *Cell* **125**:187–199.]

A powerful method to identify proteins that are spatially near to one another, either in direct physical contact or simply co-localized in the cell, is enzyme-catalyzed, proximity-dependent labeling coupled with mass spectrometry. In this method, cells are engineered with recombinant DNA to express a protein of interest that has fused to it a specialized enzyme. When the cells are exposed to the substrates of the enzyme, the enzyme converts a substrate to a highly reactive chemical that can covalently label proteins in the immediate vicinity (approximately 10–20 nm) of the fused protein of interest. One must be careful to show that the fusion of the enzyme to the protein does not disrupt its normal location and binding partners in the cell. The most common substrate used for proximity-dependent labeling contains the small molecule biotin, which is activated by the enzyme and then covalently attaches to nearby proteins. After the

labeling reaction in the living cells, all proteins can be extracted from the cells and the biotinylated proteins — those near the fused protein of interest — separated by affinity chromatography (see Figure 3-43c) using a column with beads containing instead of an antibody the protein streptavidin that binds biotin very tightly. The identities of the isolated, biotinylated proteins are then determined by proteolysis and LC-MS/MS. This approach has been used to identify 495 proteins that reside in the subcompartment of mitochondria called the *matrix* (Chapter 12) and the proteins in other subcompartments in cells that are difficult to purify biochemically. Additional experiments are required to determine the functional relationships of proteins shown to be in close proximity to each other.

Proteomics methods combined with molecular genetics methods are currently being used to identify all the protein complexes in eukaryotic cells. For example, in the yeast *Saccharomyces cerevisiae*, approximately 500 complexes, with an average of 4.9 distinct proteins per complex, have been identified. These complexes, in turn, are involved in at least 400 complex-to-complex interactions. Such systematic proteomic studies are providing new insights into the organization of proteins within cells and into how proteins work together to permit cells to live and function.

Phosphoproteomics, the identification and quantification of phosphorylation sites on the proteins in a complex mixture, is

playing a growing role in the analysis of cell metabolism and regulation. As we have already learned, the reversible phosphorylation of proteins by kinases and phosphatases is a key mechanism for regulating proteins in cells. Phosphoproteomics permits the simultaneous determination of the phosphorylation states of many proteins and thus provides an important tool for analyzing complex cellular regulatory networks. Only a fraction — in some cases, only a small fraction — of a particular protein might be phosphorylated. Thus phosphoproteomic analysis can require 50–100 times more initial cell or tissue sample material (from about 2.5 to more than 20 mg of total cellular protein per sample) than does standard proteomic analysis. As a consequence, investigators usually use affinity chromatography methods with either metal-containing (e.g., Fe^{3+} or TiO_2) or antibody-containing columns to separate phosphopeptides from nonphosphorylated peptides (phosphopeptide enrichment) prior to subjecting the phosphopeptides to LC-MS/MS analysis.

KEY CONCEPTS OF SECTION 3.6

Proteomics

- Proteomics is the systematic study of the amounts (and changes in the amounts), modifications, interactions, localization, and functions of all or subsets of all proteins in biological systems at the whole-organism, tissue, cellular, and subcellular levels.
- Proteomics provides insights into the fundamental organization of proteins within cells and how that organization is influenced by the state of the cells (e.g., differentiation into distinct cell types; responses to stress, disease, and drugs).
- A wide variety of methods are used for proteomic analyses, including two-dimensional gel electrophoresis, density-gradient centrifugation, and mass spectrometry

(particularly LC-MS/MS).

- Proteomics has helped begin to identify the proteomes of organelles ("organelle proteome profiling") as well as the organization of individual proteins into multiprotein complexes (see Figure 3-51).
- Phosphoproteomics is a specialized application of proteomics that identifies the collection of phosphorylated proteins (phosphoproteome) in cells and characterizes how the level of phosphorylation of these proteins varies as the state of the cells changes.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

active site
allostery
alpha (α) helix
autoradiography
basic helix-loop-helix (bHLH)
beta (β) sheet
beta (β) turn
biomolecular condensate
chaperone
conformation
dalton
denaturation

domain
EF hand
electrophoresis
enzyme
helix-loop-helix
helix-turn-helix
homolog
homology
immunoblotting
immunoprecipitation (IP)
intrinsically disordered region (IDR)
isoelectric point (pI)
kinase
leucine zipper
ligand
liquid chromatography (LC)
Michaelis constant, K_m
molecular complementarity
oligopeptide
peptide
peptide bond
phosphatase
phosphorylation
polypeptide
post-translational modification (PTM)
primary structure
protease
proteasome

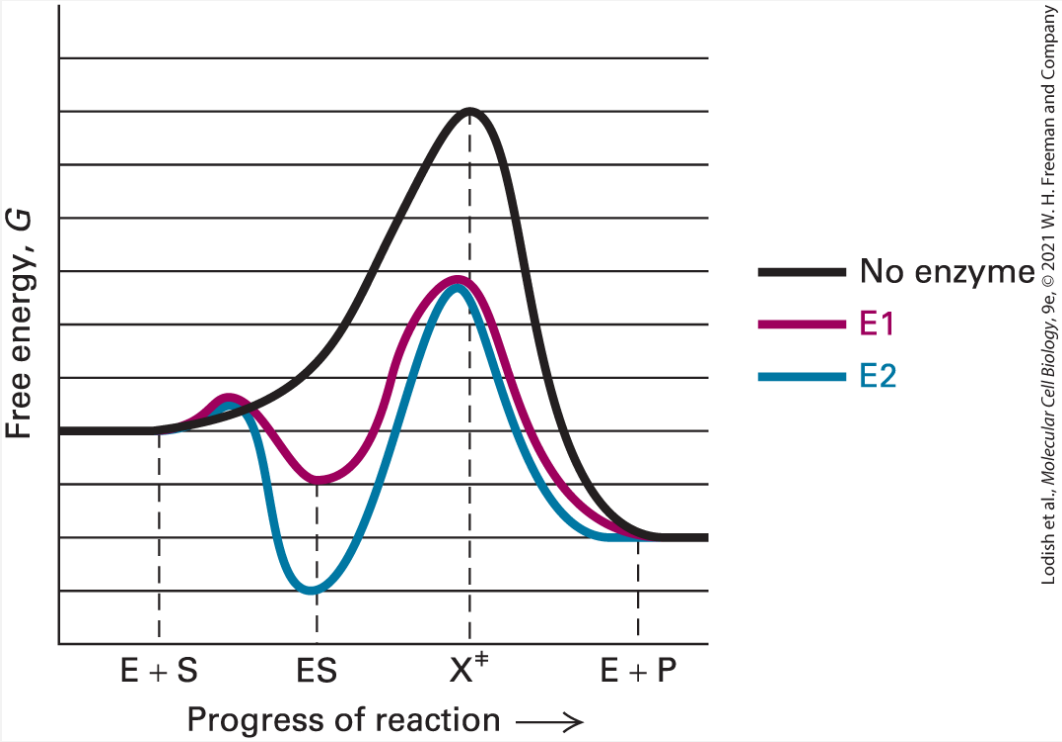
protein
proteome
proteomics
proteostasis
pulse-chase
quaternary structure
radioisotope
rate-limiting step
ribozyme
secondary structure
structural motif
substrate
tertiary structure
ubiquitinylation
V_{max} (maximal velocity)
x-ray crystallography
zinc finger

Review the Concepts

1. The three-dimensional structure of a protein is determined by its primary, secondary, and tertiary structures. Define the *primary, secondary*, and *tertiary structures*. What are some of the common secondary structures? What are the forces that hold together the secondary and tertiary structures?

2. Proper folding of proteins is essential for their biological activity. In general, the functional conformation of a protein is the conformation with lowest energy. This means that if an unfolded protein is allowed to reach equilibrium, it should assemble automatically into its native, functioning folded state. Why then is there a need for molecular chaperones and chaperonins in cells? What different roles do molecular chaperones and chaperonins play in the folding of proteins?
3. Enzymes catalyze chemical reactions. What constitutes the active site of an enzyme? What are the turnover number, the Michaelis constant (K_{m}), and the maximal velocity (V_{max}) of an enzyme? The turnover number for carbonic anhydrase is 5×10^5 molecules per second. This is a "rate constant," but not a "rate." What is the difference? By what concentration would you multiply this rate constant in order to determine an actual rate of product formation (V)? Under what circumstances would this rate become equal to the maximal velocity (V_{max}) of the enzyme?
4. The following reaction-coordinate diagram charts the energy of a substrate molecule (S) as it passes through a transition state ($X^{\ddagger}$) on its way to becoming a stable product (P) alone or in the presence of one of two different enzymes (E1 and E2). How does the addition of either enzyme affect the change in Gibbs free energy (ΔG) for the reaction? Which of the two enzymes binds with greater affinity to the substrate? Which enzyme better stabilizes

the transition state? Which enzyme functions as a better catalyst?



5. A healthy immune system can raise antibodies that recognize and bind with high affinity to almost any stable molecule. The molecule to which an antibody binds is known as an *antigen*. Antibodies have been exploited by enterprising scientists to generate valuable tools for research, diagnosis, and therapy. One clever application is the generation of antibodies that function like enzymes to catalyze complicated chemical reactions. If you wished to produce such a catalytic antibody, what would you suggest using as the antigen? Should it be the substrate of the reaction? The product? Something else?

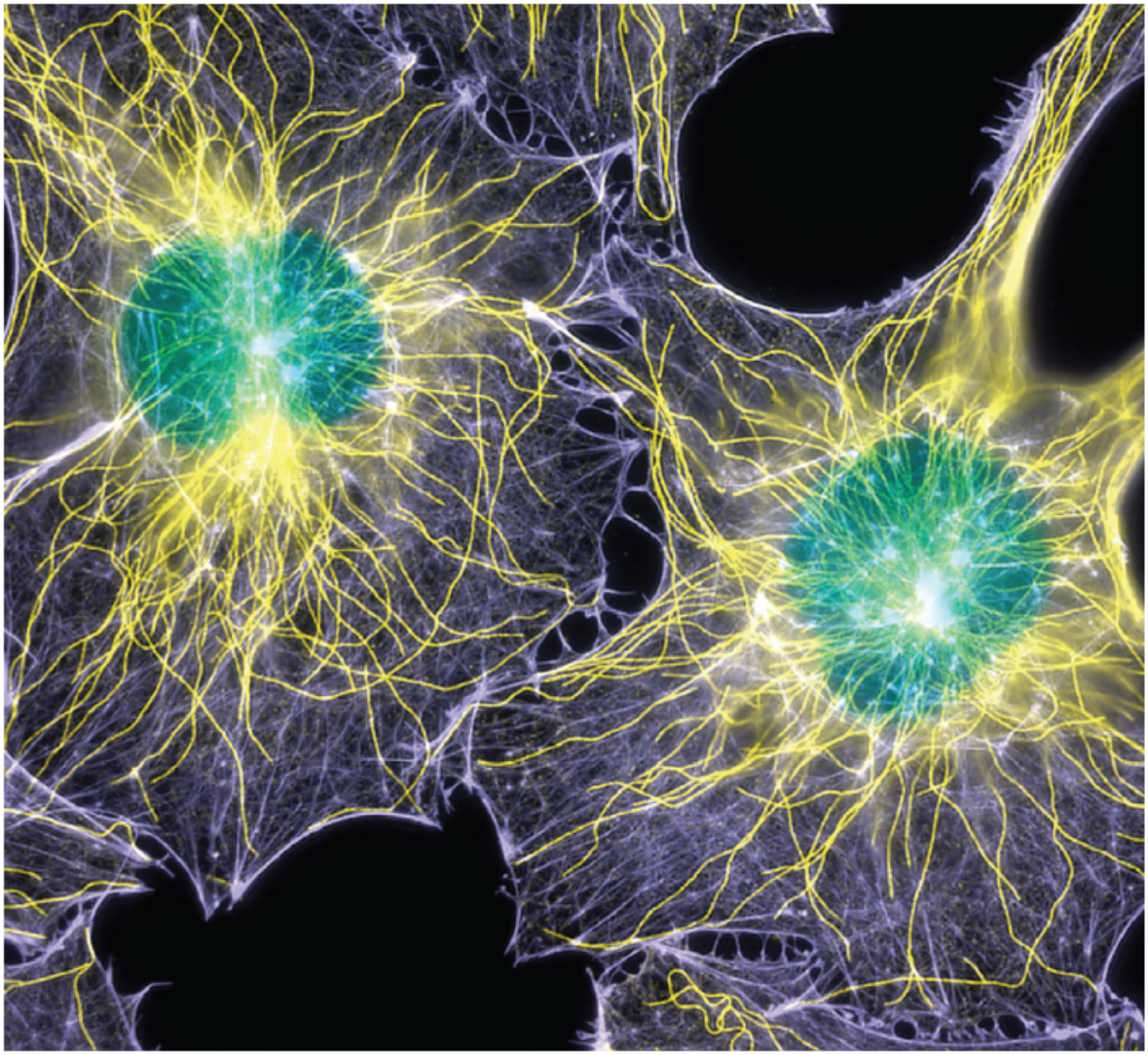
6. Proteins are degraded in cells. What is ubiquitin, and what role does it play in tagging proteins for degradation? What is the role of proteasomes in protein degradation? How might proteasome inhibitors serve as chemotherapeutic (cancer-treating) agents?
7. The function of proteins can be regulated in a number of ways. What is cooperativity, and how does it influence protein function? Describe how protein phosphorylation and proteolytic cleavage can modulate protein function.
8. A number of techniques can separate proteins on the basis of their differences in mass. Describe the use of two of these techniques, centrifugation and gel electrophoresis. The blood proteins transferrin (MW 76 kDa) and lysozyme (MW 15 kDa) can be separated by rate-zonal centrifugation or SDS-polyacrylamide gel electrophoresis. Which of the two proteins will sediment faster during centrifugation? Which will migrate faster during electrophoresis?
9. Liquid chromatography is an analytical method used to separate proteins. Describe the principles for separating proteins by gel filtration, ion-exchange, and affinity chromatography.
10. Various methods have been developed for detecting proteins. Describe how radioisotopes and autoradiography can be used for labeling and detecting proteins. How does Western blotting detect proteins?
11. Physical methods are often used to determine protein conformation. Describe how x-ray crystallography, cryoelectron microscopy, and NMR spectroscopy can be

used to determine the shapes of proteins. What are the advantages and disadvantages of each method? Which is better for small proteins? Large proteins? Huge macromolecular assemblies?

12. Mass spectrometry is a powerful tool in proteomics. What are the four key features of a mass spectrometer? Describe briefly how MALDI and two-dimensional polyacrylamide gel electrophoresis could be used to identify a protein expressed in cancer cells but not in normal healthy cells.

CHAPTER 4

Culturing and Visualizing Cells



Torsten Wittman, University of California, San Francisco.

Fluorescence microscopy showing the locations of DNA (green), microtubules (yellow), and microfilaments (purple) in two cultured cells. The cells were chemically fixed and then rendered permeable to antibodies using a gentle detergent. Microtubules were stained with

an antibody to tubulin; microfilaments were stained with a labeled toxin, phalloidin, that binds selectively to F-actin; and DNA was visualized with a DNA-binding dye.

OUTLINE

4.1 Growing and Studying Cells in Culture
4.2 Light Microscopy: Exploring Cell Structure and Visualizing Proteins Within Cells
4.3 Electron Microscopy: High-Resolution Imaging
4.4 Isolation of Cell Organelles

It is difficult to believe that just 200 years ago it was not yet appreciated that all living things are made of cells. In 1655, Robert Hooke used a primitive microscope to examine a piece of cork and saw an orderly arrangement of rectangles — the walls of the dead plant cells — that reminded him of monks' cells in a monastery, so he coined the term *cell*. Shortly after this, Antoni van Leeuwenhoek's observations of the microorganisms he saw in his simple microscope became the first description of live cells. Two hundred years later, Matthias Schleiden and Theodore Schwann proposed that cells constitute the fundamental unit of life in plants, animals, and single-celled organisms. Collectively, these discoveries were some of the greatest in biology and posed the question of how cells are organized and function.

Even today, many technical constraints hamper studies of cells in intact animals and plants. One alternative is the use of intact organs that are removed from animals and treated to maintain their physiological integrity and function. However, the organization of

organs, even isolated ones, is sufficiently complex to pose numerous problems for research. Thus molecular cell biologists often conduct experimental studies on cells isolated from an organism. In Section 4.1, we learn how to maintain and grow diverse cell types and how to isolate specific types of cells from complex mixtures.

In many cases, isolated cells can be maintained in the laboratory under conditions that permit their survival and growth, a procedure known as *culturing*. Cultured cells have several advantages over intact organisms for cell biology research. Cells of a single specific type can be grown in culture, experimental conditions can be better controlled, and in many cases a single cell can be readily grown into a colony of many identical cells. The resulting strain of cells, which are genetically homogeneous, is called a **clone**. However, cultured cells are not in their native setting, so researchers are now growing and examining cells in three-dimensional environments to more closely mimic their situation in an animal.

Discoveries about cellular organization have been intimately tied to developments in microscopy. This is as true today as it was 400 years ago. Light microscopy initially revealed the beautiful internal organization of cells, and today highly sophisticated microscopes are continually being improved to probe deeper and deeper into the molecular mechanisms by which cells function. In Section 4.2, we discuss light microscopy and the long-standing but still valuable techniques that are in use today, and then examine several clever

methods that have been developed more recently, culminating with the newest, cutting-edge technologies. A major advance came in the 1960s and 1970s with the development of immunofluorescence microscopy, which allows the localization of specific proteins within fixed cells, thus providing a static image of their location, as illustrated in the chapter-opening figure. Such studies led to the important understanding that the membranes and interior spaces of each type of organelle contain a distinctive group of proteins that are necessary for the organelle to carry out its unique functions. Another major advance came in the mid-1990s with the simple idea of expressing *chimeric proteins* — consisting of a protein of interest covalently linked to a naturally fluorescent protein — which has enabled biologists to visualize movements of individual proteins in live cells. Suddenly, the dynamic nature of cells could be appreciated. This has dramatically changed our view of cells when we were limited to only static images. Microscopes with higher sensitivity coupled with methods to enhance resolution now allow investigators to gain more information about the detailed structure and dynamics of cells. The development of fluorescence techniques also led to the development of methods to monitor protein-protein interactions in live cells, as well as a myriad of other sophisticated molecular technologies, some of which we discuss in this section. For decades, light microscopy was constrained by the resolution of the light microscope — to about 200 nm — due to the limitations imposed by the wavelength of visible light. We discuss methods that

have been developed over the last few years to beat this resolution barrier with the development of *super-resolution microscopy*.

Despite the amazing developments in light microscopy, visible light still provides too low a resolution to examine cells in fine, ultrastructural detail. The electron microscope gives a much higher resolution, but the technology generally requires that the cell be fixed and sectioned, and therefore all cell movements are frozen in time. Electron microscopy also allows investigators to examine the structure of macromolecular complexes or single macromolecules. In Section 4.3, we outline the various approaches for preparing specimens for observation in the electron microscope and describe the types of information that can be derived from them.

Light and electron microscopy revealed that all eukaryotic cells contain a similar repertoire of membrane-limited compartments termed *organelles*. In parallel with the developments in microscopy, subcellular fractionation methods were developed that have enabled cell biologists to isolate individual organelles to a high degree of purity. These techniques, detailed in Section 4.4, continue to provide important information about the protein composition and biochemical function of organelles.

4.1 Growing and Studying Cells in Culture

The study of cells is greatly facilitated by growing them in culture, so that they can be examined by microscopy and subjected to specific treatments under controlled conditions. It is generally quite easy to grow unicellular bacteria, fungi, or protists, for example, by placing them in a rich medium that supports their growth. However, animal cells come from multicellular organisms, which makes it more difficult to culture single cells or small groups of cells. In this section, we discuss how animal cells are grown in culture and how different cell types can be purified for study.

Culture of Animal Cells Requires Nutrient-Rich Media and Special Solid Surfaces

To permit the survival and normal function of cultured tissues or cells, the temperature, pH, ionic strength, and access to essential nutrients must simulate as closely as possible the conditions within an intact organism. Isolated animal cells are typically placed in a nutrient-rich liquid, called the *culture medium*, within specially coated plastic dishes or flasks. The cultures are kept in incubators in which the temperature, atmosphere, and humidity can be

controlled. To reduce the chances of bacterial or fungal contamination, antibiotics are often added to the culture medium. To further guard against contamination, investigators usually transfer cells between dishes, add reagents to the culture medium, and otherwise manipulate the specimens within special sterile cabinets containing circulating air that is filtered to remove microorganisms and other airborne contaminants.

Media for culturing animal cells must supply the nine amino acids (phenylalanine, valine, threonine, tryptophan, isoleucine, methionine, leucine, lysine, and histidine) that cannot be synthesized by adult vertebrate animal cells. In addition, most cultured cells require three other amino acids (cysteine, tyrosine, and arginine) that are synthesized only by specialized cells in intact animals, as well as glutamine, which serves as a nitrogen source. The other necessary components of a medium for culturing animal cells are vitamins, various salts, fatty acids, glucose, and serum — the fluid remaining after the noncellular part of blood (plasma) has been allowed to clot. Serum contains various protein factors that are needed for the proliferation of mammalian cells in culture, including the polypeptide hormone insulin; transferrin, which supplies iron in a bioaccessible form; and numerous growth factors. In addition, certain cell types require specialized protein growth factors not present in serum. For instance, progenitors of red blood cells require erythropoietin, and T lymphocytes require interleukin 2 (see Chapter 16). A few mammalian cell types can be grown in a chemically defined, serum-free medium containing amino acids,

glucose, vitamins, and salts plus certain trace minerals, specific protein growth factors, and other components.

Unlike bacterial and yeast cells, which can be grown in suspension, most animal cell types will grow only when attached to a solid surface. This requirement highlights the importance of the cell-surface proteins, called **cell-adhesion molecules (CAMs)**, that cells use to bind to adjacent cells and to components of the extracellular matrix such as collagen, laminin, or fibronectin (see Chapter 20). The solid growth surface (usually glass or plastic) is either pre-coated with these extracellular-matrix proteins, or they come from the serum or are secreted by the cells in culture. A single cell cultured on a glass or plastic dish proliferates to form a visible mass, or *colony*, containing thousands of genetically identical cells in 4 to 14 days, depending on the growth rate. Although most normal animal cells require a surface to grow on, some specialized blood cells, and especially tumor cells, can be grown in suspension as single cells.

Primary Cell Cultures and Cell Strains Have a Finite Life Span

Primary cells are cells isolated directly from tissues. Normal animal tissues (e.g., skin, kidney, liver) or whole embryos are commonly used to establish *primary cell cultures*. To prepare individual tissue

cells for a primary culture, the cell-cell and cell-matrix interactions must be broken. To do so, tissue fragments are treated with a combination of a protease (e.g., trypsin, the collagen-hydrolyzing enzyme collagenase, or both) and a divalent cation chelator (e.g., EDTA) that depletes the medium of free calcium (Ca^{2+}). Many CAMs require calcium and are thus inactivated when calcium is removed; other CAMs that are not calcium dependent need to be cleaved by a protease for the cells to separate. The released cells are then placed in a nutrient-rich, serum-supplemented medium in dishes, where they can adhere to the surface and to one another. The same protease-chelator solution is used to remove adherent cells from a culture dish for biochemical studies or subculturing (transfer to another dish).

Fibroblasts are the predominant cells in connective tissue and normally produce extracellular-matrix components, such as collagen, that bind to CAMs, thereby anchoring cells to a surface. In culture, fibroblasts usually divide more rapidly than other cells from a tissue, eventually becoming the predominant cell type in a primary culture unless special precautions are taken to remove them when isolating other types of cells.

When cells removed from an embryo or an adult animal are cultured, most of the adherent cells will divide a finite number of times and then cease growing (a phenomenon called *cell senescence*). For instance, human fetal fibroblasts divide about 50 times before

they cease growth. Starting with 10^6 cells, 50 doublings has the potential to produce $10^6 \times 2^{50}$, or more than 10^{20}, cells, whose weight would be equivalent to that of about a thousand people. Normally, only a very small fraction of these cells are used in any one experiment. Thus even though its lifetime is limited, a single culture, if carefully maintained, can be studied through many cell generations. Such a lineage of cells originating from one initial primary culture is called a **cell strain**.

One important exception to the finite life span of normal cells is the *embryonic stem cell*, which, as its name implies, is derived from an embryo and will divide and give rise to all tissues during development. As we discuss in Chapter 22, embryonic stem cells can be cultured indefinitely under the appropriate conditions.

Research with cell strains is simplified by our ability to freeze them and successfully thaw them at a later time for experimental analysis. Cell strains can be frozen in a state of suspended animation and stored for extended periods at liquid nitrogen temperature, provided that a preservative that prevents the formation of damaging ice crystals is used. Although not all cells survive thawing, many do survive and resume growth.

Transformed Cells Can Grow Indefinitely in Culture

To be able to clone individual cells, modify cell behavior, or select mutants, biologists often want to maintain cell cultures for many more than 50 doublings. Such prolonged growth is exhibited by cells derived from some tumors. In addition, rare cells in a population of primary cells may undergo spontaneous oncogenic mutations, leading to *oncogenic transformation* (see Chapter 25). Such cells, said to be oncogenically transformed, or simply *transformed*, are able to grow indefinitely. A culture of cells with an indefinite life span is considered immortal and is called a **cell line**.

The HeLa cell line, the first human cell line established, was originally obtained in 1952 from a malignant tumor (carcinoma) of the uterine cervix and is still used extensively today. The cells were obtained from Henrietta Lacks without her consent, which has raised ethical issues concerning ownership of biological materials. Many other human cell lines are derived from cancers, and biologists have rendered others immortal by transforming them to express oncogenes.

Regardless of their source, cells in immortal lines often have chromosomes with abnormal DNA sequences. In addition, the number of chromosomes in such cells is usually greater than that in the normal cell from which they arose, and the chromosome number changes as the cells continue to divide in culture. A noteworthy exception is a human cell line of hematopoietic origin that is haploid for all chromosomes except chromosome 8. Since

inactivation of one of the two copies of a gene in a diploid cell generally does not generate a phenotype, a line with a single copy of most genes is very useful for genetic analysis, making possible the types of genetic screens employed in model organisms (see Chapter 6). Cells with an abnormal number of chromosomes are said to be *aneuploid*.

Flow Cytometry Separates Different Cell Types

Some cell types differ sufficiently in density that they can be separated on the basis of this physical property. White blood cells (leukocytes) and red blood cells (erythrocytes), for instance, have very different densities because erythrocytes have no nucleus; thus these cells can be separated by equilibrium density-gradient centrifugation (described in Section 4.4). Most cell types cannot be differentiated so easily, so other techniques, such as flow cytometry, must be used to separate them.

To separate one type of cell from a complex mixture, it is necessary to have some way to mark and then sort out the desired cells. As we will see below, it is possible to mark cells by expressing a fluorescent protein in them, but if only a few cells in the population express the protein, how can we sort them from the nonfluorescent ones? The cells can be analyzed in a *flow cytometer*. This machine flows cells

past a laser beam that measures the light that they scatter and the fluorescence that they emit; thus it can quantify the cells expressing the fluorescent protein in a mixture. A **fluorescence-activated cell sorter (FACS)**, which is based on flow cytometry, can both analyze the cells and select the few fluorescent cells from thousands of others and sort them into a separate culture dish (Figure 4-1). To achieve this, the cells are mixed with a buffer and forced through a vibrating nozzle to generate tiny droplets. The concentration of cells is adjusted so that most of the droplets do not contain cells, and the ones that do contain only one. Just before the nozzle, the stream of cells passes through a laser beam so that the presence and size of a cell can be recorded from the scattered light using one detector, and the amount of fluorescent light emitted can be quantified using a second, fluorescent light detector. If a cell is present in a droplet, the droplet is given a negative electric charge as it emerges from the nozzle. The stream of droplets then passes through two plates that generate an electric field proportional to the fluorescence detected from the cell in the droplet. This field generates a force that moves charged droplets out of the stream of uncharged droplets and into a collection tube. Since the amount of force applied is proportional to the fluorescence emitted by the cell in the droplet, cells with different levels of fluorescence can be collected. Having been sorted from other cells, the selected cells can be grown in culture.

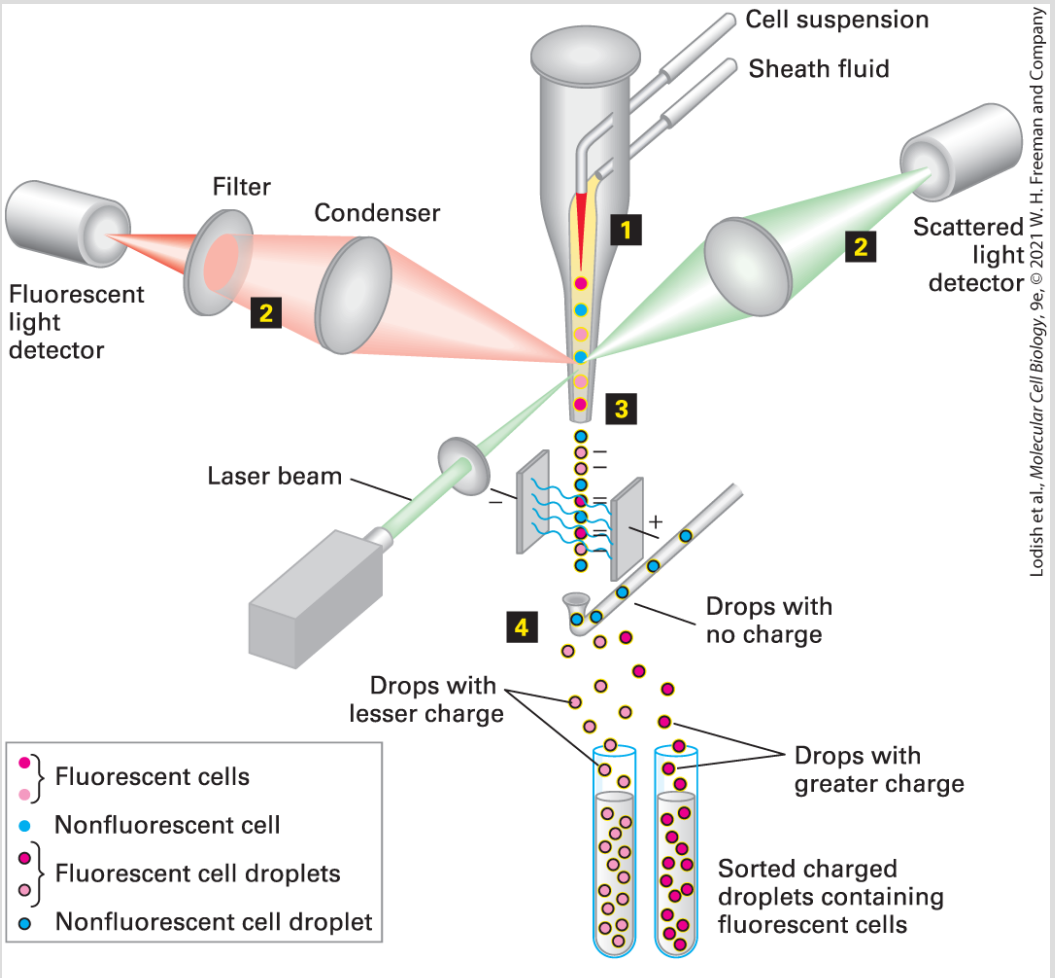
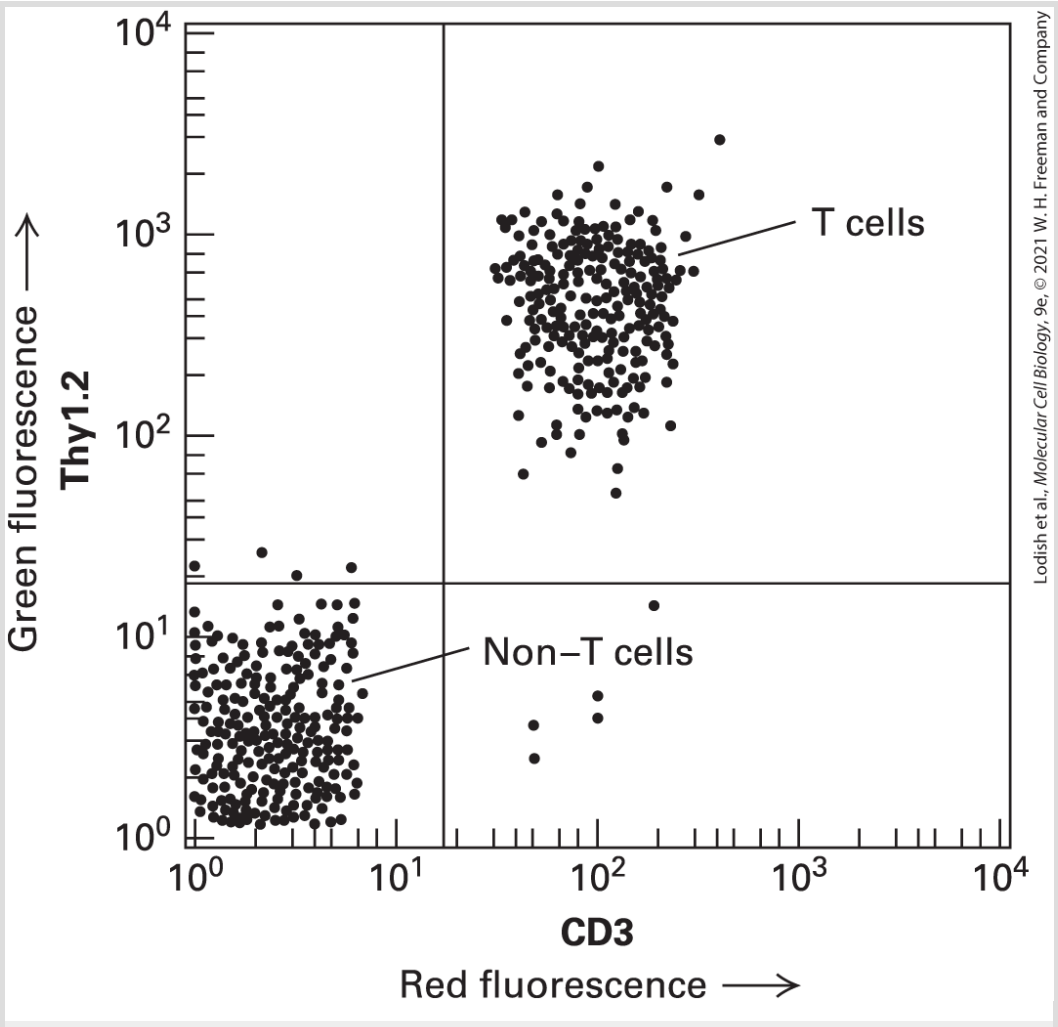


FIGURE 4-1 A fluorescence-activated cell sorter (FACS) separates cells having different levels of fluorescence. Step 1: A concentrated suspension of labeled cells is mixed with a buffer so that the cells pass single file through a laser light beam. Step 2: Both the fluorescent light emitted and the light scattered by each cell are measured; from measurements of the scattered light, the size and shape of the cell can be determined. Step 3: The suspension is then forced through a nozzle, which forms tiny droplets containing at most a single cell. At the time of formation at the nozzle tip, each droplet containing a cell is given a negative electric charge proportional to the fluorescence of that cell determined from the earlier measurement. Step 4: Droplets now pass through an electric field, so that those with no charge are discarded, whereas those with different electric charges are separated and collected. Because it takes only milliseconds to sort each droplet, as many as 10 million cells per hour can pass through the machine.

The FACS procedure is commonly used to purify different types of white blood cells, each of which bears on its surface one or more distinctive proteins and so will bind monoclonal antibodies specific for its proteins. If a cell mixture is incubated with a fluorescent dye linked to the antibody to a specific cell-surface protein, only the desired cells will be fluorescent. Only the T cells of the immune system, for instance, have both CD3 and Thy1 proteins on their surfaces. The presence of these surface proteins allows T cells to be separated easily from other types of blood cells or spleen cells (Figure 4-2).



EXPERIMENTAL FIGURE 4-2 FACS separates T cells with bound fluorescence-tagged antibodies to two cell-surface proteins from other white blood cells. Spleen cells from a mouse were treated with a red fluorescent monoclonal antibody specific for the CD3 cell-surface protein and with a green fluorescent monoclonal antibody specific for a second cell-surface protein, Thy1. As the cells passed through a FACS, the intensity of the green and red fluorescence emitted by each cell was recorded. Each dot represents a single cell. This plot of green fluorescence (vertical axis) versus red fluorescence (horizontal axis) for thousands of spleen cells shows that about half of them — the T cells — express both CD3 and Thy1 proteins on their surfaces (upper-right quadrant). The remaining cells, which exhibit low fluorescence (lower-left quadrant), express only background levels of these proteins and are other types of white blood cells. Note the logarithmic scale on both axes.

[Data from Chengcheng Zhang, Whitehead Institute.]

Flow cytometry can measure many different parameters in cells. For example, it can measure a cell's DNA content (from the amount of fluorescence emitted from a DNA-binding dye) and determine the general shape and size of a cell (from the amount of scattered light). Measurements of the DNA content of individual cells are used to follow replication of DNA as cells progress through the cell cycle (see Chapter 19). Flow cytometry can also be used to identify specific cell types in a population by using cell-type specific antibodies, each with a different fluorescent molecule attached. An example is distinguishing different cell types in blood samples. It can also be used to study signaling pathways at the single-cell level. Stimulation of many signaling pathways results in the phosphorylation of specific proteins (discussed in Chapters 15 and 16), and these can often be detected in cell populations by western

blot analysis using phospho-specific antibodies that recognize individual sites of phosphorylation on specific proteins (see Chapter 3). However, this yields a population average, so flow cytometry has been adapted to quantify the phosphorylation level of specific proteins in individual cells. Use is made of phospho-specific antibodies to which different fluorescent reporters have been covalently linked. For example one antibody might be linked to a fluorochrome that emits green fluorescence, and another antibody linked to one that emits red fluorescence. Cells are treated so that they are made permeable to antibodies, but their phosphorylated proteins are retained. The permeabilized cells are then incubated with phospho-specific antibodies and analyzed by flow cytometry to detect the degree of phosphorylation of the proteins, as reported by the amount of green or red signal. In this way, the level of protein phosphorylation can be assessed in response to activation of different signaling pathways.

An alternative method for separating specific types of cells uses small magnetic beads coupled to antibodies to a specific cell-surface molecule. For example, to isolate T cells, the beads are coated with a monoclonal antibody specific for a surface protein such as CD3 or Thy1. Only cells with these proteins will stick to the beads, which can be recovered from the preparation by attraction to a small magnet on the side of a test tube.

Growth of Cells in Two-Dimensional and Three-Dimensional Culture Mimics the In Vivo Environment

While much has been learned using cells grown on plastic or glass surfaces, these surfaces are far removed from cells' normal tissue environment. As detailed in Chapter 20, many cell types function only when closely linked to other cells. Key examples are the sheet-like layers of epithelial tissue, called **epithelia** (singular, **epithelium**), that cover the external and internal surfaces of organs. Typically, epithelial cells have distinct surfaces, called the *apical* (top), *basal* (base or bottom), and *lateral* (side) surfaces (see Figure 20-11). The basal surface usually contacts an underlying extracellular matrix called the *basal lamina*, whose composition and function are discussed in Section 20.3. Epithelial cells often function to transport specific classes of molecules across the epithelial sheet; for example, the epithelial lining of the intestine transports nutrients into the cell through the apical surface and out toward the bloodstream across the basolateral surface. When grown on plastic or glass, epithelial cells cannot easily perform this function. Therefore, special containers have been designed with a porous surface that acts as a basal lamina to which epithelial cells attach and form a uniform two-dimensional sheet (Figure 4-3). A commonly used cultured cell line derived from dog kidney epithelium, called the *Madin-Darby canine kidney* (*MDCK*) cell line, is often used to study the formation and function of epithelial sheets.

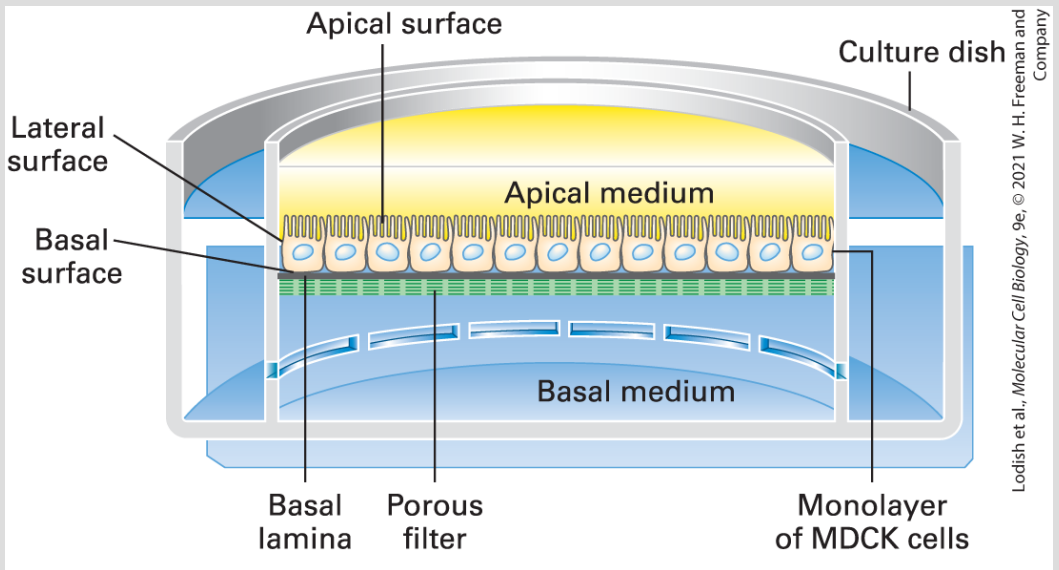
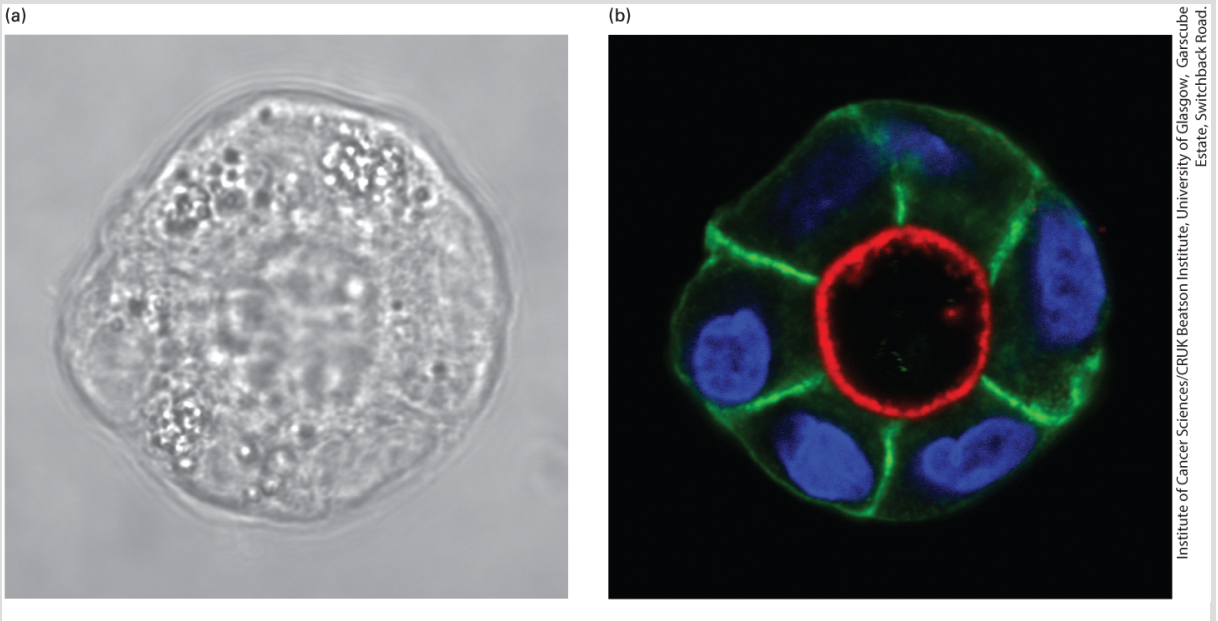


FIGURE 4-3 Madin-Darby canine kidney (MDCK) cells grown in specialized containers provide a useful experimental system for studying epithelial cells. MDCK cells form a polarized epithelium when grown on a porous membrane filter coated on one side with collagen and other components of the basal lamina. With the use of the special culture dish shown here, the medium on each side of the filter (apical and basal sides of the monolayer) can be experimentally manipulated and the movement of molecules across the layer monitored.

However, even a two-dimensional sheet often does not allow cells to fully mimic behavior in their normal environment. Methods have now been developed to grow cells in three dimensions by providing a support infiltrated with appropriate components of the extracellular matrix (discussed in Chapter 20). If MDCK cells are cultured under appropriate conditions, they will form a tubular sheet mimicking a tubular organ or the duct of a secretory gland. In these three-dimensional structures, the apical side of the epithelial sheet lines the lumen, whereas the basal side of each cell is in contact with the extracellular matrix (Figure 4-4).



EXPERIMENTAL FIGURE 4-4 MDCK cells can form cysts in culture. (a) MDCK cells grown on a supported extracellular matrix will form groups of cells that polarize to form a tubular single layer of cells with a lumen in the middle, called a *cyst*. (b) By examining the localization of proteins found in the apical (red) and basolateral membranes (green), we can see that these cells are fully polarized, with the apical side facing the lumen, which recapitulates their organization in the kidney tubules from which they are derived. The nuclear DNA is stained blue.

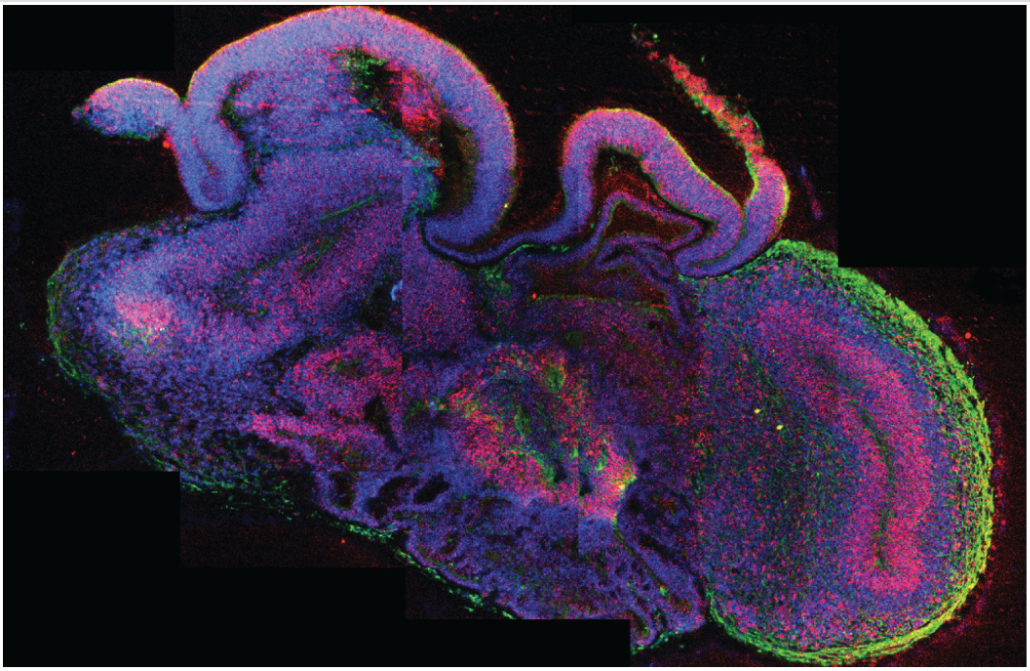
In a similar manner, primary hepatocytes (cells isolated from a liver) can now be isolated and cultured. Under proper conditions, they will form hepatocyte spheroids that can be maintained in culture and respond to drugs in much the same way that an intact liver does in living animals. These are just two examples of a growing field in which organ culture systems are being developed to more closely correspond to cells and tissues in living animals — and thereby expand the number of experiments that can be performed in a physiologically relevant manner.

Stem Cells Can Differentiate in Culture to Make Organoids

As we discuss in detail in Chapter 22, embryonic stem cells can be isolated from the inner cell mass of an early mammalian embryo. These cells are *pluripotent*, that is, they can give rise to all tissues of the embryo. Mature animals have *adult stem cells* in their tissues. These cells are not pluripotent, but are nevertheless crucial for cell regeneration or tissue repair. The division and differentiation of cells in an animal is orchestrated by many factors, including interactions with the extracellular matrix and other cells, response to soluble factors, and temporal changes in gene expression. With major advances in the study of the extracellular matrix (discussed in Chapter 20), signaling pathways elucidated by developmental studies (see Chapters 15 and 16), and the regulation of gene expression (discussed in Chapter 8), much has been learned about the molecular details of these differentiation pathways. It is now possible to recapitulate some of these developmental events in culture.

When adult stem cells from the intestine were placed in culture under appropriate conditions in a three-dimensional matrix, the cells formed intestinal organoids that had the normal crypt and villous structures of the small intestine. In a dramatic advance, researchers have used human pluripotent cells in culture to generate brain organoids whose organization and structure is

remarkably similar to the fetal brain (Figure 4-5). Organoids have now been developed for many organs, including the brain, intestines, stomach, kidney, heart, and tongue. So far, organoids advance to a stage of development characterized in the fetus, but do not proceed to the organization seen in an adult organ. These astonishing results reveal that cells possess the intrinsic program to divide, differentiate, and organize themselves into complex structures.



EXPERIMENTAL FIGURE 4-5 Brain organoid derived from pluripotent stem cells. In this section of the organoid, neural stem cells are labeled red and neurons are green. All cell nuclei are stained blue.

Besides being a remarkable achievement, what uses can be made of organoids? First, cells in an organoid can be followed much more easily than in a living animal. For example, in brain organoids, it

will be possible to examine early human brain development in much more detail. Second, stem cells used to generate organoids can be manipulated by introducing mutations and then observed to see how they affect the development of the organoid. Third, it is possible to test the effects of drugs on organoid development, which is much easier than in animals. Fourth, it has been found that tumor cells removed from a patient can be used to generate a morphologically characteristic tumor organoid in culture — which can be used to test the efficacy of potential drugs before they are administered to a patient. These are just a few of the remarkable uses for organoids; many more creative uses for both basic science and medicine are being developed.

Can we grow a whole organ that could be transplanted into a patient? As indicated above, organoids grown in culture so far do not develop into adult organs. However, recent advances in biomedical engineering are developing promising strategies to do this, initially in experimental animals. In an example of one approach, a 3-D printer is used to help make a replacement ear. First, an exact computer image of an ear is generated. This image is used to program a 3-D printer to assemble a pliable matrix — containing support material that is biodegradable, together with appropriate components of the extracellular matrix — in the precise shape of an ear. This matrix provides support for the growth of skin cells, either in culture or after transplantation under the skin, so that ultimately the synthetic organ can be surgically attached to a

living animal. Other approaches make use of 3-D printers to assemble the matrix and seed it with appropriate cells. An exciting and ambitious goal of this technology is to generate synthetic organs containing many different types of cells by printing each of several layers with the appropriate matrix and cells to generate complex three-dimensional organs that might one day be used to replace defective ones in patients. Many hurdles still need to be overcome, but the ability to generate stem cells from patients and then induce differentiation in culture is overcoming the major obstacle of immunological rejection and will probably be key to providing cells for the assembly of synthetic organs. ■

Hybridomas Produce Abundant Monoclonal Antibodies

In addition to serving as research models for studies of cell function, cultured cells can be converted into "factories" for producing specific proteins. For example, special cultured cells can be used to generate monoclonal antibodies, which are experimental tools widely used in many aspects of cell biological research. They are also used for diagnostic and therapeutic purposes in medicine, as we discuss in later chapters.

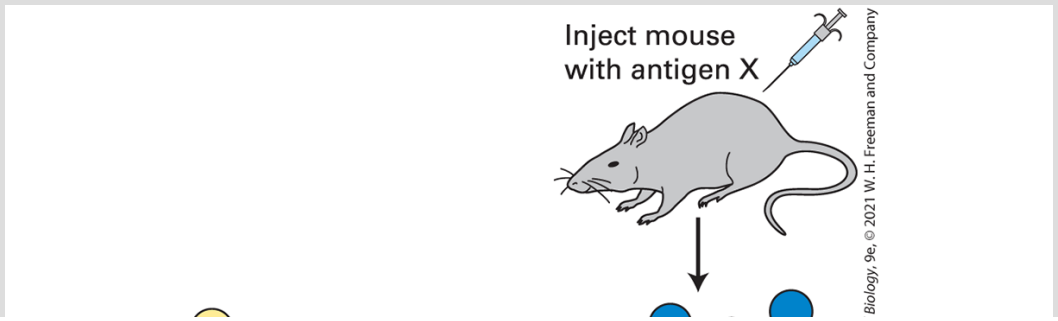
To understand the challenge of generating monoclonal antibodies, we must briefly review how mammals produce antibodies; more

detail is provided in Chapter 24. Recall that antibodies are proteins secreted by white blood cells that bind with high affinity to their antigen (see Figure 3-22). Each normal antibody-producing B lymphocyte in a mammal is capable of producing a single type of antibody that can bind to a particular determinant, or **epitope**, on an antigen molecule. An epitope is generally a small region on the antigen, consisting, for example, of just a few amino acids. If an animal is injected with an antigen, the B lymphocytes that make antibodies recognizing that antigen are stimulated to grow and secrete those antibodies. Each antigen-activated B lymphocyte forms a clone of cells in the spleen or lymph nodes, with each cell producing the identical antibody — that is, a **monoclonal antibody**. Because most natural antigens contain multiple epitopes, exposure of an animal to an antigen usually stimulates the formation of multiple B-lymphocyte clones, each producing a different antibody. The resulting mixture of antibodies from the many B-lymphocyte clones that recognize different epitopes on the antigen is said to be *polyclonal*. Such **polyclonal antibodies** circulate in the blood and can be isolated as a group.

Although polyclonal antibodies are very useful, monoclonal antibodies are more suitable for many types of experiments and medical applications in which we need a reagent that binds to just one site on a protein; for example, one that competes with a ligand on a cell-surface receptor. Unfortunately, the biochemical purification of any one type of monoclonal antibody from blood is

not feasible for two main reasons: the concentration of any given antibody is quite low, and all antibodies have the same basic molecular architecture (see Figure 3-22).

To produce and then purify monoclonal antibodies, one first needs to be able to grow the appropriate B-lymphocyte clone. However, primary cultures of normal B lymphocytes are of limited usefulness for the production of monoclonal antibodies because they have a limited life span. Thus the first step in producing a monoclonal antibody is to generate a library of immortal antibody-producing cells (Figure 4-6). Immortality is achieved by fusing normal B lymphocytes from an immunized animal with transformed, immortal lymphocytes called *myeloma cells* that themselves do not synthesize antibodies, and selecting for hybrid cells called a **hybridoma**. Like myeloma cells, hybridoma cells grow rapidly and are immortal. Each hybridoma produces the monoclonal antibody encoded by its B-lymphocyte parent. The library of hybridomas is then screened for clones that produce the desired antibody; any clone producing that antibody is then grown in large cultures, from which a substantial quantity of pure monoclonal antibody can be obtained.



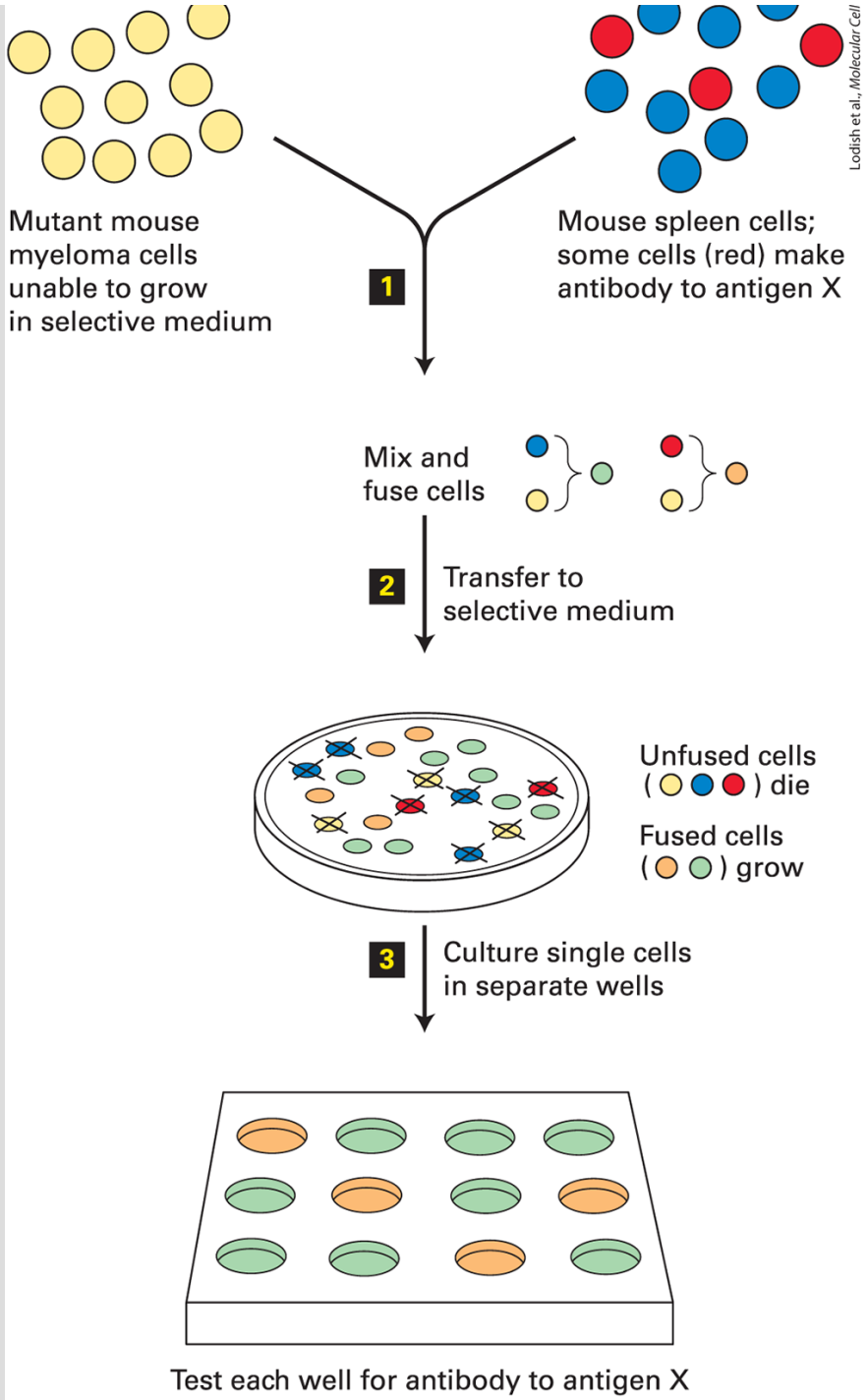


FIGURE 4-6 Use of cell fusion and selection to obtain hybridomas producing a monoclonal antibody to a specific protein. Step 1: Immortal myeloma cells that cannot synthesize purines under special conditions because they lack thymidine kinase are fused with normal antibody-producing spleen cells from an animal that was immunized with antigen X. Step 2: When cultured in a

special selective medium, unfused and self-fused cells do not grow: the myeloma cells do not grow because the selective medium does not contain purines, and the spleen cells do not grow because they have a limited life span in culture. Thus only fused cells formed from a myeloma cell *and* a spleen cell survive in the selective medium, proliferating into clones called *hybridomas*. Each hybridoma produces a single antibody. Step **3**: Testing of individual clones identifies those that recognize antigen X. After a hybridoma that produces a desired antibody has been identified, the clone can be cultured to yield large amounts of that antibody.

Monoclonal antibodies are valuable research tools. They are commonly employed in affinity chromatography to isolate and purify proteins from complex mixtures (see Figure 3-43c). As we discuss later in this chapter, they are also employed in immunofluorescence microscopy to locate a particular protein within cells, and they are used in immunoblotting to identify specific proteins in cell fractions (see Figure 3-44).

Rabbit, mouse, and human immunoglobulins are built from two heavy chains and two light chains, with both the light and heavy chains contributing to the antigen binding site (see Figure 24-14). Remarkably, camelids (including camels and llamas) have these conventional immunoglobulins and immunoglobulins consisting of just two heavy chains. The antigen-combining site of these antibodies is relatively small, about 15 kDa compared with about 150 kDa for a whole mouse monoclonal antibody. Thus the camelid immunoglobulins are easily expressed in, and purified from, bacteria. Such so-called *nanobodies* with a single high-affinity

binding site for specific antigens are being developed for therapeutic, diagnostic, and research purposes.

Monoclonal antibodies have become important diagnostic and therapeutic tools in medicine as well. For example, monoclonal antibodies that bind to and inactivate toxins secreted by bacterial pathogens are used to treat certain infectious diseases.

The first monoclonal antibodies used for therapeutic purposes were made in mice. However, when used in humans these were recognized as foreign proteins and removed by the immune system. Subsequently, antibodies were engineered to contain mostly human sequences fused to the mouse antigen-binding regions. An early example of such a therapeutic antibody was the development of Herceptin for treatment of an aggressive form of breast cancer. It was known that these cancers overexpress HER2, a plasma membrane protein of the EGF receptor tyrosine kinase family (see Chapter 16). Monoclonal antibodies to the extracellular domain of HER2 were generated and screened for those that reduced HER2-overexpressing tumors in mice. Refining and humanizing this mouse monoclonal antibody resulted in the development of Herceptin, which has now been used to treat about a half million patients with HER2-positive breast cancer.

Therapeutic antibodies are available today with about 85–90 percent human sequences, although generating these is a technologically

complicated process. A much better solution would be to generate monoclonal antibodies directly from human B cells. Recently, such an approach has been made possible.

To make human monoclonal antibodies, many millions of naive B cells are isolated from removed tonsils. Since the isolated cells cannot be propagated indefinitely, the population is immortalized by infecting them with Epstein-Barr virus. The immortal naive B cells are induced to undergo class-switching and somatic mutations (discussed in Chapter 24) to generate an extensive library of B cells in which each clone secretes a specific monoclonal antibody. The library of B cells can then be screened with antigen for those that secrete the desired monoclonal antibody. In this way, it will be possible to isolate human-based therapeutic monoclonal antibodies.

A Wide Variety of Cell Biological Processes Can Be Studied with Cultured Cells

As we discussed in the introduction to this chapter, studying animal cells in culture is much easier than studying cells in intact animals, partly because they can easily be subjected to a variety of manipulations. Cultured cells are particularly useful for the elucidation of fundamental processes. One way to understand a

biological process is to interfere with a specific constituent in the cell and assess the outcome — this approach is like trying to understand how a car works by removing components and seeing what goes wrong. In some cases, human diseases associated with genetic defects in specific cell components can be analyzed using cells cultured from the patients. For example, analysis of cultured cells from patients with a genetic defect resulting in hypercholesterolemia — who have elevated blood cholesterol leading to heart disease and stroke — was critical in elucidating the basic steps of receptor-mediated endocytosis (see Chapter 14). In addition to relying on naturally occurring genetic lesions, we can manipulate cultured cells to interfere with expression of specific components. As we will see in Chapter 6, it is possible to decrease the expression of a specific protein in cultured cells by selectively knocking down the corresponding mRNA and then assess what effect this change has on particular processes in the cell. Chapter 6 also describes the more recently developed gene editing techniques that can be used to inactivate specific genes in the genomes of cultured cells and thereby assess how the complete loss of specific RNAs and proteins affects cell functions.

Drugs Are Commonly Used in Cell Biological Research

Another powerful way to analyze biological processes is to treat cells with drugs that bind to specific cell components and inactivate or activate them. In this section, we discuss how new drugs affecting specific cell processes can be developed.

Naturally occurring drugs have been used for centuries, but how they worked was often unknown. For example, extracts of meadow saffron were used to treat gout, a painful disease resulting from inflammation of joints. Today we know that this plant contains colchicine, a drug that depolymerizes microtubules (see Chapter 18) and interferes with the ability of white blood cells to move to sites of inflammation. Alexander Fleming discovered that certain fungi secrete compounds that kill bacteria (antibiotics), and his discovery resulted in the development of penicillin. Only later was it discovered that penicillin inhibits cell division by blocking the assembly of the cell walls of certain bacteria.

Discoveries like these have resulted in a wide range of drugs that can inhibit specific and essential processes of cells. In most cases, researchers have eventually been able to identify the molecular targets of these drugs. For example, there are many antibiotic drugs that affect aspects of prokaryotic protein synthesis.

How does one discover a new drug? One widely used approach is to search *chemical libraries*, consisting of tens to hundreds of thousands of different compounds, for chemicals that inhibit a specific process. The screening of chemical libraries in conjunction with

high-throughput microscopic techniques has now become one of the major routes for new leads in drug discovery. Here we give just one case to illustrate how this type of approach works.

In our example (Figure 4-7a), researchers wanted to identify compounds that inhibit mitosis, the process by which duplicated chromosomes are accurately segregated by a microtubule-based machine called the *mitotic spindle* (discussed in Chapter 18). It was known that if spindle assembly is compromised, cells are arrested in mitosis. Therefore, the screen first used an automated robotic method to look for compounds that arrest cells in mitosis. The basis for the inhibition of mitosis by the candidate compounds was then explored to see if they affected assembly of the microtubules. Since inhibition of microtubule assembly was not of interest, the effect of the remaining candidate compounds on the structure of the spindle was determined by immunofluorescence microscopy using antibodies to tubulin, the major protein of microtubules. Over 16,000 compounds were screened, and a compound was identified that resulted in cells with abnormal spindles — instead of having two asters, they had a single aster, resulting in what is called a *mono-astral array* (Figure 4-7b). This drug, now called *monastrol*, was found to interfere with the assembly of the spindle by inhibiting a microtubule-based motor protein called *kinesin-5*. (See Chapter 18 for more details about the mitotic spindle.) While monastrol did not have the appropriate properties for a clinically useful drug, other

kinesin-5 inhibitors have been developed that are now in clinical trials.

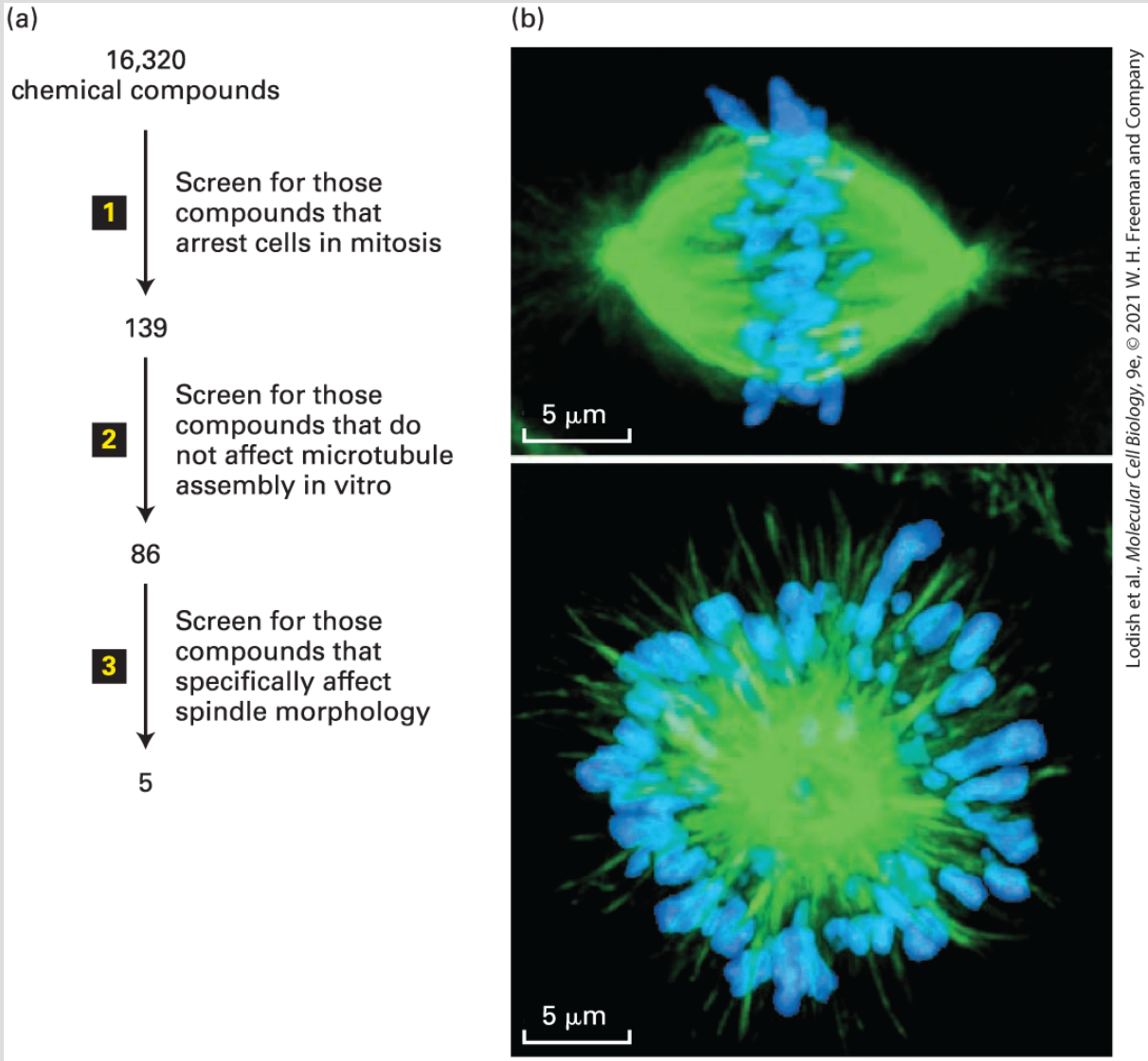


FIGURE 4-7 Screening for drugs that affect specific biological processes. (a) In this example, a chemical library of 16,320 different chemicals was subjected to a series of screens for inhibitors of mitosis. Since such an inhibitor is expected to arrest cells at the mitotic stage of the cell cycle, the first screen (step 1) was to see if any of the chemicals enhanced the level of a marker specific for mitotic cells; this screen yielded 139 candidates. Microtubules make up the structure of the mitotic spindle, and the researchers were not interested in new drugs that target microtubules, so in the second screen (step 2) they

tested the 139 compounds for their ability to affect microtubule assembly; this test eliminated 53 candidates. Immunofluorescence microscopy with antibodies to tubulin (the major subunit of microtubules), together with a stain for DNA, was then used in the third screen (step **3**) to identify compounds that disrupt the structure of the spindle. (b) Localization of tubulin (green) and DNA (blue) for an untreated mitotic spindle (*top*) and one treated with one of the five compounds remaining after screen, now called *monastrol*. Monastrol inhibits a microtubule-based motor protein called *kinesin-5*, discussed in Chapter 18, that is necessary to separate the poles of the mitotic spindle. When kinesin-5 is inhibited, the two poles remain associated to give a monopolar spindle.

[Part (b) Republished with permission of American Association for the Advancement of Science from Mayer, T.U. et al., "Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen," 1999, *Science* **286**:971–974; permission conveyed through Copyright Clearance Center, Inc.]

KEY CONCEPTS OF SECTION 4.1

Growing and Studying Cells in Culture

- Animal cells have to be grown in culture under conditions that mimic their natural environment, which generally requires them to be supplied with necessary amino acids and growth factors.
- Most animal cells need to adhere to a solid surface to grow.
- Primary cells — those isolated directly from tissue — have a finite life span.
- Transformed cells, such as cells derived from tumors, can grow indefinitely in culture.
- Cells that can be grown indefinitely are called a *cell line*.
- Many cells lines are aneuploid, having a different number of chromosomes than the parent cell from which they were derived.
- Cells expressing a fluorescent protein can be sorted on a machine called a *fluorescence-activated cell sorter* (*FACS*) (see Figure 4-2).
- Different cell types express different marker proteins on their cell surfaces, which can be labeled with fluorescent markers, allowing them to be sorted on a FACS machine.
- Epithelial cells are often grown in special containers to mimic their functional polarity. Cells can also be grown on three-dimensional matrices to more accurately reflect their normal environment.

- Monoclonal antibodies, which bind one epitope on an antigen, can be secreted by cultured cells called *hybridomas*. These hybrid cells are made by fusing antibody-producing B lymphocytes with immortal myeloma cells and then identifying those clones that produce the desired antibody. Monoclonal antibodies are important for basic research and as therapeutic agents (see Figure 4.6).
- Cells in culture can be much more easily manipulated than cells in an intact animal.
- Basic biological processes can be studied by interfering with specific cell components, either through genetic mechanisms or by the application of specific drugs.
- Large chemical libraries can be screened for compounds that target specific processes to study those processes and to identify new drugs.

4.2 Light Microscopy: Exploring Cell Structure and Visualizing Proteins Within Cells

The cellular basis of life was first appreciated using primitive light microscopes. Since then, progress in cell biology has paralleled, and has often been driven by, technological advances in light microscopy (Figure 4-8). Here we discuss each of these major developments and how they advanced the study of cellular processes. First we describe basic uses of a light microscope to observe unstained cells and structures, or material treated with chemical stains. Next we describe how fluorescence microscopy is used to localize specific proteins in fixed cells. By using molecular genetic techniques to fuse a protein of interest with a naturally fluorescent protein and express the resulting chimeric protein in cells, it is possible to follow the movement of specific proteins in live cells — an ability that has revealed the dynamic organization of live cells. In parallel with these advances in specimen preparation, optical advances were being made that enhance and sharpen the images provided by fluorescence microscopy, revealing cellular structure with unprecedented resolution. Many specialized technologies have emerged from these advances, and we describe some of the more important ones.

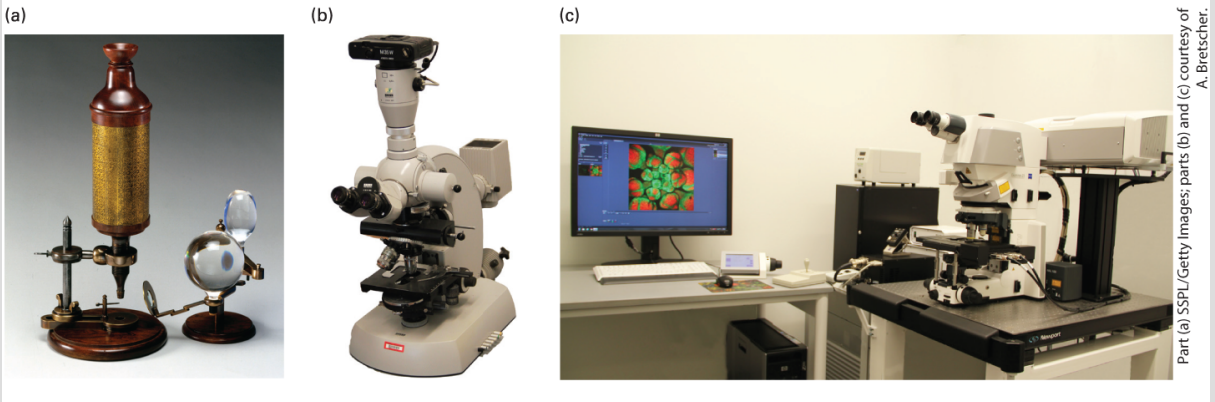


FIGURE 4-8 Development of the light microscope. (a) Early microscopes, like the ones used by Robert Hooke in the 1660s, used lenses or a mirror to illuminate the specimen. (b) Optics in general, and light microscopes in particular, developed enormously during the nineteenth century. By the middle of the twentieth century, highly sophisticated microscopes limited only by the resolution of light were common. (c) In the second half of the twentieth century, fluorescence microscopy and digital imaging, together with confocal techniques, were developed to yield the versatile microscopes of today.

Many of the microscopy techniques we discuss allow one to examine live cells. These include *live cell imaging* as well as examining responses of live cells to a specific stimulus, such as a growth factor, or their interactions with other cells. These techniques have provided scientists with the ability to probe the functioning of individual components in live cells.

The Resolution of the Conventional Light Microscope Is About 0.2 μm

All microscopes produce a magnified image of a small object, but the nature of the image depends on the type of microscope

employed and on the way the specimen is prepared. The compound microscope, used in conventional *bright-field light microscopy*, contains several lenses that magnify the image of a specimen under study (Figure 4-9a, b). The total magnification is a product of the magnification of the individual lenses: if the *objective lens*, the lens closest to the specimen, magnifies hundredfold (a 100 × lens, the maximum usually employed), and the *projection lens* that focuses the image on a camera magnifies tenfold, the final magnification will be 1000-fold. Alternatively, if the light is directed to an ocular or *eyepiece* lens that magnifies tenfold, the final magnification recorded by the human eye will also be 1000-fold.

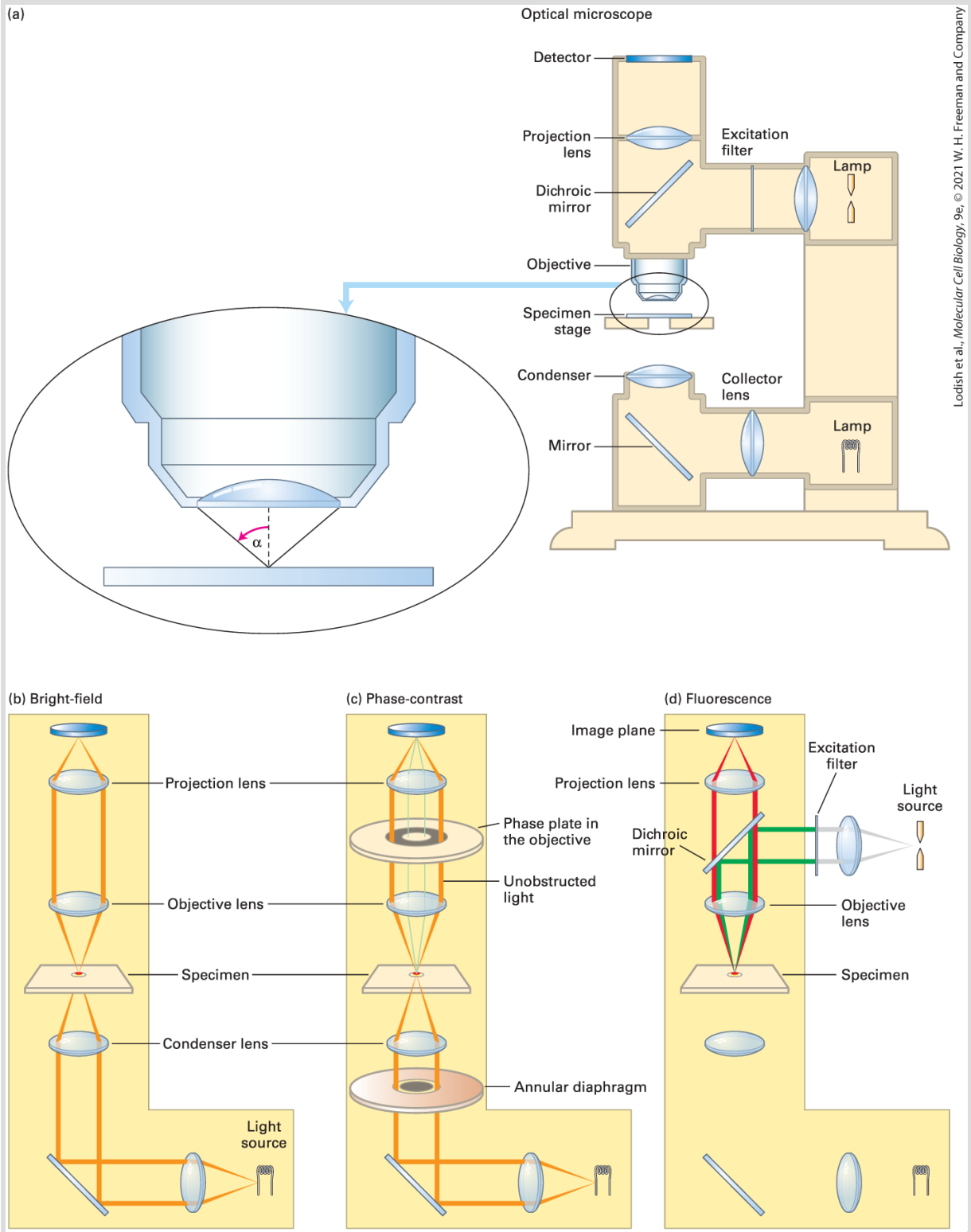


FIGURE 4-9 Optical microscopes are commonly configured for bright-field, phase-contrast, or fluorescence microscopy. (a) In a typical light microscope, the specimen is

usually mounted on a transparent glass slide and positioned on the movable specimen stage. (b) In bright-field light microscopy, light from a tungsten lamp is focused on the specimen by a condenser lens below the stage; the light travels the pathway shown in yellow. (c) In phase-contrast microscopy, incident light passes through an annular diaphragm, which focuses a circular annulus (ring) of light on the specimen. Light that passes unobstructed through the specimen is focused by the objective lens onto the thicker gray ring of the phase plate, which absorbs some of the direct light and alters its phase by one-quarter of a wavelength. If a specimen refracts (bends) or diffracts the light, the phase of some light waves is altered (green lines), and the light waves pass through the clear region of the phase plate. The refracted and unrefracted light is recombined at the image plane to form the image. (d) In fluorescence microscopy, a beam of light from a mercury lamp (gray lines) is directed to the excitation filter, which allows only the correct wavelength of light to pass through (green lines). The light is then reflected off a dichroic mirror and through the objective lens, which focuses it on the specimen. The fluorescent light emitted by the specimen (red lines) passes up through the objective lens, then through the dichroic mirror, and is focused and recorded on the detector at the image plane.

The most important property of any microscope, however, is not its magnification, but its resolving power, or **resolution**: the ability to distinguish between two very closely positioned objects. Merely enlarging the image of a specimen accomplishes nothing if the image is blurred. The resolution of a microscope lens is numerically equivalent to D, the minimum distance between two distinguishable objects. The smaller the value of D, the better the resolution. The value of D is given by the equation

$$D = \frac{0.61\lambda}{N \sin \alpha} \tag{4-1}$$

where α is the angular aperture, or half-angle, of the cone of light entering the objective lens from the specimen (Figure 4-9a), N is the refractive index of the medium between the specimen and the objective lens (i.e., the relative velocity of light in the medium compared with the velocity in air), and λ is the wavelength of the incident light. Resolution is improved by using shorter wavelengths of light (decreasing the value of λ) or by gathering more light (increasing either N or α). Lenses for high-resolution microscopy are designed to work with oil between the lens and the specimen under a thin glass coverslip, since oil has a higher refractive index (1.56, compared with 1.0 for air and 1.3 for water), and matching the refractive index more closely to that of glass (about 1.52) maximizes the recovered light. To maximize the angle α, and hence $\sin \alpha$, the lenses are also designed to focus very close to the thin coverslip covering the specimen. The term $N \sin \alpha$ is known as the *numerical aperture* (*NA*) and is usually marked on the objective lens. A good high-magnification lens has an NA of about 1.4, and the very best lenses — which cost as much as a medium-sized car — have a value approaching 1.5. Notice that the magnification is not part of this equation.

Owing to limitations in the values of α, λ, and N based on the physical properties of light, the *limit of resolution* of a microscope using light of wavelength 400 nm is about 0.2 μm (200 nm). No matter how many times the image is magnified, a conventional light microscope can never resolve objects that are closer than about

0.2 μm apart or reveal details smaller than about 0.2 μm in size. However, some newer and sophisticated technologies devised to overcome this resolution barrier can resolve objects that are just a few 10s of nanometers apart. We discuss such super-resolution microscopes in a later section.

Phase-Contrast and Differential-Interference-Contrast Microscopy Visualize Unstained Live Cells

Cells are about 70 percent water, 15 percent protein, and 6 percent RNA, and contain smaller amounts of lipids, DNA, and small molecules. Since none of these major classes of molecules are colored, and since they hardly impede the transmission of light, special methods must be used to see cells in a microscope. For example, the simplest microscopes view cells under *bright-field* optics (Figure 4-9b), and little detail can be seen (Figure 4-10, *left*). Two common methods for imaging live cells and unstained tissues to generate contrast take advantage of differences in the refractive index and thickness of cellular materials. These methods, called *phase-contrast microscopy* and *differential-interference-contrast (DIC) microscopy* (or *Nomarski interference microscopy*), produce images that differ in appearance and reveal different features of cell architecture. Figure 4-10 compares images of live, cultured cells obtained with these two methods and with standard bright-field

microscopy. Since optical microscopes are expensive, they are often set up to perform many different types of microscopy on the same microscope stand (see Figure 4-9a–d).

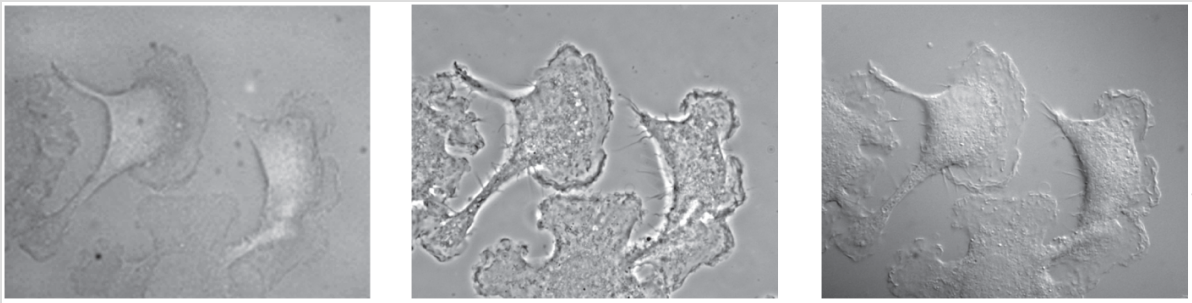


FIGURE 4-10 Live cells can be visualized by microscopy techniques that generate contrast by interference. These micrographs show live, cultured macrophage cells viewed by bright-field microscopy (*left*), phase-contrast microscopy (*middle*), and differential-interference-contrast (DIC) microscopy (*right*). In a phase-contrast image, cells are surrounded by alternating dark and light bands; in-focus and out-of-focus details are simultaneously imaged in a phase-contrast microscope. In a DIC image, cells appear in pseudorelief. Because only a narrow in-focus region is imaged, a DIC image is an optical slice through the object.

Phase-contrast microscopy generates an image in which the degree of darkness or brightness of a region of a specimen depends on the *refractive index* of that region. Light moves more slowly in a medium with a higher refractive index. Thus a beam of light is refracted (bent) once as it passes from the medium into a transparent object and again when it departs. In a phase-contrast microscope, a cone of light generated by an annular diaphragm in the condenser lens illuminates the specimen (see Figure 4-9c). The light passes through the specimen into the objective lens, and the unobstructed direct light passes through a region of the phase plate that both transmits

only a small percentage of the light and changes its phase slightly. The part of a light wave that passes through a specimen will be refracted and will be out of phase (out of synchrony) with the part of the wave that does not pass through the specimen. How much their phases differ depends on the difference in refractive index along the two paths and on the thickness of the specimen. The refracted and unrefracted light is recombined at the image plane to form the image. If the two parts of the light wave are recombined, the resultant light will be brighter if they are in phase and less bright if they are out of phase (Figure 4-10, *middle*). Phase-contrast microscopy is suitable for observing single cells or thin cell layers, but not thick tissues. It is particularly useful for examining the location and movement of larger organelles in live cells.

DIC microscopy, which is based on splitting the light into two perpendicular components before passing them through the specimen and then recombining them to observe their interference pattern, is the method of choice for visualizing extremely small details and thick objects. Contrast is generated by differences in the refractive index of the object and of its surrounding medium. In DIC images, objects appear to cast a shadow to one side. The shadow primarily represents a difference in the refractive index of a specimen rather than its topography. DIC microscopy easily defines the outlines of large organelles, such as the nucleus and vacuole. In addition to having a three-dimensional appearance, a DIC image is a thin *optical section*, or slice, through the object (Figure 4-10, *right*).

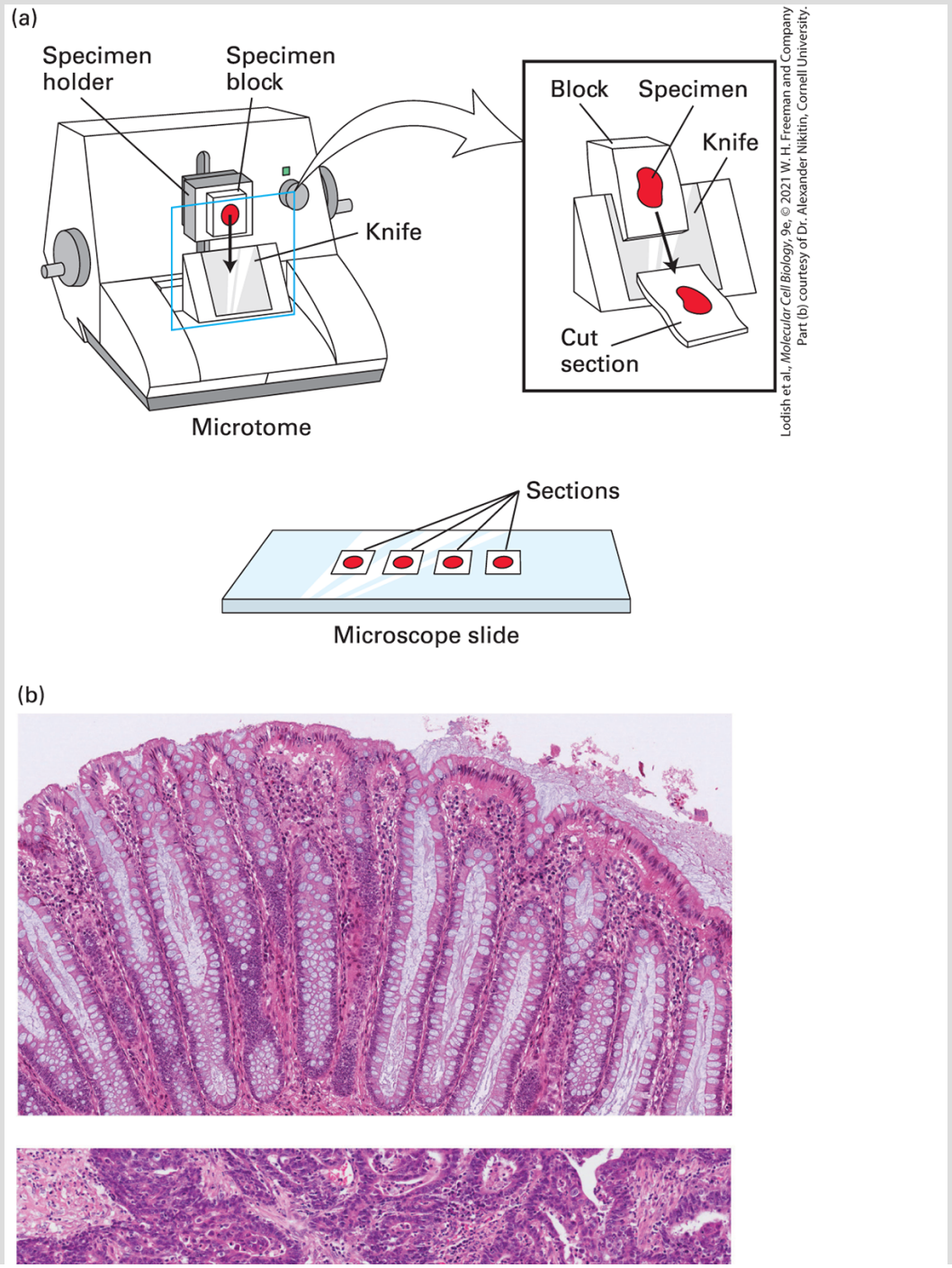
Thus details of the nucleus in thick specimens (e.g., an intact *Caenorhabditis elegans* roundworm; see Figure 22-26d) can be observed in a series of such optical sections, and the three-dimensional structure of the object can be reconstructed by combining the individual DIC images. Both phase-contrast and DIC microscopy can be used in *live cell microscopy*, in which the same cell is imaged at regular intervals over time to generate a movie that allows the observer to study cell movement.

Imaging Subcellular Details Often Requires That Specimens Be Fixed, Sectioned, and Stained

As we have seen, live cells and tissues generally do not absorb light, so they are nearly invisible in a light microscope. Although cells can be visualized by the special techniques we have just discussed, these methods do not reveal the fine details of structure.

Specimens for light microscopy are commonly fixed with a solution containing chemicals that cross-link most proteins and nucleic acids. Formaldehyde, a common fixative, cross-links amino groups on adjacent molecules; these covalent bonds stabilize protein-protein and protein–nucleic acid interactions and render the molecules insoluble and stable for subsequent procedures. After fixation, a tissue sample for examination by light microscopy is

usually embedded in paraffin and cut into sections about 5-μm thick (Figure 4-11a). Cultured cells growing on glass coverslips are thin enough to be fixed in situ and visualized by light microscopy without the need for sectioning.



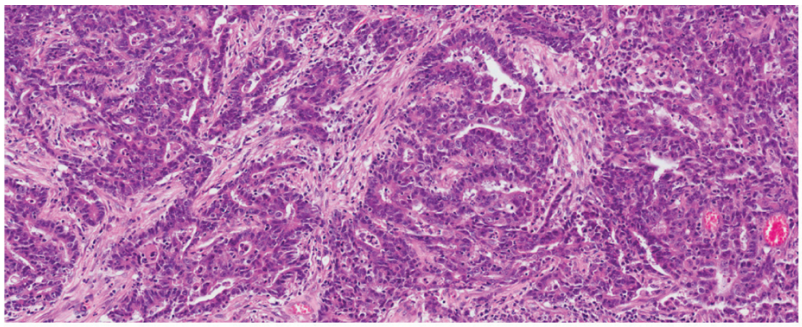


FIGURE 4-11 Tissues for light microscopy are commonly fixed, embedded in a solid medium, and cut into thin sections. (a) A fixed tissue is dehydrated by soaking in a series of alcohol-water solutions, ending with an organic solvent compatible with the embedding medium. To embed the tissue for sectioning, the tissue is placed in liquid paraffin. After the block containing the specimen has hardened, it is mounted on the arm of a microtome, and slices are cut with a knife. Typical sections cut for light microscopy are $5\text{-}\mu\text{m}$ thick. The sections are collected on microscope slides and stained with an appropriate agent. (b) Sections of normal (*top*) and cancerous (adenocarcinoma, *bottom*) human colon stained with H&E stain. Notice the disorganization of the cells in the cancer tissue.

Finally, a specimen for light microscopy is stained, allowing the main structural features of the cell or tissue to be clearly seen. Many chemical stains bind to molecules that have specific features. For example, histological samples are often stained with *hematoxylin* and *eosin* ("H&E stain"). Hematoxylin binds to basic amino acids (lysine and arginine) on many different kinds of proteins, whereas eosin binds to acidic molecules (e.g., DNA and side chains of the amino acids aspartate and glutamate). Because of their different binding properties, these dyes stain various cell types sufficiently differently that they are distinguishable visually (Figure 4-11b). If an enzyme catalyzes a reaction that produces a colored or otherwise

visible precipitate from a colorless precursor, that enzyme can be detected in cell sections by their colored reaction products. Such staining techniques, although once quite common, have been largely replaced by other techniques for visualizing particular proteins, as we discuss next.

Fluorescence Microscopy Can Localize and Quantify Specific Molecules in Live Cells

Perhaps the most versatile and powerful technique for localizing molecules within a cell by light microscopy is **fluorescent staining** of cells and observation by *fluorescence microscopy*. A chemical is said to be *fluorescent* if it absorbs light at one wavelength (the excitation wavelength) and emits light (fluoresces) at a specific longer wavelength. Fluorescence microscopes allow the excitation light to pass through the objective lens into the sample and then permit the user to selectively observe the emitted fluorescent light coming back through the objective lens from the sample. This is achieved by using a special type of filter, called a *dichroic mirror*, that reflects the excitation light into the sample and allows the longer wavelength light emitted by the fluorescent sample to pass through to the observer (see Figure 4-9d). Here we discuss several ways in which fluorescence microscopy can be used to examine specific molecules in cells.

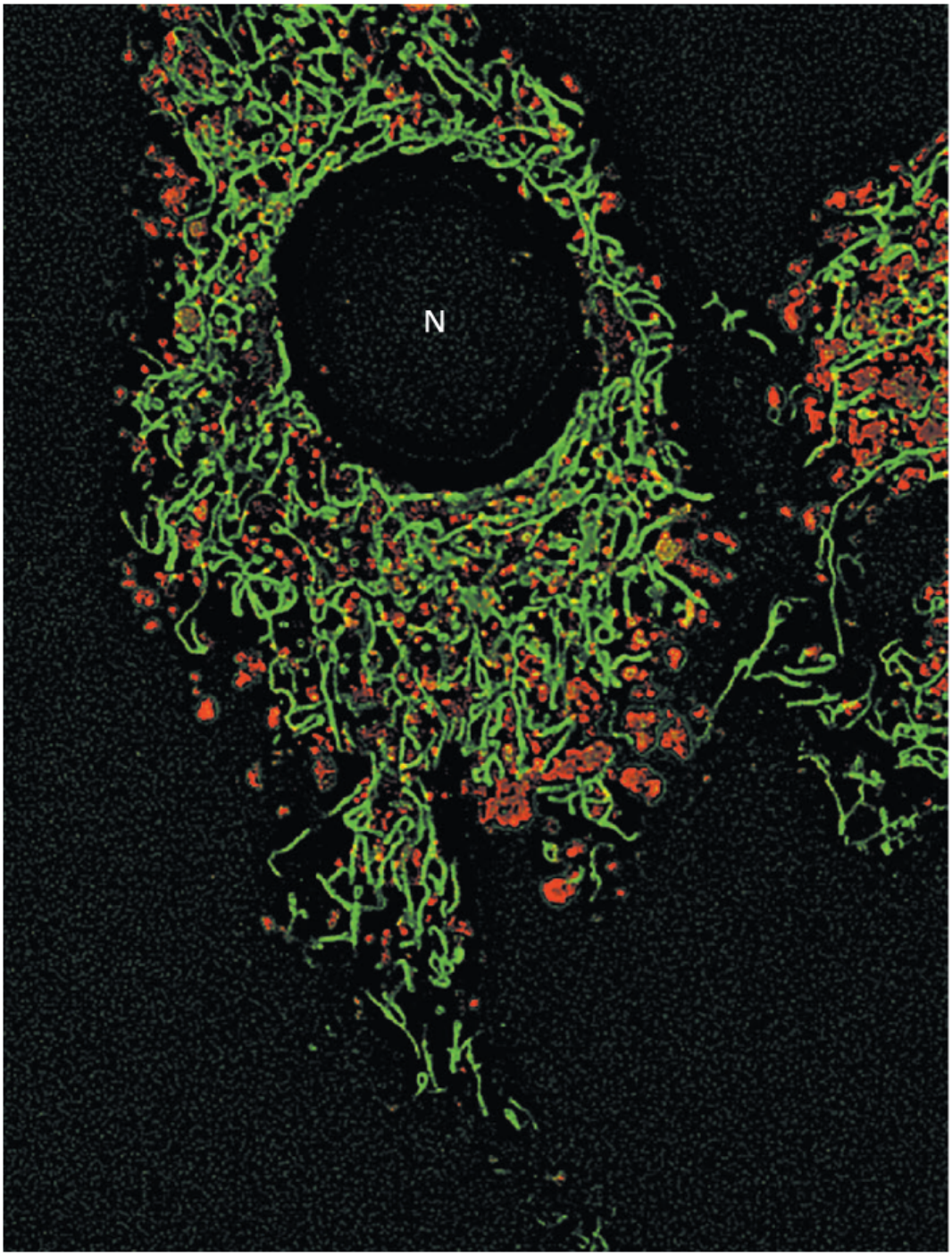
Intracellular Ion Concentrations Can Be Determined with Ion-Sensitive Fluorescent Dyes

The concentration of Ca^{2+} or H^{+} within live cells can be measured with the aid of fluorescent dyes, or *fluorochromes*, whose fluorescence depends on the concentration of these ions. As discussed in later chapters, intracellular Ca^{2+} and H^{+} concentrations have pronounced effects on many cellular processes. For instance, many hormones and other stimuli cause a rise in cytosolic Ca^{2+}, from the resting level at less than 10^{-7} M up to 10^{-5} or 10^{-4} M, that induces various cellular responses, such as the contraction of muscle.

Introduction of the fluorochrome *Fura-2*, which binds a single Ca^{2+}, can be used to follow transient increases of Ca^{2+} in living cells. This binding, which is proportional to the cytosolic Ca^{2+} concentration over a certain range, increases the fluorescence of Fura-2 at one particular wavelength. At a second wavelength, the fluorescence of Fura-2 is the same whether or not Ca^{2+} is bound and thus provides a measure of the total amount of Fura-2 in a region of the cell. By examining cells continuously in the fluorescence microscope and measuring rapid changes in the ratio of Fura-2 fluorescence at the two wavelengths, one can quantify rapid changes in the fraction of Fura-2 that has bound a Ca^{2+} ion, and thus in the concentration of

cytosolic Ca^{2+}. Although Fura-2 has been used extensively to study Ca^{2+} levels, as we discuss later in this chapter, heterologous expression of a protein Ca^{2+} sensor, which becomes fluorescent upon binding Ca^{2+}, has largely replaced the use of Ca^{2+}-sensitive dyes.

Fluorescent dyes that are sensitive to H^+ concentrations can be used to monitor the cytosolic pH of live cells. Other useful probes consist of a fluorochrome linked to a weak base that is only partially protonated at neutral pH and thus can freely permeate cellular membranes. In acidic organelles, however, these probes become protonated. Because the protonated probes cannot recross the organelle membrane, they accumulate in the lumen at concentrations much greater than in the cytosol. Thus this type of fluorescent dye can be used to specifically stain particular organelles in live cells, such as lysosomes and mitochondria shown in Figure 4-12.



© 2020 Thermo Fisher Scientific, Inc. Used under permission.

EXPERIMENTAL FIGURE 4-12 Location of lysosomes and mitochondria in a cultured living bovine pulmonary artery endothelial cell. The cell was stained with a green-fluorescing dye that is specifically bound to mitochondria and a red-fluorescing dye that is specifically incorporated into lysosomes. The image was

sharpened using a deconvolution computer program discussed later in the chapter. N, nucleus.

Immunofluorescence Microscopy Can Detect Specific Proteins in Fixed Cells

The common chemical dyes mentioned above stain nucleic acids or broad classes of proteins, but it is much more informative to detect the presence and location of specific proteins. *Immunofluorescence microscopy*, the most widely used method of detecting specific proteins, uses an antibody to which a fluorescent dye has been covalently attached. To use this method, one must first generate antibodies to the specific protein of interest. As discussed briefly in Section 4.1 and in detail in Chapter 24, when an experimental animal is injected with an antigen, the vertebrate immune system responds by generating polyclonal antibodies that recognize different epitopes on the antigen. The antibodies that specifically recognize the antigen can be purified from all the other antibodies and proteins in blood serum by affinity chromatography over a resin on which the antigen is immobilized (see Figure 3-43c). Alternatively, as we described earlier in this chapter, it is possible to generate a clonal cell line that secretes antibodies to a specific epitope on the antigen; these are called *monoclonal antibodies*.

To use either type of antibody to localize a cytosolic protein, the cells or tissue must first be fixed to ensure that all components remain in place, and the cell must be permeabilized to allow entry of the antibody, which is commonly done by incubating the cells with a non-ionic detergent or by extracting the lipids with an organic solvent. To localize just a surface protein, the cell is fixed, but there is no need to permeabilize it. In one version of immunofluorescence microscopy, the antibody is covalently linked to a fluorochrome. Classically used fluorochromes include rhodamine and Texas red, which emit red light; Cy3, which emits orange light; and fluorescein, which emits green light; but newer and more photostable fluorochromes, with emission wavelengths from blue to far-red, have now been developed. When a fluorochrome-antibody complex is added to a permeabilized cell or tissue section, the complex will bind to the corresponding antigen, then light up when illuminated at the excitation wavelength. Staining a specimen with different dyes that fluoresce at different wavelengths allows multiple proteins as well as DNA to be localized within the same cell.

The most commonly used variation of this technique is called *indirect immunofluorescence microscopy* because the antibody specific to the protein of interest is detected indirectly. In this technique, an unlabeled monoclonal or polyclonal antibody is applied to the specimen, followed by a second, fluorochrome-tagged antibody that binds to the constant (Fc) segment of the first antibody. For example,

a secondary antibody can be generated by immunizing a goat with the Fc segment that is common to all rabbit immunoglobulin G antibodies. When coupled to a fluorochrome, this second antibody preparation (called *goat–anti-rabbit*) will detect any rabbit antibody used to stain a tissue or cell (Figure 4-13). Because several goat–anti-rabbit antibody molecules can bind to a single rabbit antibody molecule in a specimen, the fluorescence is generally much brighter than if a single fluorochrome-tagged antibody were used. This approach is often extended to do *double-label fluorescence microscopy*, in which two proteins can be visualized simultaneously. For example, two proteins can be visualized by indirect immunofluorescence microscopy using primary antibodies made in different animals (e.g., rabbit and chicken) and secondary antibodies (e.g., goat–anti-rabbit and sheep–anti-chicken) labeled with different fluorochromes. In another variation, one protein can be visualized by indirect immunofluorescence microscopy and the second protein by a dye that specifically binds to it. Once the individual images are taken on the fluorescence microscope, they can be merged electronically (Figure 4-14, and see chapter-opening figure).

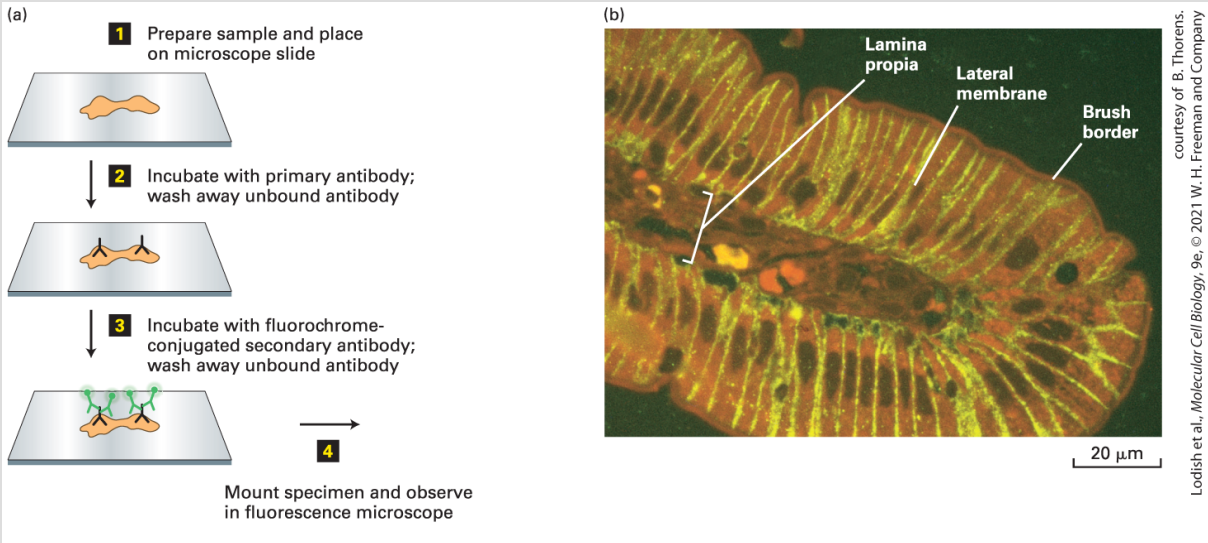
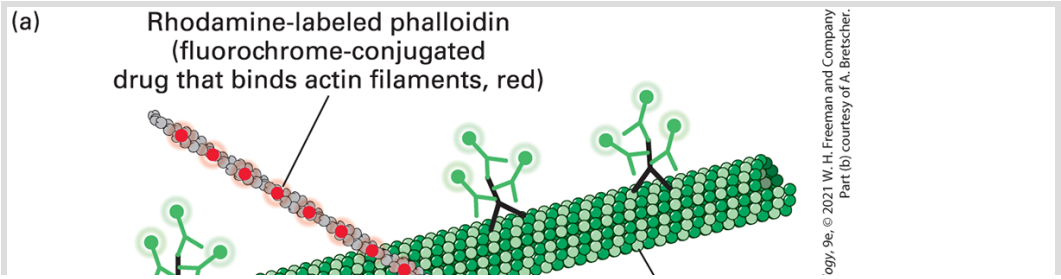
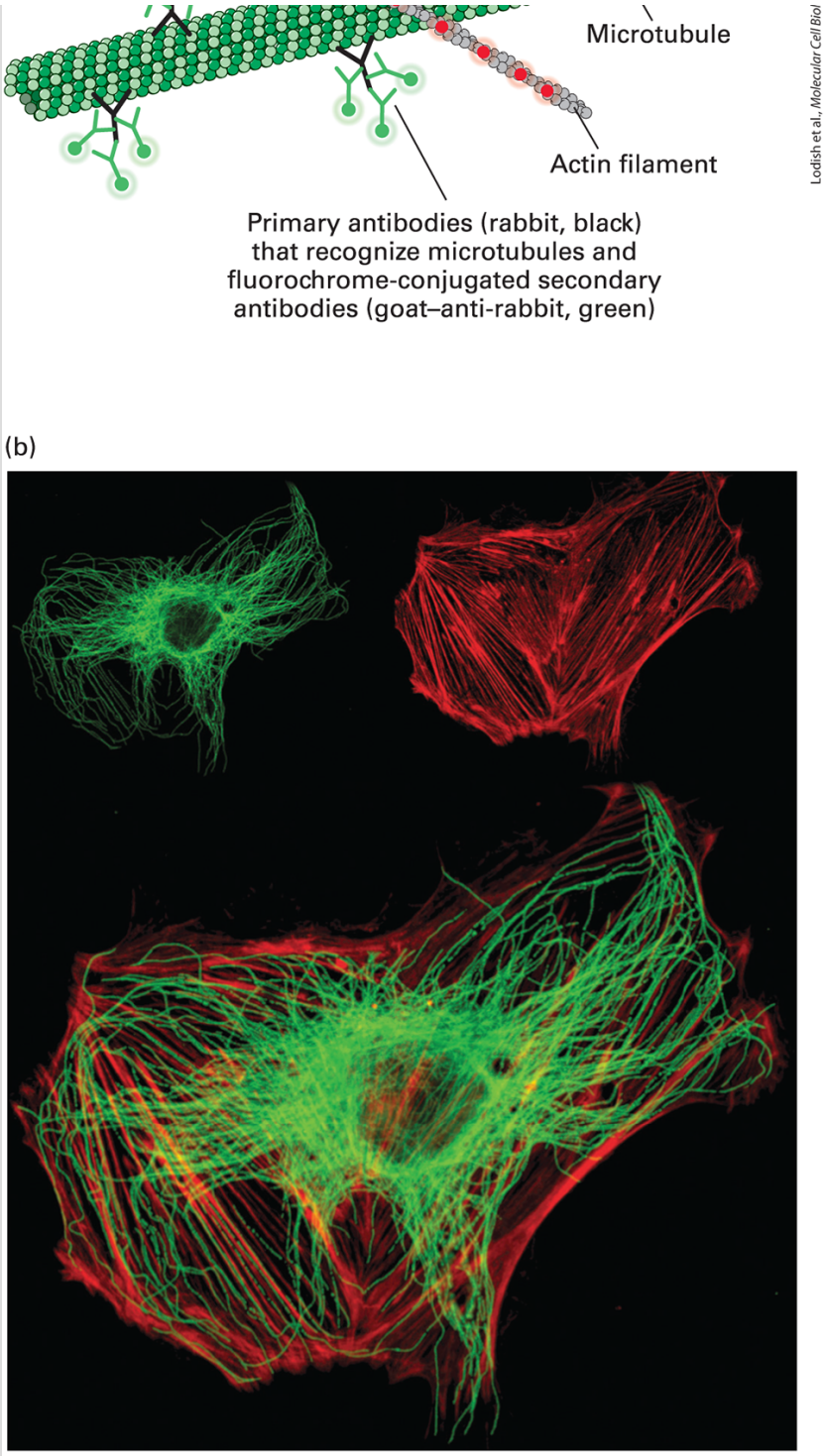


FIGURE 4-13 A specific protein can be localized in fixed tissue sections by indirect immunofluorescence microscopy. (a) To localize a protein by immunofluorescence microscopy, a tissue section, or sample of cells, must be chemically fixed and made permeable to antibodies (step 1). The sample is then incubated with a primary antibody that binds specifically to the antigen of interest, and unbound antibody is then removed by washing (step 2). The sample is next incubated with a fluorochrome-labeled secondary antibody that specifically binds to the primary antibody, and again, excess secondary antibody is removed by washing (step 3). The sample is then mounted in specialized mounting medium and examined in a fluorescence microscope (step 4). (b) In this example, a section of the rat intestinal wall was stained with Evans blue, which generates a nonspecific red fluorescence, and GLUT2, a glucose transport protein, was localized by indirect immunofluorescence microscopy. GLUT2 (yellow) is seen to be present in the basal and lateral sides of the intestinal cells, but is absent from the brush border, composed of closely packed microvilli on the apical surface facing the intestinal lumen. Capillaries run through the lamina propria, a loose connective tissue beneath the epithelial layer.

[Part (b) from B. Thorens et al., 1990, *Am. J. Physiol.* **259**:C279;]





EXPERIMENTAL FIGURE 4-14 Double-label fluorescence microscopy can visualize the relative distributions of two proteins. In double-label fluorescence microscopy, each protein must be labeled with a different fluorochrome. (a) A cultured cell was fixed and permeabilized and then

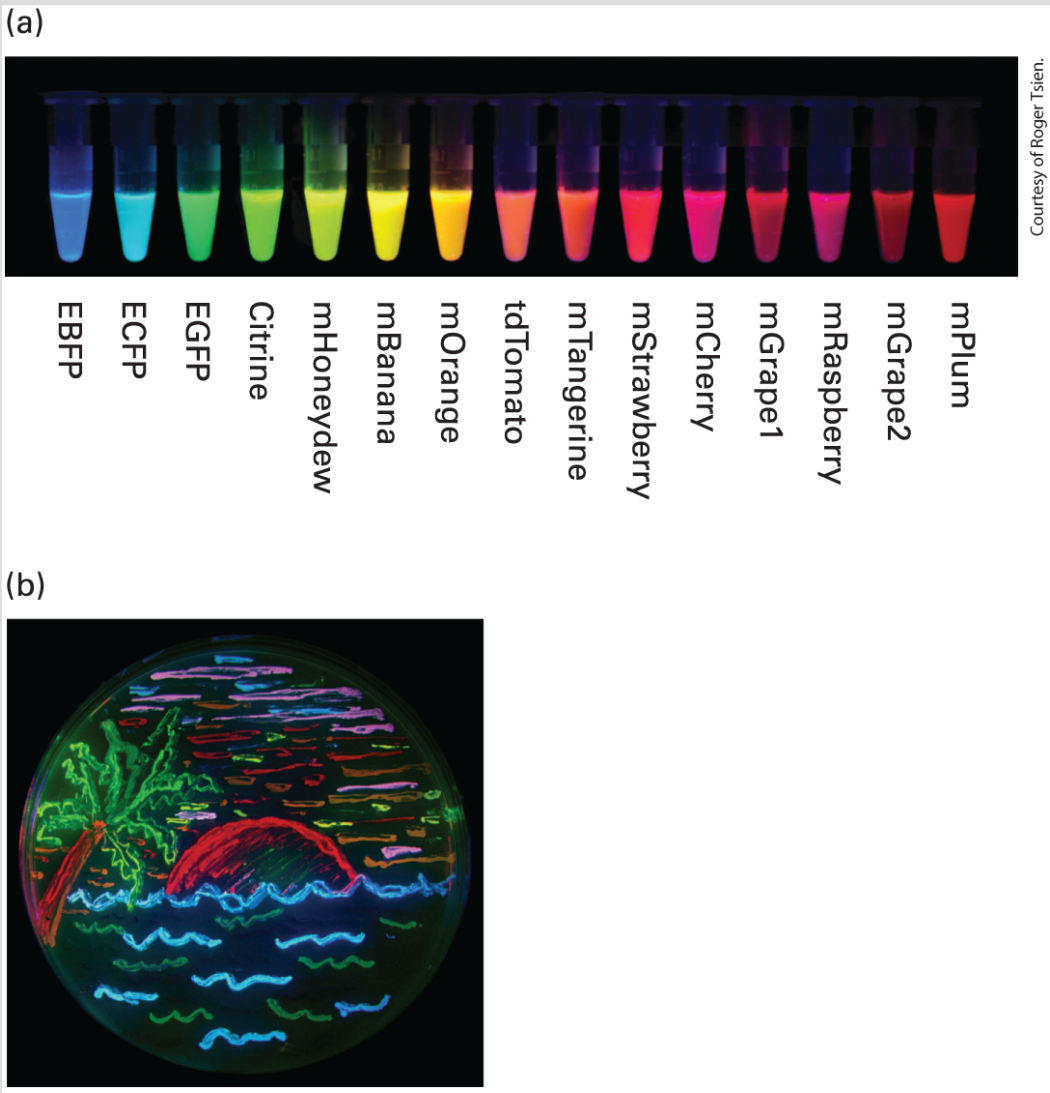
incubated with Rhodamine-labeled phalloidin, a reagent that specifically binds to filamentous actin. It was also incubated with rabbit antibodies to tubulin, the major component of microtubules, followed by a fluorescein-labeled secondary goat–anti-rabbit antibody. (b) The upper panels show the fluorescein-stained tubulin (*left*) and Rhodamine-stained actin (*right*), and the lower panel shows the electronically merged images.

In another widely used version of this technology, molecular genetic techniques are used to make a cDNA encoding a recombinant protein to which is fused a short sequence of amino acids called an *epitope tag*. When expressed in cells, this cDNA will generate the protein linked to the specific tag. Two commonly used epitope tags are FLAG, which encodes the amino acid sequence DYKDDDDK (single-letter amino acid code), and Myc, which encodes the sequence EQKLISEEDL. Commercial fluorochrome-coupled monoclonal antibodies to the FLAG or Myc epitopes can then be used to detect the recombinant protein in the cell. In an extension of this technology to allow the simultaneous visualization of two proteins, one protein can be tagged with FLAG and a different protein with Myc. Each tagged protein is then visualized with a different color, for example, with a labeled antibody emitting green light for the Myc epitope and a labeled antibody emitting red light to the FLAG epitope. One has to use caution in studying expressed tagged proteins to ensure that the tag does not interfere with the protein's function or location.

Tagging with Fluorescent Proteins Allows the Visualization of Specific Proteins in Live Cells

The jellyfish *Aequorea victoria* expresses a naturally fluorescent protein, called *green fluorescent protein* (*GFP*, ~27 kDa). GFP contains a serine, tyrosine, and glycine sequence whose side chains spontaneously cyclize to form a green-fluorescing fluorochrome when illuminated with blue light. Using recombinant DNA technology, it is possible to make a DNA construct in which the coding sequence of GFP is fused to the coding sequence of a protein of interest. When this construct is introduced into and expressed in cells, a GFP-tagged protein is made in which the protein of interest is covalently linked to GFP as part of the same polypeptide. Although GFP is a medium-sized protein, the function of the protein of interest is often not changed by fusing it to GFP. This technique allows one to visualize GFP — and hence the location of the protein of interest. Not only can we see the location of the GFP-tagged protein immediately but we can also view its distribution in a live cell over time and thereby assess its dynamics or track its localization following various treatments, such as the application of specific drugs. The simple idea of tagging specific proteins with GFP revolutionized cell biology and led to the subsequent development of many different fluorescent proteins (Figure 4-15). Use of this colorful variety of fluorescent proteins allows one to visualize two or more proteins simultaneously if they are each tagged with a

different-colored fluorescent protein. We describe additional techniques that exploit fluorescent proteins in the following sections.



Courtesy of Roger Tsien.

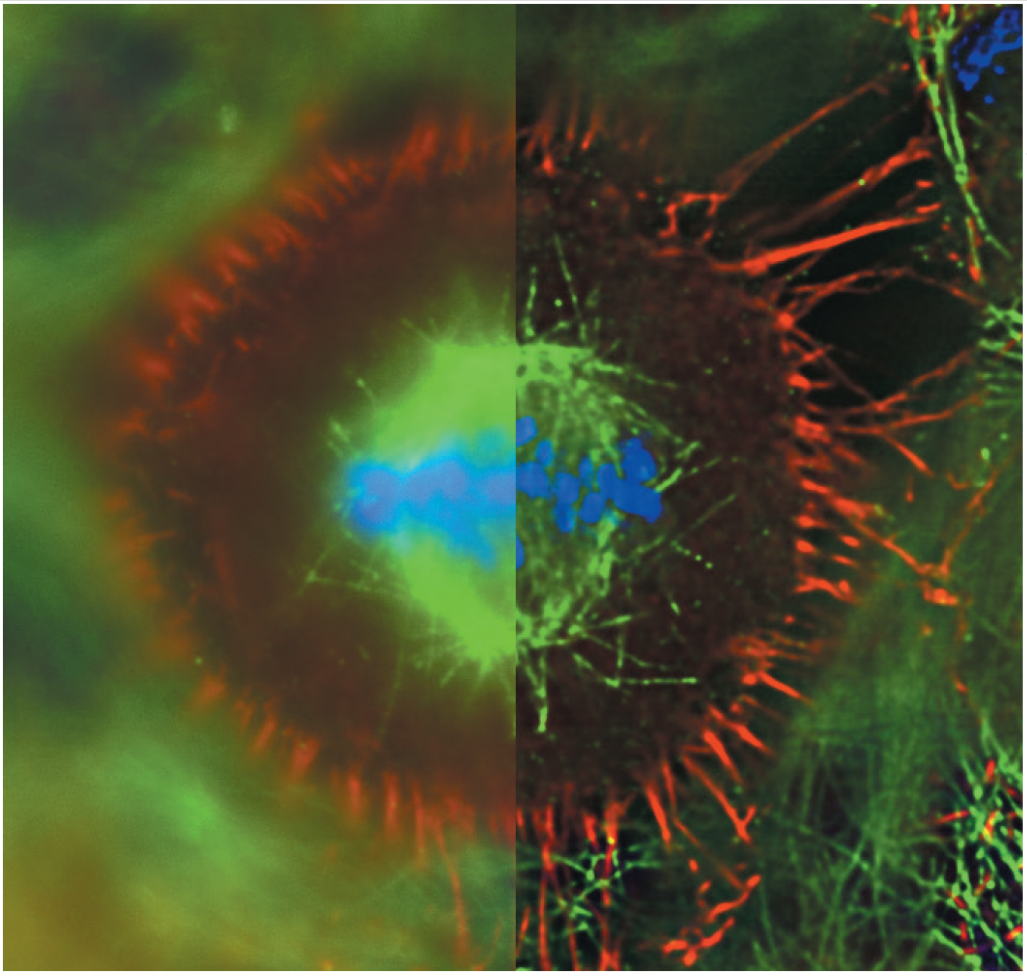
FIGURE 4-15 Many different colors of fluorescent proteins are available. (a) Tubes show the emission colors and names of many different fluorescent proteins. (b) An agar dish is illuminated to show growing bacteria expressing several different-colored fluorescent proteins.

Deconvolution and Confocal Microscopy Enhance Visualization of Three-Dimensional Fluorescent Objects

Conventional fluorescence microscopy has two major limitations. First, the fluorescent light emitted by a sample comes not only from the plane of focus but also from molecules above and below it; thus the observer sees a blurred image caused by the superposition of fluorescent images from molecules at many depths in the cell. The blurring effect makes it difficult to determine the actual spatial arrangements. Second, to visualize thick specimens, consecutive (serial) images at various depths throughout the sample must be collected and then aligned to reconstruct structures in the original thick tissue. Two general approaches have been developed to obtain high-resolution three-dimensional information. Both of these methods require that the image be collected electronically so that it can then be computationally manipulated as necessary.

The first approach, called *deconvolution microscopy*, uses computational methods to remove fluorescence contributed by out-of-focus parts of the sample. Consider a three-dimensional sample in which images from three different focal planes are recorded. Since the whole sample is illuminated, the image from plane 2 will contain out-of-focus fluorescence from planes 1 and 3. If we knew

exactly how out-of-focus fluorescence from planes 1 and 3 contributed to the light collected in plane 2, we could computationally remove it. To obtain this information for a particular microscope, we can make a series of images of focal planes from a test slide containing tiny fluorescent beads. Each bead represents a pinpoint of light that becomes a blurred object outside its focal plane; from these images we can determine a *point spread function* that enables us to calculate the distribution of fluorescent point sources that contributed to the blur when out of focus. Once we have calibrated the microscope in this manner, the experimental series of images can be computationally deconvolved. Microscopes with automated stages to collect the images, and associated software programs to deconvolve the images, are available commercially. Images restored by deconvolution display impressive detail without any blurring, as illustrated in Figure 4-16.



EXPERIMENTAL FIGURE 4-16 Deconvolution fluorescence microscopy yields high-resolution optical sections. A cultured cell in mitosis was stained with fluorochrome-labeled reagents specific for DNA (blue), microtubules (green), and the microfilament-membrane linking protein ezrin (red). A series of fluorescent images were obtained at consecutive focal planes (optical sections) through the cell. A single plane is shown in which the left half is a raw image and the right half has been subjected to deconvolution processing. Notice how much clearer the chromosomes (blue) and microtubules (green) are after processing.

The second approach to obtaining better three-dimensional information is called *confocal microscopy* because it uses optical

methods to obtain images from a specific focal plane and exclude light from other planes. Confocal microscopes collect a series of images focused through the vertical depth of the sample, from which an accurate three-dimensional representation can be computationally generated. Two types of confocal microscopes are in common use today, a *point-scanning* confocal microscope (also known as a *laser-scanning confocal microscope*, or *LSCM*) and a *spinning disk* confocal microscope. The idea behind each microscope is to both illuminate and collect emitted fluorescent light from just one small area of a focal plane at a time in such a way that out-of-focus light is excluded. This can be achieved by collecting the emitted light through a pinhole before it reaches the detector — light from the illuminated focal plane passes through, whereas light from other focal planes is largely excluded. The illuminated area is then moved across the whole focal plane to build up the image electronically. The two types of microscopes differ in how they cover the image. The point-scanning microscope uses a point laser light source at the excitation wavelength to rapidly scan the focal plane in a raster pattern, collects the emitted fluorescence in a photomultiplier tube, and thereby builds up an image (Figure 4-17a). It can then take a series of images at different depths in the sample to generate a three-dimensional reconstruction. A point-scanning confocal microscope can provide exceptionally high-resolution images in both two and three dimensions (Figure 4-18), although it has two minor limitations. First, it can take significant time to scan each focal plane, so if a very dynamic live cell process is being

imaged, the microscope may not be able to collect images fast enough to follow the dynamics. Second, it illuminates each spot with intense laser light, which can bleach the fluorochrome being imaged and damage live cells by phototoxicity, thereby limiting the number of images that can be collected.

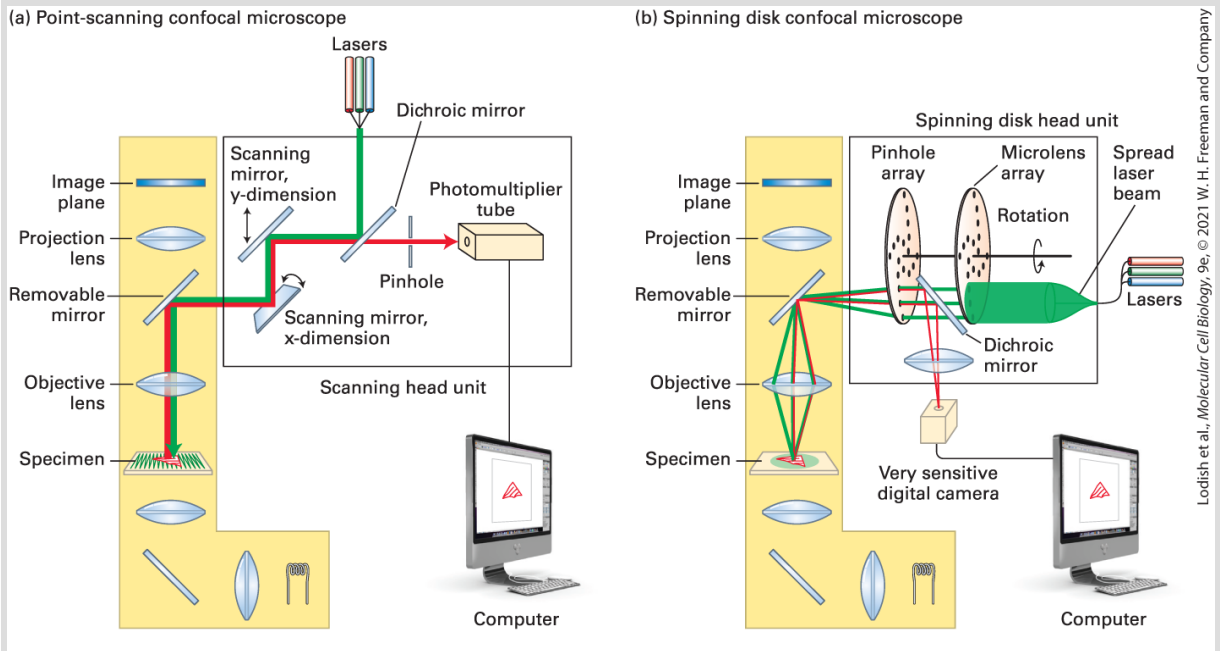


FIGURE 4-17 Light paths for two types of confocal microscopy. Both types of microscopy are assembled around a conventional fluorescence microscope (yellow shading). (a) Light path in a point-scanning confocal microscope. A single-wavelength point of light from an appropriate laser is reflected off a dichroic mirror and bounces off two scanning mirrors and from there passes through the objective lens to illuminate a spot in the specimen. The scanning mirrors rock back and forth in such a way that the light scans the specimen in a raster fashion (see green lines in the specimen). The fluorescence emitted by the specimen passes back through the objective lens and is bounced off the scanning mirrors and through the dichroic mirror. This allows the light to pass toward the pinhole. This pinhole excludes light from out-of-focus focal planes, so the light reaching the photomultiplier tube comes almost exclusively from the illuminated spot in the focal plane. A computer then takes these signals and reconstructs the image. (b) Light path in a spinning disk confocal microscope. Here, instead of using two scanning mirrors, the beam from the laser is spread to illuminate pinholes on the coupled spinning disks, the first consisting of microlenses to focus the light

on pinholes in the second disk. The excitation light passes through the objective lens to provide point illumination of a number of spots in the specimen. The fluorescence emitted passes back through the objective lens and through the holes in the spinning disk, and is then bounced off a dichroic mirror into a sensitive digital camera. The pinholes in the disk are arranged so that as it spins, it rapidly illuminates all parts of the specimen several times. As the disk spins as fast as 5000 rpm, very dynamic events in live cells can be recorded.

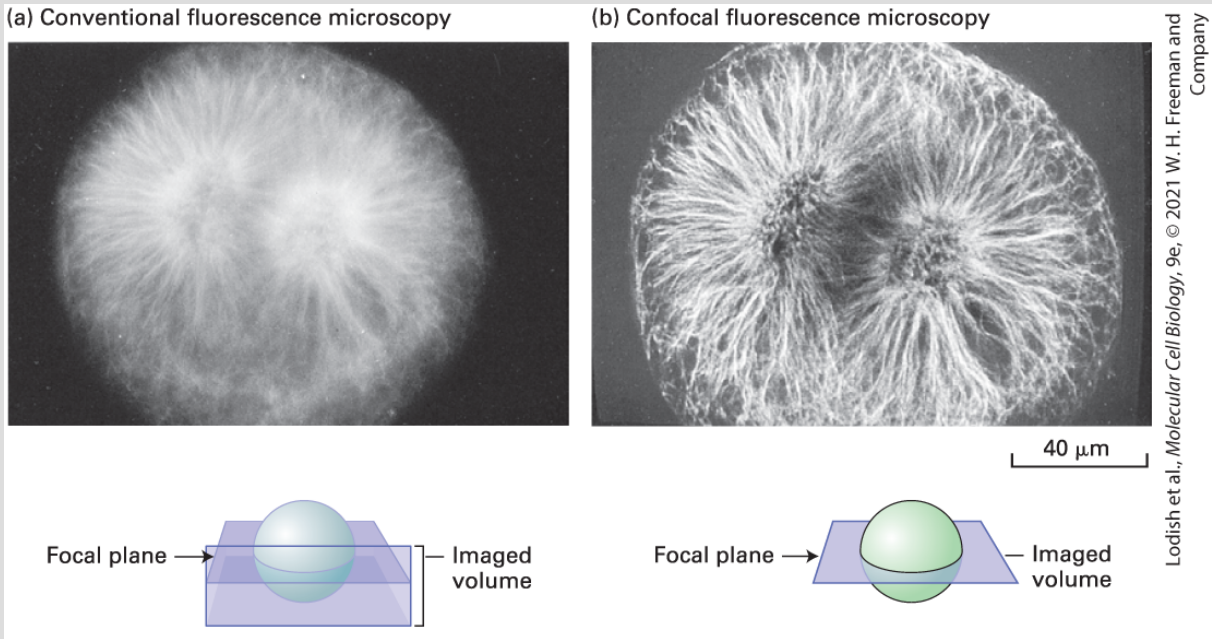
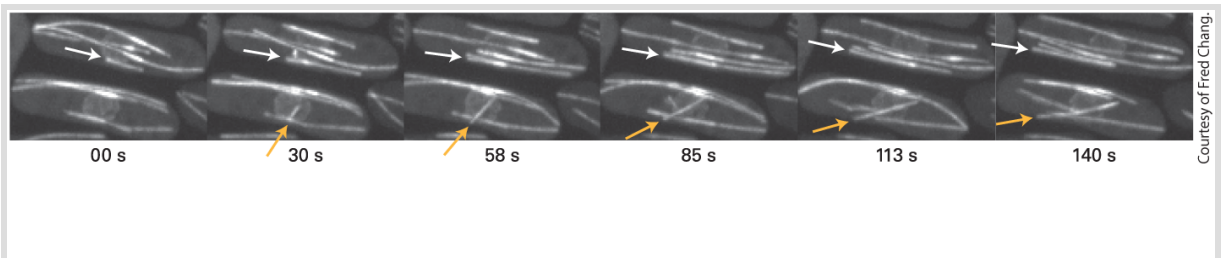


FIGURE 4-18 Confocal microscopy produces an in-focus optical section through thick cells. A mitotic fertilized egg from a sea urchin (*Psammechinus*) was lysed with a detergent, exposed to a tubulin antibody, and then exposed to a fluorescein-tagged antibody that binds to the anti-tubulin antibody. (a) When viewed by conventional fluorescence microscopy, the mitotic spindle is blurred. This blurring occurs because background fluorescence is detected from tubulin above and below the focal plane as depicted in the sketch. (b) The confocal microscopic image is sharp, particularly in the center of the mitotic spindle. In this case, fluorescence is detected only from molecules in the focal plane, generating a very thin optical section.

[Micrographs © 1987 White et al., *J. Cell Biol.* **105**:41–48. https://doi.org/10.1083/jcb.105.1.41]

The spinning disk microscope circumvents these two limitations (Figure 4-17b). The excitation light from a laser is spread out and illuminates a small part of a disk spinning at high speed; for example, at 5000 rpm. The disk in fact consists of two linked disks: one with 20,000 lenses that precisely focuses the laser light on 20,000 pinholes of the second disk. The pinholes are arranged in such a way that they completely scan the focal plane of the sample several times with each turn of the disk. The emitted fluorescent light returns through the pinholes of the second disk and is reflected by a dichroic mirror and focused onto a highly sensitive digital camera. In this way, the sample is scanned in less than a millisecond, so the real-time location of a fluorescent reporter can be captured even if it is highly dynamic (Figure 4-19). A limitation of the spinning disk microscope is that the pinhole size is generally fixed and has to be matched to the magnification of the objective lens, so it is generally configured for use with a 63× or 100× objective and is less useful for the lower magnification imaging that might be required in tissue sections. However, recently, spinning disk head units have become available in which the pinhole size can be changed to better accommodate objectives with other magnifications. Thus the point-scanning and spinning disk confocal microscopes have overlapping and complementary strengths.



EXPERIMENTAL FIGURE 4-19 The dynamics of microtubules can be imaged on the spinning disk confocal microscope. Six frames from a movie of GFP-tubulin in two rod-shaped cells of fission yeast are shown. Arrowheads indicate examples of shrinking and growing microtubules.

Two-Photon Excitation Microscopy Allows Imaging Deep into Tissue Samples

We have just seen how point-scanning confocal microscopy can help reduce fluorescence from out-of-focus planes. To achieve this, it focuses a cone of laser light on a spot that scans across the focal plane. Regions above and below the focal plane are also illuminated by this cone of light, generating an out-of-focus signal that must be removed by collecting the light through a pinhole. This intense cone of light can also lead to photobleaching (rendering the fluorescent protein inactive) or damage to the sample by phototoxicity. If the sample is very thin, these are not significant problems, but as the sample gets thicker, they become more relevant. To circumvent these problems, use was made of the finding that a fluorochrome can be excited either by a single photon — for example, at 488 nm — or by two photons of half the energy at 960 nm, either of which will generate the same emission wavelength (Figure 4-20a). Thus if a 960-nm cone of laser light is focused on a spot in one plane so that only at the focal point is there sufficient density of photons to excite the

fluorochrome (Figure 4-20b), no out-of-focus signal will be obtained and less photobleaching or phototoxicity will occur. Because only fluorochromes in the focal plane are excited, two-photon microscopy can be used to explore much thicker samples, and there is no need for a pinhole to exclude out-of-focus light. Moreover, 960-nm light can penetrate more deeply into biological material than can 488-nm light. However, very high laser intensity is required for two-photon excitation microscopy, as the two photons must arrive within about a femtosecond of each other, which is achieved using rapid laser pulses. If individual cells in an animal express different color variants of fluorescent proteins, two-photon microscopy can be used to observe events in living animals (called *intravital imaging*) within about 1 mm of the surface (Figure 4-20c).

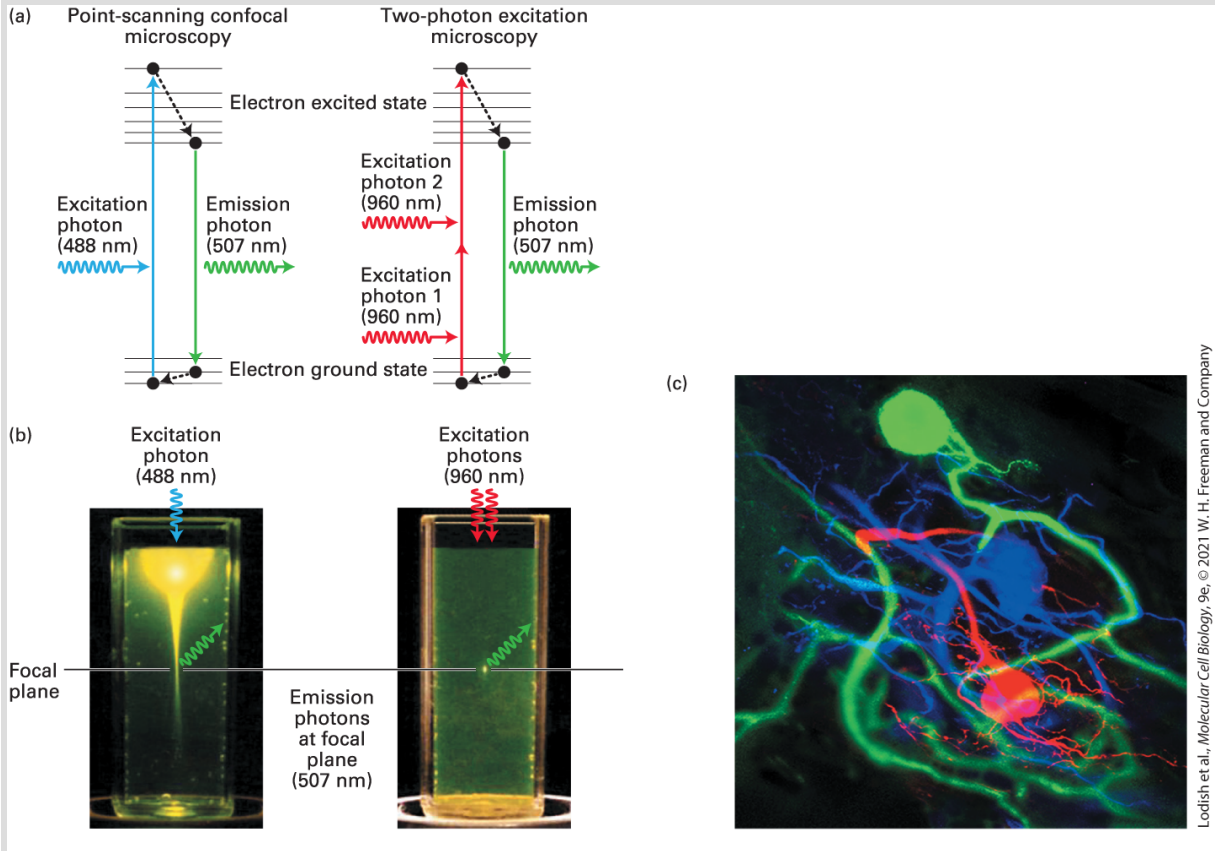


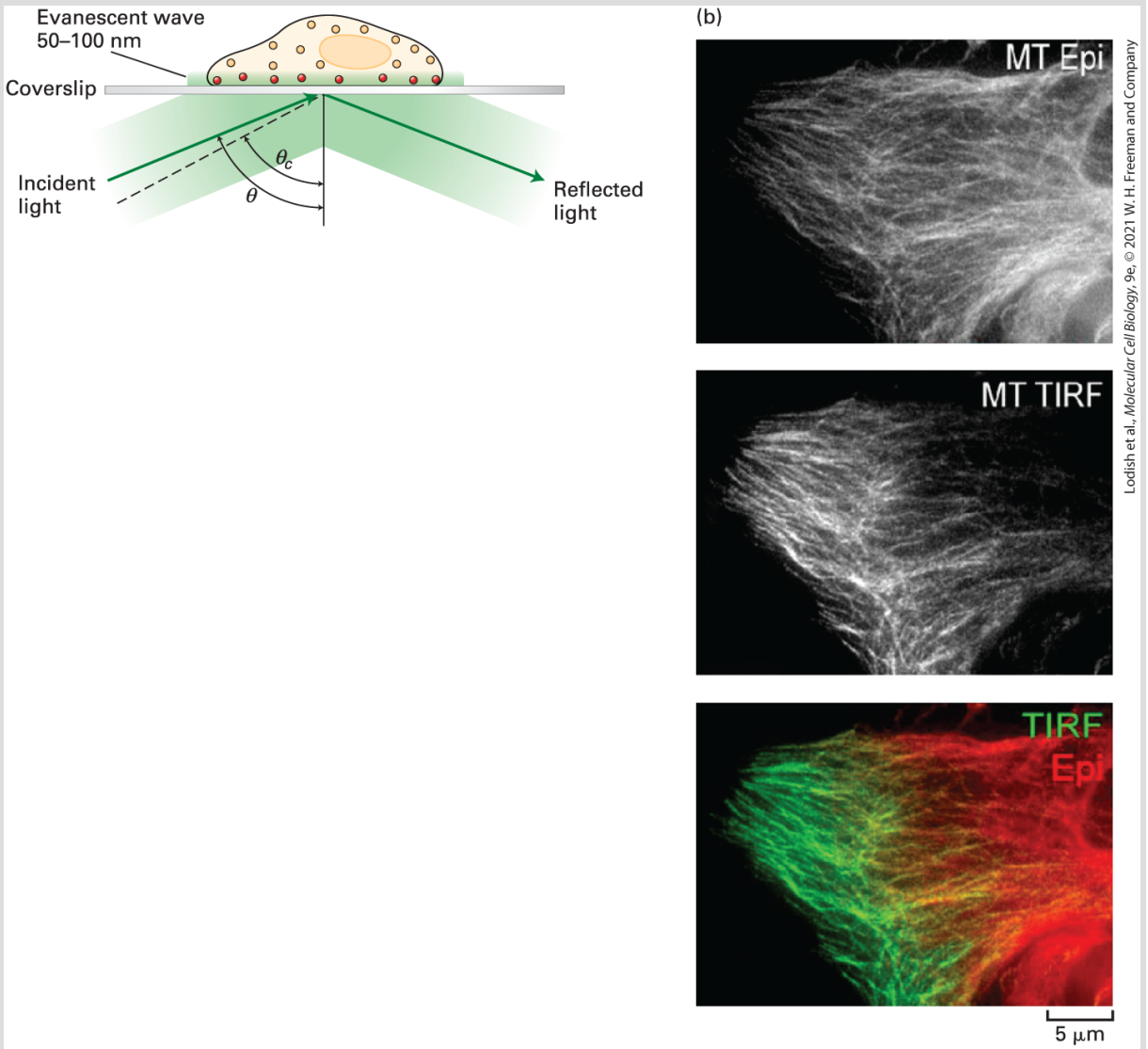
FIGURE 4-20 Two-photon excitation microscopy restricts illumination to the focal plane to allow deep penetration for intravital imaging. (a) Different excitation methods are used for conventional point-scanning confocal microscopy and for two-photon excitation microscopy. In the conventional system, absorption of a single photon of the appropriate wavelength (here at 488 nm, blue arrow) causes an electron to jump to the excited state. After undergoing vibrational relaxation (black dashed arrow), the electron falls back to the ground state with emission of one photon at a longer (lower energy) wavelength, in this case 507 nm (green arrow). In two-photon excitation, when two photons of the appropriate wavelength (shown here at 960 nm, red arrows) arrive almost instantaneously, they can both be absorbed and induce the electron to jump to the excited state. The electron undergoes some vibrational relaxation (black dashed arrow) and falls back to the ground state with the emission of a photon (507 nm). (b) A cuvette of fluorescent material is illuminated with 488-nm light (*left*), as in conventional confocal microscopy, or with intense 960-nm light, as in two-photon microscopy. Notice that the conventional system produces a bright cone of excitation outside the focal plane, whereas two-photon excitation illuminates just one spot in the focal plane. (c) An example of intravital imaging in which labeled neurons in a lobster were imaged.

[Part (b) Reprinted by permission from Nature Publishing Group, from Zipfel, W. R., et al. 2003. "Nonlinear magic: multiphoton microscopy in the biosciences," *Nat. Biotechnol.*, **21**:1369–1377; permission conveyed through Copyright Clearance Center, Inc.; part (c) unpublished data from Peter Kloppenburg and Warren R. Zipfel.]

TIRF Microscopy Provides Exceptional Imaging in One Focal Plane

The confocal microscopes we have just described provide amazing and informative images, but they are not perfect. Scientists continue to develop systems that are optimized for special circumstances. Some experimental situations call for fluorescence imaging in a thin focal plane adjacent to a surface, where it would be optimal to minimize out-of-focus background. For example, confocal imaging is not ideal for exploring the details of proteins at adhesion sites between a cell and a coverslip, or for following the kinetics of assembly of microtubules attached to a coverslip. Both of these situations can be imaged at high sensitivity using *total internal reflection fluorescence* (*TIRF*) microscopy, in which only the portion of the specimen immediately adjacent to the coverslip is illuminated. In the most common configuration of TIRF microscopy, the excitation light comes through the objective lens. However, the angle at which the light arrives at the coverslip is adjusted to be slightly larger than the critical angle θ_c so that the light is reflected off the coverslip and returns up through the objective (Figure 4-21a). This generates a narrow band of light, called an *evanescent wave*, that

illuminates only about 50–100 nm of the sample adjacent to the coverslip (2–4 times the thickness of a microtubule), with no illumination of the rest of the sample. Thus if you have a complex mixture of fluorescent structures in a specimen, the TIRF microscope will show you only those that are within 50–100 nm of the coverslip. TIRF has been exceptionally useful in identifying structures on the bottoms of cells grown on a coverslip (Figure 4-21b) and for measuring the kinetics of assembly and disassembly of structures such as microtubules and actin filaments (see Chapters 17 and 18).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

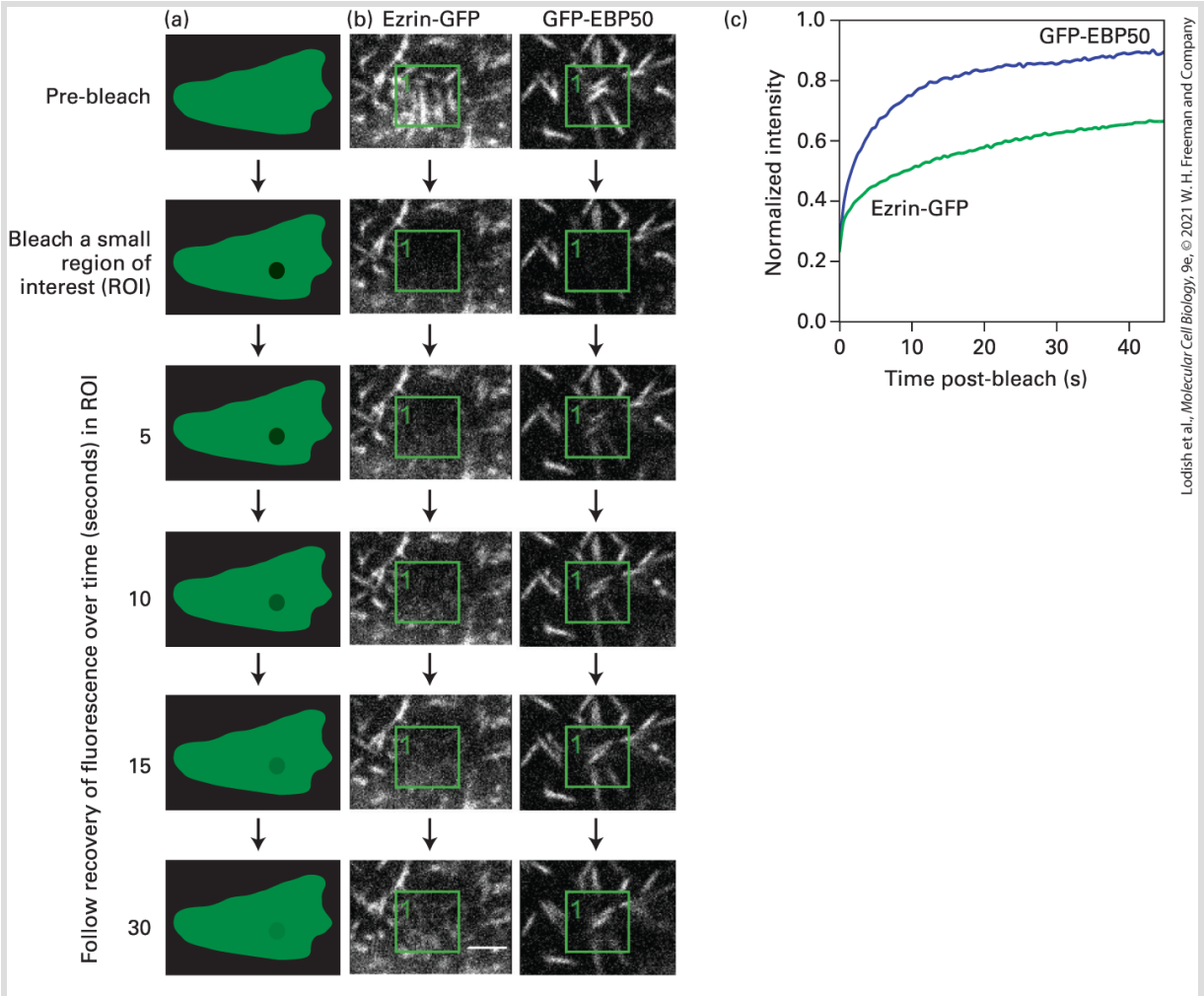
EXPERIMENTAL FIGURE 4-21 Fluorescent samples in a restricted focal plane can be imaged by total internal reflection fluorescence (TIRF) microscopy. (a) Illumination setup for TIRF microscopy. The angle at which light becomes reflected from a glass-water interface is called the *critical angle* θ_c. Below this angle, the beam of light will pass into the specimen; above θ_c, it will be reflected. In TIRF microscopy, the angle of the incident beam θ_c, resulting in most of the light being reflected, but also generating a very thin region of illumination called the evanescent wave (depicted in light green). (b) Immunofluorescence microscopy with tubulin antibody was used to visualize microtubules viewed by conventional fluorescence microscopy (*top*) and by TIRF (*middle*), and a merged image was created from the two views (*bottom*). The two images were collected and false-colored red and green so that the merge could highlight those microtubules that are close to the coverslip (green).

[Part (b) © 2010 J. B. Manneville et al., 2010, *J. Cell Biol.* **191**:585–598.
https://doi.org/10.1083/jcb.201002151]

FRAP Reveals the Dynamics of Cellular Components

Live cell fluorescence imaging reveals the locations and bulk dynamics of populations of fluorescent molecules, but it doesn't tell you how dynamic individual molecules are. For example, if we see that a GFP-labeled protein forms a patch at the surface of a cell, does this represent a stable collection of fluorescent protein molecules or a dynamic equilibrium, with fluorescent proteins coming in and out of the patch? We can investigate this question by observing the dynamics of the molecules in the patch (Figure 4-22). If we use a high-intensity light to permanently bleach the fluorochrome (e.g., GFP) in the patch, initially there will be no fluorescence coming from it, and it will look dark in the fluorescence microscope. However, if the components in the patch are in dynamic equilibrium with unbleached molecules elsewhere in the cell, the bleached molecules will be replaced by unbleached ones, and the fluorescence will begin to come back. The rate of fluorescence recovery is a measure of the dynamics of the molecules. This technique, known as *fluorescence recovery after photobleaching* (*FRAP*), has revealed how very dynamic many components of cells are. For example, it has been used to determine

the diffusion coefficient of cytoskeletal and membrane proteins (see Figure 10-10). In another approach to measuring the dynamics of tagged proteins, variants of fluorescent proteins have been developed that can be switched, using a laser of an appropriate wavelength, from emitting green light to emitting red light. In this way, the dynamics of the switched population of red-emitting molecules can be imaged in live cells.



EXPERIMENTAL FIGURE 4-22 Fluorescence recovery after photobleaching (FRAP) reveals the dynamics of molecules. In a live cell, following the distribution of a GFP-labeled protein provides a view of the overall distribution of the protein, but it doesn't tell us how dynamic

populations of individual molecules might be. (a) In FRAP, the GFP fluorochrome is bleached by a short burst of strong laser light focused on the region of interest (ROI). The light rapidly bleaches the molecules irreversibly, so they are not detected again. Restoration of fluorescence in the region tells us that unbleached molecules have moved into the ROI. (b) FRAP was used to determine the exchange rate of EBP50 and ezrin, two components of microvilli (seen as white lines), on the apical surface of epithelial cells. In cells expressing either GFP-EBP50 or ezrin-GFP, the GFP in a small region indicated by the green box was bleached, and recovery by exchange with unbleached protein was followed. The fluorescence of GFP-EBP50 returns very fast, indicating it has a fast exchange rate, whereas the fluorescence of ezrin-GFP returns slowly, indicating a slower exchange rate. (c) By quantifying the recovery, the dynamic properties of EBP50 and ezrin can be established.

[Parts (b) and (c) ©2012 D. Garbett and A. Bretscher, 2012, *J. Cell Biol.* **198**:195–203. https://doi.org/10.1083/jcb.201204008]

FRET Measures Distance Between Fluorochromes

Fluorescence microscopy can also be used to determine if two proteins interact in vivo by taking advantage of a phenomenon called **Förster resonance energy transfer (FRET)**. This technique uses a pair of fluorescent proteins in which the emission wavelength of the first is close to the excitation wavelength of the second (Figure 4-23a). For example, when cyan fluorescent protein (CFP) is excited with 433-nm light, it fluoresces and emits light at 475 nm. If yellow fluorescent protein (YFP) is close by, however, instead of emitting 475-nm light, CFP transfers energy to YFP by FRET, and YFP emits light at 530 nm. The efficiency of FRET is proportional to R^{-6}, where R is the distance between the fluorochromes; it is therefore very

sensitive to small changes in distance and in practice is not detectable at distances greater than 10 nm. Thus by illuminating an appropriately prepared sample with 433-nm light and observing at 530 nm, one can tell if proteins separately tagged with CFP and YFP are in very close proximity. For example, FRET sensors have been developed to determine where signaling between a small GTP-binding regulatory protein and its effector occurs in the cell (Figure 4-23b).

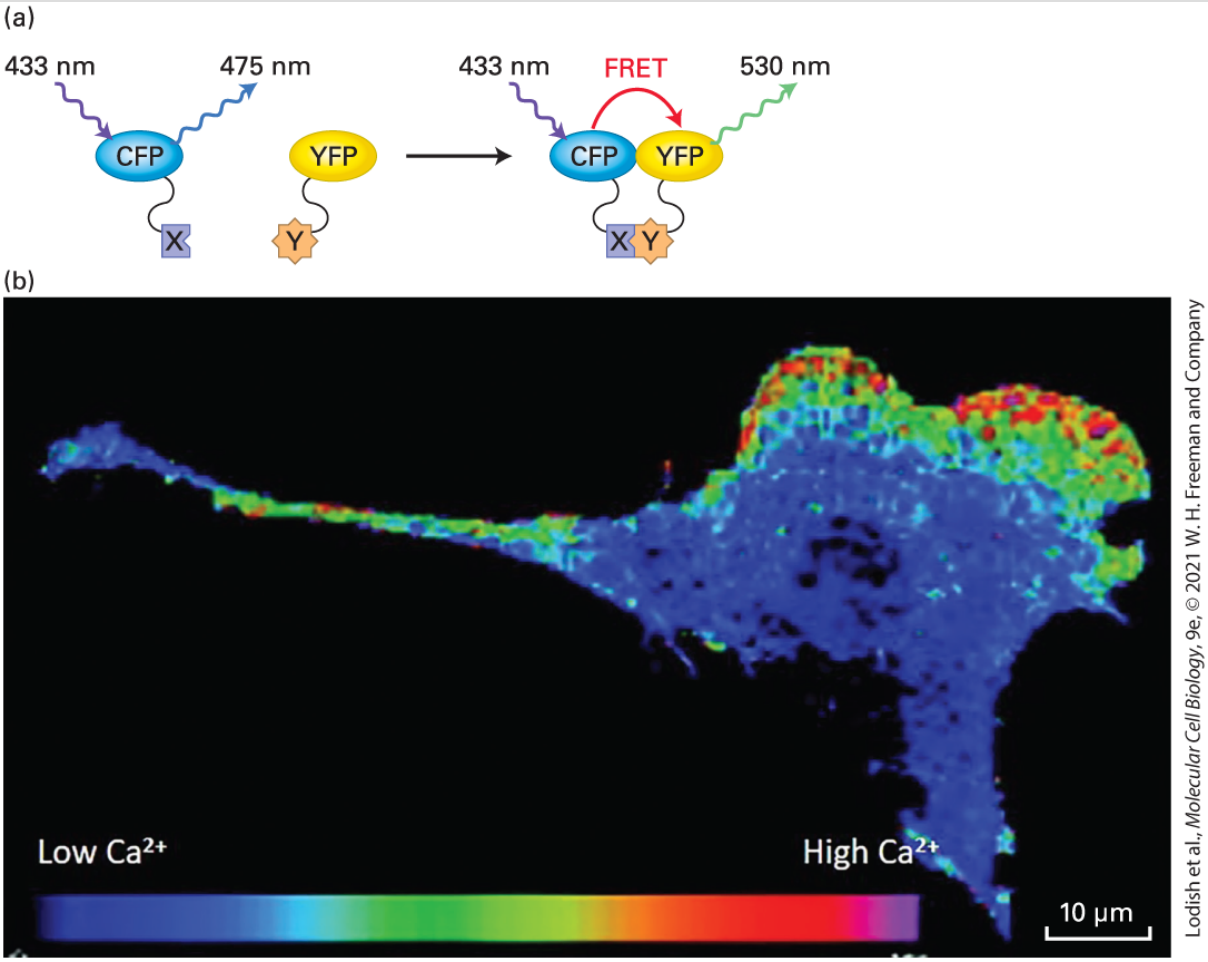


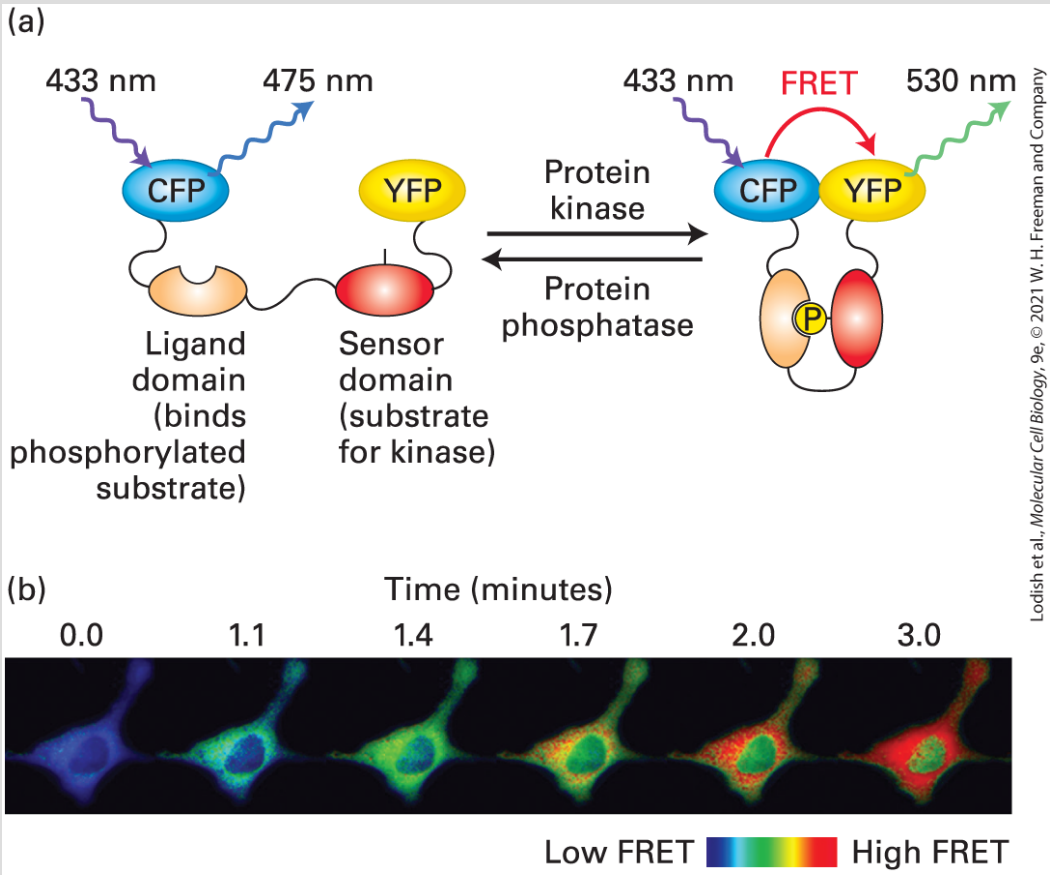
FIGURE 4-23 Protein-protein interactions can be visualized by FRET. The idea behind FRET is to use two different fluorescent proteins so that when one is excited, energy will be transferred to the second one by FRET, provided that they are sufficiently close. (a) In this

example, cyan fluorescent protein (CFP) is fused to protein X, yellow fluorescent protein (YFP) is fused to protein Y, and both proteins are expressed in a live cell. If the cell is now illuminated with 433-nm light, the CFP will emit a fluorescent signal at 475 nm. If YFP is not close by (*left*), energy transfer will not occur, and no 530-nm light will be emitted. However, if protein X interacts with protein Y (*right*), it will bring CFP close to YFP, energy will be transferred to YFP by FRET, and YFP will emit light at 530 nm. (b) In this mouse fibroblast, FRET has been used to reveal that the interaction between an active regulatory protein (Rac) and its binding partner is localized to the front of the migrating cell.

[Part (b) ©2003 R. B. Sekar and A. Periasamy, 2003, *J. Cell Biol.* **160**:629–633. https://doi.org/10.1083/jcb.200210140]

A clever application of FRET, called a *FRET biosensor*, can be used to sense local biochemical environments in live cells. The idea is to express a single polypeptide containing both CFP and YFP separated by a region that undergoes a conformational change when it senses a biochemical signal. In the absence of the signal, CFP and YFP are too far apart to undergo significant FRET. However, when the signal is detected, the conformational change brings CFP and YFP close enough together to generate FRET. A version of this technique can be used to measure the local activity of a specific protein kinase. In this case, between the CFP and YFP lies a region of polypeptide containing the substrate for the protein kinase — the *sensor domain* — and a domain that binds specifically to the phosphorylated substrate — the *ligand domain* (Figure 4-24a). When the sensor domain is phosphorylated by the kinase, the ligand domain binds to it and brings the CFP and YFP sufficiently close to undergo FRET. Since protein phosphorylation is a dynamic process, dephosphorylation of the sensor domain by the appropriate

phosphatase will deactivate the FRET biosensor. Thus the FRET signal will reflect the regions of the cell where there is excess kinase over phosphatase activity. As an example, scientists have developed a FRET biosensor for protein kinase A activity, which is activated by elevation of the signaling molecule cAMP (see Section 15.1). In cells expressing the protein kinase A biosensor, pharmacological elevation of cAMP induces rapid FRET (Figure 4-24b). Creative researchers are developing FRET biosensors to illuminate many different types of local environments; for example, FRET biosensors exist to measure the concentration and location of Ca^{2+} (see Figure 4-27) and the activation state of GTPase switch proteins (see Figure 3-35).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 4-24 FRET biosensors can detect local biochemical environments. (a) A FRET biosensor is a fusion protein containing two fluorescent proteins linked by a region sensitive to the environment under study. In this example, a protein construct consists of CFP linked to YFP by a region that contains a particular sequence that can be phosphorylated by a specific kinase (the sensor domain) and a region (the ligand domain) that binds the sensor domain when it is phosphorylated. In the absence of kinase activity, the two fluorescent proteins are too far apart to undergo FRET, whereas when locally phosphorylated by the active kinase, the sensor domain becomes phosphorylated, the ligand domain binds to it, and CFP and YFP are brought sufficiently close to undergo FRET. The sensor can also be deactivated when it encounters the appropriate phosphatase that removes the added phosphate. Thus the biosensor reports on the ratio of kinase to phosphatase activity in the local environment. (b) An example of the use of a FRET biosensor for protein kinase A, which is activated by elevation of cAMP. In this example, forskolin, a drug that induces the generation of cAMP, was added to cells at $t = 0$ and the

images collected at various times thereafter. Imaging shows both the rate of activation and localization of the active kinase.

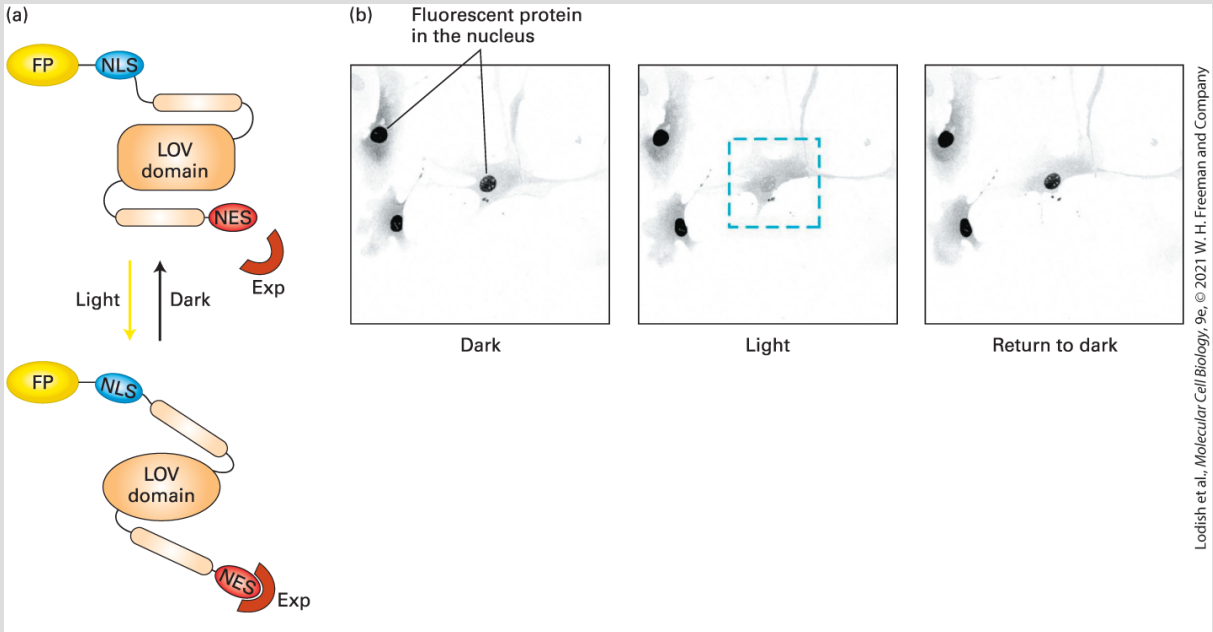
[Part (b) from J. Zhang et al., 2001, *Proc. Nat'l. Acad. Sci. USA* **98**:14997–15005. © 2001 National Academy of Sciences, U.S.A.]

Optogenetics Allows Light to Regulate Events in a Spatial and Temporal Manner

The idea that light can be used to control biochemical reactions in living cells arose from the study of phototaxis (attraction to light) of the green alga *Chlamydomonas reinhardtii*. Scientists discovered that a pulse of bright light induced the opening of a cation channel in a protein now known as *channelrhodopsin*. When channelrhodopsin is expressed in neurons, the opening of the channel by light allows the very rapid influx of Na^+ that depolarizes the membrane. As we will see in Chapter 23, expression of channelrhodopsins in specific neurons combined with localized pulses of light can be used to dissect neural circuits.

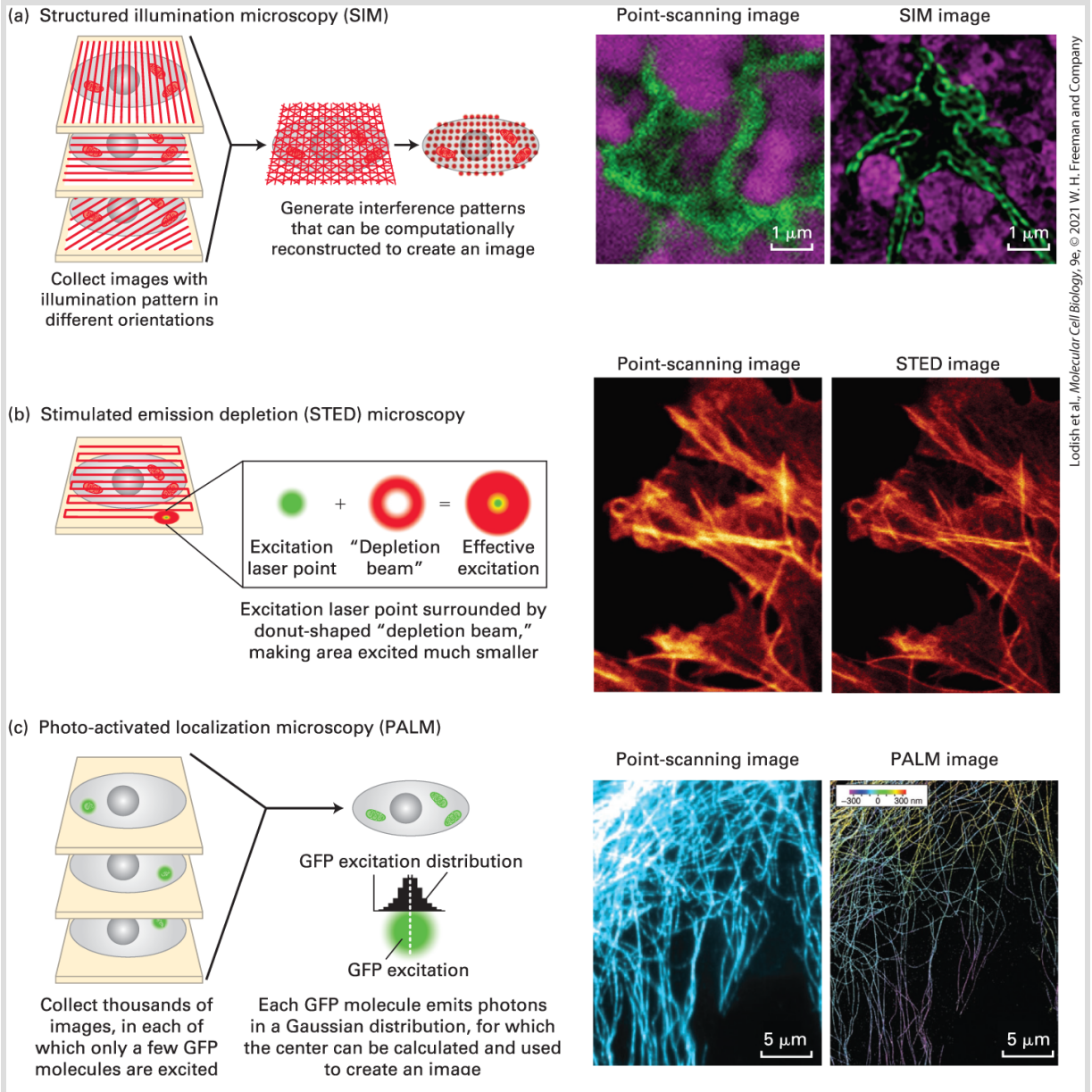
In a subsequent advance, methods were developed to use light to control biochemical reactions locally. This technology grew out of the identification of proteins that control plant phototropism (orientation toward, or away, from light). One of these, using the light sensitive LOV domain, has been extensively used. In an

example to demonstrate the utility of this approach, light was used to induce an engineered protein located in the nucleus to move into the cytoplasm. As described in Chapter 13, proteins that are targeted to the nucleus contain a short nuclear localization sequence (NLS), and those that are exported from the nucleus have a short nuclear export sequence (NES) that can interact with a protein called an *exportin* to transport it out into the cytoplasm. A multidomain construct was made containing the following parts: Fluorescent protein-NLS-LOV-NES (Figure 4-25a). The construct was made so that in the dark the NLS sequence was accessible, whereas the NES was masked. When expressed in cells in the dark, the protein is located in the nucleus as expected because NES cannot interact with exportin (Figure 4-25a, *top*). However, when illuminated, the LOV domain changes conformation, releasing the NES, which can now bind exportin (Figure 4-25a, *bottom*). This change results in export of the protein from the nucleus. When the illumination is turned off, the LOV domain reverses its conformation, masking the NES, and the protein returns to the nucleus (Figure 4-26b). If instead of the fluorescent protein a nuclear regulator was made part of the construct, this would allow light-mediated regulation of its activity. This type of approach has been used in many different settings, for example, to recruit proteins to the plasma membrane or to activate signaling pathways.



EXPERIMENTAL FIGURE 4-25 Use of the LOV domain and light to alter the localization of a protein. (a) In this example, a construct was made in which a fluorescent protein was fused to a nuclear localization sequence (NLS), the LOV domain, and a nuclear export sequence (NES). In the dark the NES sequence is not accessible to interact with exportin (Exp), but when illuminated a conformational change in the LOV domain releases the NES that can now interact with exportin. (b) When this construct is expressed in cells, the construct localizes to the nucleus because the NLS sequence is accessible, whereas the NES sequence is not. However, upon illumination of a single cell with blue light (box), the LOV domain undergoes a conformational change, releasing the NES so that it is accessible to exportin and the protein is then transported out of the nucleus. This is reversible, as the protein returns to the nucleus when the illumination is turned off. The images in (b) show an inverted fluorescence image of the cells expressing the construct.

[Republished with permission of Nature Publishing Group from Yumerefendi H, et all, Light-induced nuclear export reveals rapid dynamics of epigenetic modifications. *Nat. Chem. Biol.,* 17 Apr 2016, **12** (6): 399–401. Permission conveyed through Copyright Clearance Center, Inc.]



EXPERIMENTAL FIGURE 4-26 Super-resolution microscopy can generate light-microscope images with up to nanometer resolution. The theoretical resolution of the light microscope can be circumvented by super-resolution microscopy. (a) In structured illumination microscopy (SIM), the sample is illuminated by a pattern of light and dark stripes and several images are taken as the illumination is rotated. This technique generates interference patterns that can be mathematically reconstructed to generate a higher-resolution image. The images on the right show similar fields of the nucleus (lamin, green; DNA, magenta) imaged by conventional point-scanning confocal microscopy and by SIM, which improves resolution about twofold. (b) In stimulated emission depletion (STED) microscopy, the sample is scanned as in point-scanning microscopy, but with a very small

point of light, generated by an emission laser and confined by a donut-shaped stimulated emission depletion zone. The sample at the right shows part of a cell stained for actin fibers after imaging by point-scanning microscopy and by STED. (c) In photoactivated localization microscopy (PALM), use is made of a variant of GFP that can be photoactivated by a wavelength different from its excitation wavelength. When a small number of GFP molecules are activated, and then excited, each will emit thousands of photons that can be collected. This generates a Gaussian curve centered on the location of the emitting GFP; the center provides the location of the GFP to nanometer accuracy. This process is reiterated hundreds of times to excite other GFP molecules, and a high-resolution image emerges. At right, a confocal image of microtubules is compared with a corresponding super-resolution image in which the three-dimensional arrangement of the microtubules is color coded.

[Micrographs in (a) republished by permission of the American Association for the Advancement of Science, from: Schermelleh, L. et al., 2008. "Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy," *Science*, **320**: 1332–1336; permission conveyed through the Copyright Clearance Center, Inc.; in (b) from Dr. Elise Stanley, Toronto Western Research Institute; (c) republished by permission of the American Association for the Advancement of Science, from: Huang, B. et al., 2008. "Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy," *Science*, **319**: 810–813; permission conveyed through the Copyright Clearance Center, Inc.]

Point Source Fluorescent Objects Can Be Located at Nanometer Resolution

Although a conventional microscope cannot resolve two objects closer together than 200 nanometers, it can localize a stationary single object to within a few nanometers. If we have a point-source of fluorescence — say, a single fluorescently tagged protein — and use a sensitive camera to rapidly collect and image multiple photons, this will generate a Gaussian curve of fluorescence

intensity around the precise position of the tagged protein. A computer can then calculate the average position to reveal the center of the object to within a few nanometers. In this way, computer algorithms can be used to track the position of single objects, or even moving objects as long as the exposure time for each image is sufficiently short so that the object does not move appreciably. As we discuss in a later chapter, this technique has been used to measure nanometer-sized steps as molecules and vesicles move along cytoskeletal filaments (see Figures 17-28 and 17-29).

Super-Resolution Microscopy Can Localize Proteins to Nanometer Accuracy

As we discussed earlier, the theoretical resolution limit of the fluorescence microscope is about 0.2 μm (200 nm). Two new general approaches, collectively known as *super-resolution microscopy*, have been developed to get around this limitation. In the first type of approach, the illuminating light is patterned in such a way that higher resolution images are obtained. In *structured illumination microscopy* (*SIM*), the specimen is illuminated with a pattern of light and dark stripes, and several images are taken as the illuminating pattern is rotated. Computational analysis of the images gives a resolution of about 100 nm, twice that of a conventional confocal

microscope, as can be seen in a micrograph of invaginations in part of the nuclear envelope (Figure 4-26a). SIM is especially good for live cell imaging, as three-dimensional images can be collected every 4 seconds. In *stimulated emission depletion* (*STED*) *microscopy*, the sample is scanned just as in point-scanning microscopy, but with an important difference: the focused excitation laser point is surrounded by a donut-shaped depletion beam, which effectively makes the area excited much smaller. Since the computer records the precise position of the spot excited and can record the emission from it, it can build up an impressive image, optimally yielding a resolution of 30 nm, greatly enhancing the detail one can see, for example, in an image of actin fibers (Figure 4-26b).

The second general approach uses point-source detection and localization as described in the previous section. To understand how this works, consider two fluorescent spots separated by 75 nm. When you try to image them together, they each generate a Gaussian distribution of fluorescence that overlaps so much that they look like one spot. However, if you could image each spot individually and find the center of each Gaussian curve, you could beat the resolution limit and detect the two spots 75 nm apart. One way to do this, called *photoactivated localization microscopy* (*PALM*), also known as *stochastic optical reconstruction microscopy* (*STORM*), relies on the ability of a variant of GFP to be photoactivated; that is, it can become fluorescent only after being activated by a specific wavelength of light, different from its excitation wavelength.

Consider what happens when we activate just one such GFP molecule. When we then excite the sample, the one activated GFP emits many hundreds of photons, giving rise to a Gaussian distribution (Figure 4-26c). Although analysis of each photon does not tell us precisely where the GFP is, the center of the distribution can tell us where the GFP is located with nanometer accuracy. If we now activate another GFP, we can localize it individually with the same precision. In PALM/STORM, a small percentage of GFPs are activated and each localized with high precision, and then another set is activated and localized, and as additional cycles of activation and localization are recorded, a high-resolution image emerges. In practice, super-resolution images have a resolution of about 30 nm in the focal plane, and 50 nm perpendicular to it. For example, the three-dimensional distribution of microtubules can be seen with much greater clarity than with conventional fluorescence microscopy (see Figure 4-26c). These types of images can take significant time to generate, so their use on samples of live cells is so far restricted. Nevertheless, because of the tremendous benefits of super-resolution microscopy, enormous efforts are being made to improve its sensitivity, speed, and spatial resolution, and we can expect rapid progress in the development of these approaches. Indeed, a new variation of STED combined with PALM/STORM, called *MINFLUX*, has been reported to result in a resolution close to 1 nm, although it has so far only been used in very specialized situations.

Light-Sheet Microscopy Can Rapidly Image Cells in Living Tissue

Most of the confocal imaging approaches we have discussed above illuminate and detect fluorochromes through the same objective lens; this ensures that the excitation of the sample and the imaging of the emitted fluorescence occur in the same focal plane. As a result of this approach, there are limitations on the depth of sample that can be imaged. A new technology has been developed to get around this limitation. In *light-sheet microscopy*, the sample is illuminated from the side and then viewed in an orthogonal direction (Figure 4-27a). A focused laser beam sweeps back and forth across the sample, illuminating a single plane. A detection objective, at right angles to the illuminated plane, then images the sample. To generate a three-dimensional image, all the planes in the sample have to be imaged. This is achieved by coordinately stepping the illuminating sheet and detection objective through the sample to generate a stack of images, which can then be assembled computationally into a three-dimensional rendering.

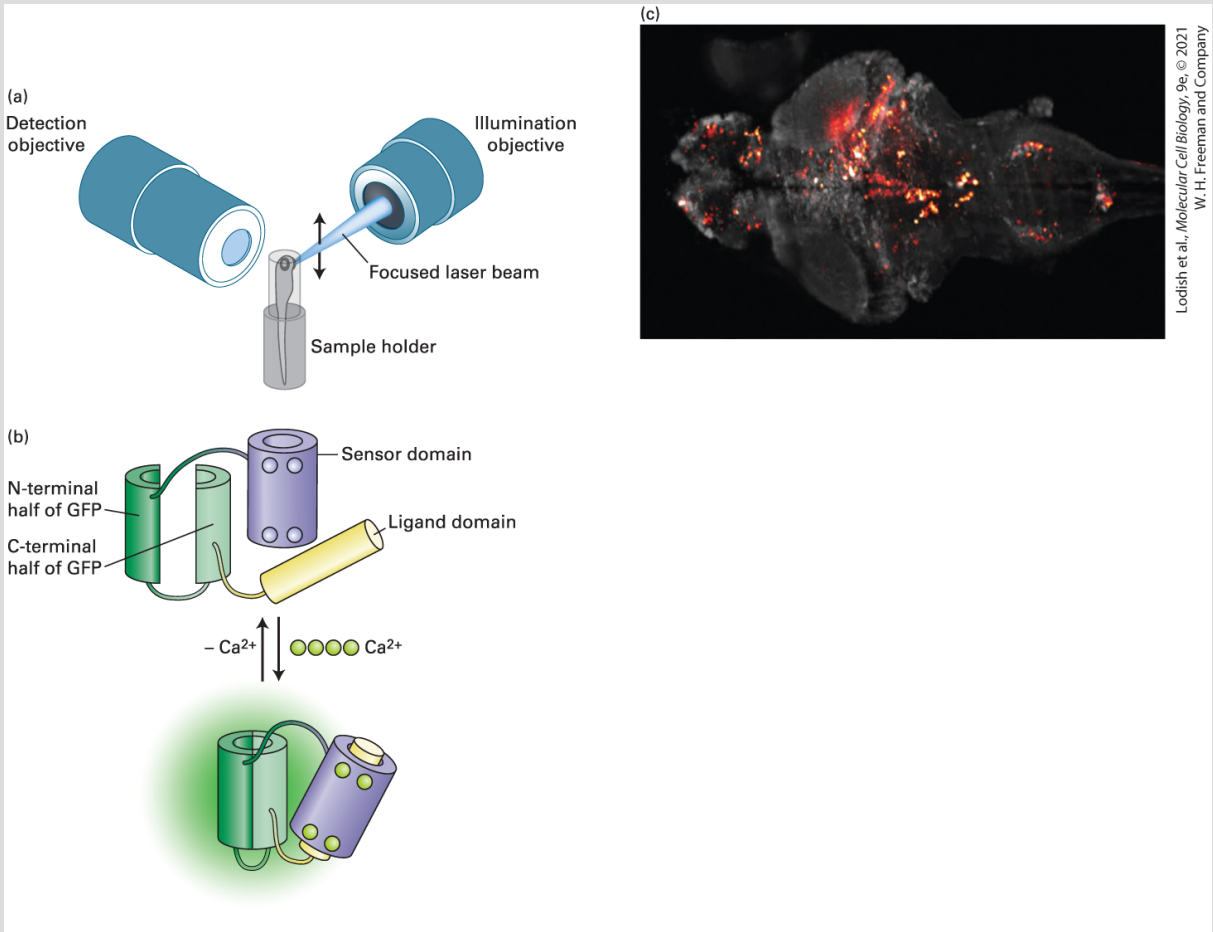


FIGURE 4-27 Light-sheet microscopy can image rapid events in living tissue. (a) In light-sheet microscopy, a tissue sample is illuminated from the side by a focused laser beam that scans the sample (arrows) to generate a sheet of light. The sample is observed in the orthogonal direction through the detection objective. To get a three-dimensional image, the illuminating and detection objectives are moved coordinately, taking images throughout the depth of the sample. (b) The Ca^{2+} biosensor known as GCaMP. This biosensor is made using recombinant DNA techniques to generate a polypeptide to which the N- and C-termini of GFP are fused and the middle interrupted. On one side of the interruption is the sensor domain, consisting of the Ca^{2+}-binding protein calmodulin. On the other side is the ligand domain, consisting of a target sequence to which calmodulin will bind in the presence of Ca^{2+}. In the absence of Ca^{2+}, the GFP is not functional due to the interruption. In the presence of Ca^{2+}, calmodulin binds four Ca^{2+} ions and undergoes a conformational change that allows it to bind the ligand domain. This conformational change brings the two parts of GFP into close proximity so that the fluorescent protein is functional. (c) Local increases in the level of Ca^{2+}, false-colored red, in cells of the brain of a living zebrafish.

[Part (c) reprinted by permission from Nature Publishing Group, from Ahrens, M.B., et al., 2013. "Whole-brain functional imaging at cellular resolution using light-sheet microscopy," *Nat. Methods* **10**:413–420; permission conveyed through the Copyright Clearance Center, Inc.]

In one example of the use of this technology, scientists have imaged concentrations of Ca^{2+} in the cells of a living zebrafish brain. The neurons in the brain of the zebrafish were made to express a Ca^{2+} biosensor, called *GCaMP*, which is based on GFP that does not fluoresce in the absence of Ca^{2+} but is designed to fluoresce in its presence (Figure 4-27b). This biosensor consists of GFP in which the N- and C-terminal domains are connected, but the middle of the protein is interrupted. A Ca^{2+} sensor domain is attached to one end of the GFP and a ligand domain to the other end. Because the structure of GFP is interrupted, the protein no longer fluoresces when excited by the appropriate illuminating light. The sensor domain is derived from the small Ca^{2+}-binding protein calmodulin (see Figure 3-34), which changes its conformation when it binds four Ca^{2+} ions. The ligand domain binds calmodulin only after calmodulin is activated by binding Ca^{2+}. When active calmodulin binds the ligand domain, it changes the conformation of the interrupted GFP in such a way that it can now be excited and fluoresce. In this way, a fluorescent signal is reported whenever Ca^{2+} levels rise sufficiently to activate calmodulin. As we discuss in Chapter 23, when neurons communicate, part of this communication involves an elevation in Ca^{2+} levels. This neuronal

communication can be nicely imaged by light-sheet microscopy in vivo in the brain of a living zebrafish expressing the GCaMP biosensor (Figure 4-27c).

KEY CONCEPTS OF SECTION 4.2

Light Microscopy: Exploring Cell Structure and Visualizing Proteins Within Cells

- The resolution of the light microscope, about $0.2\ \mu\mathrm{m}$, is limited by the wavelength of light.
- Differences in refractive index can be used to observe parts of single cells by employing phase-contrast and differential-interference-contrast microscopy.
- Tissues generally have to be fixed, sectioned, and stained for cells and subcellular structures to be observed.
- Fluorescence microscopy makes use of compounds that absorb light at one wavelength and emit it at a longer wavelength.
- Ion-sensitive fluorescent dyes can measure intracellular concentrations of ions, such as Ca^{2+}.
- Immunofluorescence microscopy makes use of antibodies to localize specific components in fixed and permeabilized cells.
- Indirect immunofluorescence microscopy uses an unlabeled primary antibody that recognizes a specific protein, followed by a fluorescently labeled secondary antibody that recognizes the primary one and allows it to be visualized.
- Short sequences encoding epitope tags can be appended to protein-coding sequences to allow localization of the expressed protein using an antibody to the epitope tag.
- Green fluorescence protein (GFP) and its derivatives are naturally occurring fluorescent proteins.
- Expressing a fusion of GFP to a protein of interest allows its localization and dynamics to be explored in a live cell.
- Deconvolution and confocal microscopy provide greatly improved clarity in fluorescent images by removing out-of-focus fluorescent light.
- Total internal reflection fluorescence (TIRF) microscopy allows fluorescent samples adjacent to a coverslip to be seen with great clarity.
- Fluorescence recovery after photobleaching (FRAP) allows the dynamics of a population of molecules to be analyzed.

- Förster resonance energy transfer (FRET) is a technique in which light energy is transferred from one fluorescent protein to another when the proteins are very close, thereby revealing when two molecules are close in the cell.
- Optogenetics allows in vivo imaging to follow the effect of local and rapid biochemical changes induced by light.
- Super-resolution microscopy allows for detailed fluorescent images at nanometer resolution.
- Light-sheet microscopy can provide fluorescent images of thick samples by illuminating the sample with a sheet of light from the side.

4.3 Electron Microscopy: High-Resolution Imaging

Electron microscopic imaging of biological samples, such as single proteins, organelles, cells, and tissues, offers a much higher resolution of ultrastructure than can be obtained by light microscopy. The short wavelength of electrons means that the limit of resolution for a transmission electron microscope is theoretically 0.005 nm (much less than the diameter of a single atom) or 40,000 times better than that of a light microscope and 2 million times better than that of the unaided human eye. However, the effective resolution of the transmission electron microscope in the study of biological systems is considerably less than this ideal. Under optimal conditions, a resolution of 0.10 nm can be obtained with transmission electron microscopes, about 2000 times better than with conventional light microscopes.

The fundamental principles of electron microscopy are similar to those of light microscopy; the major difference is that in electron microscopes, electromagnetic lenses focus a high-velocity electron beam instead of the visible light used by optical lenses. In **transmission electron microscopy (TEM)**, electrons are emitted from a filament and accelerated in an electric field (Figure 4-28, *left*). A condenser lens focuses the electron beam onto the sample; objective and projector lenses focus the electrons that pass through the specimen and project them onto a viewing screen or other

detector. Because atoms in air absorb electrons, the entire tube between the electron source and the detector is maintained under an ultrahigh vacuum. Thus living material cannot be imaged by electron microscopy.

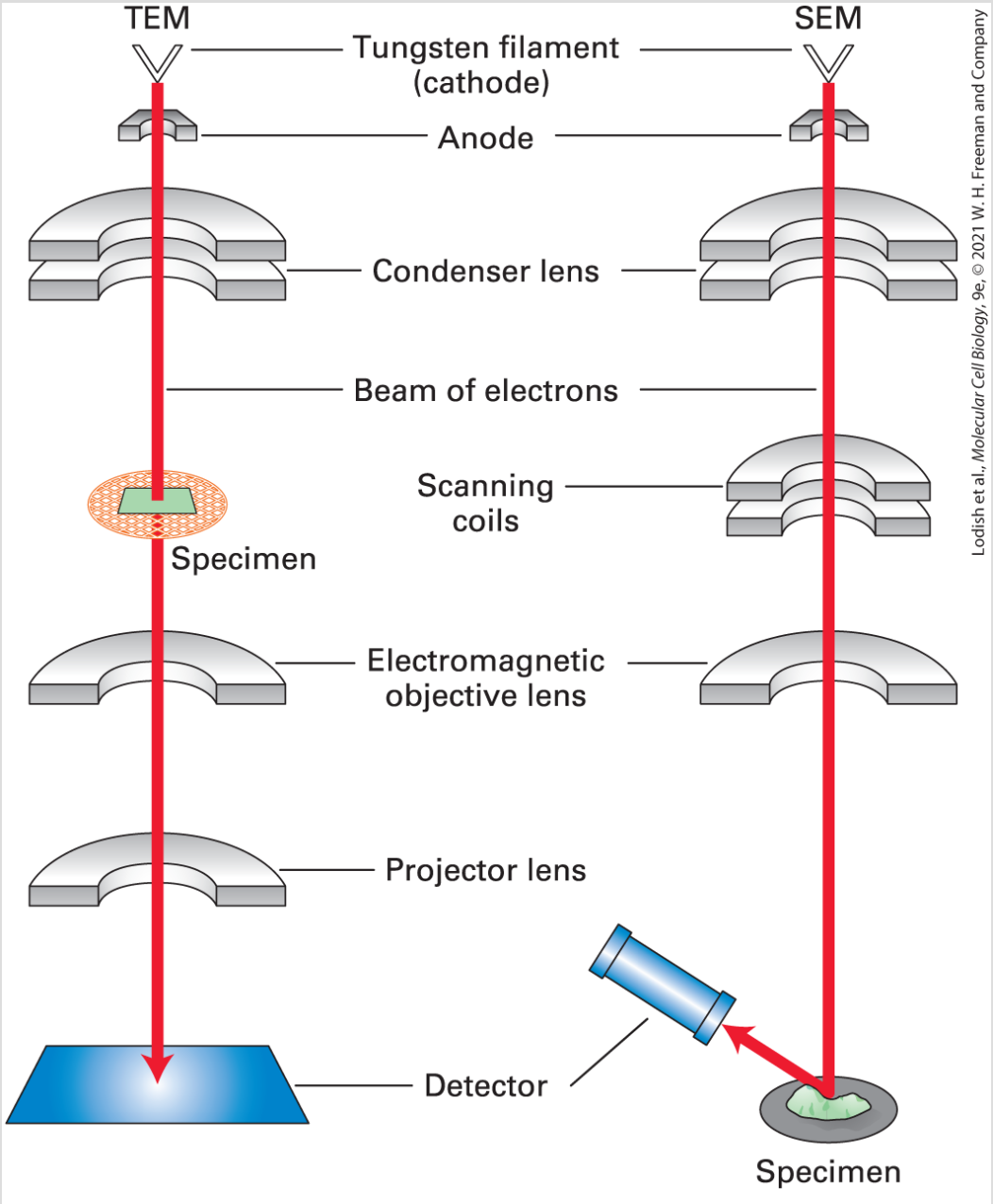


FIGURE 4-28 In electron microscopy, images are formed from electrons that pass through a specimen or are scattered from a metal-coated specimen. In a

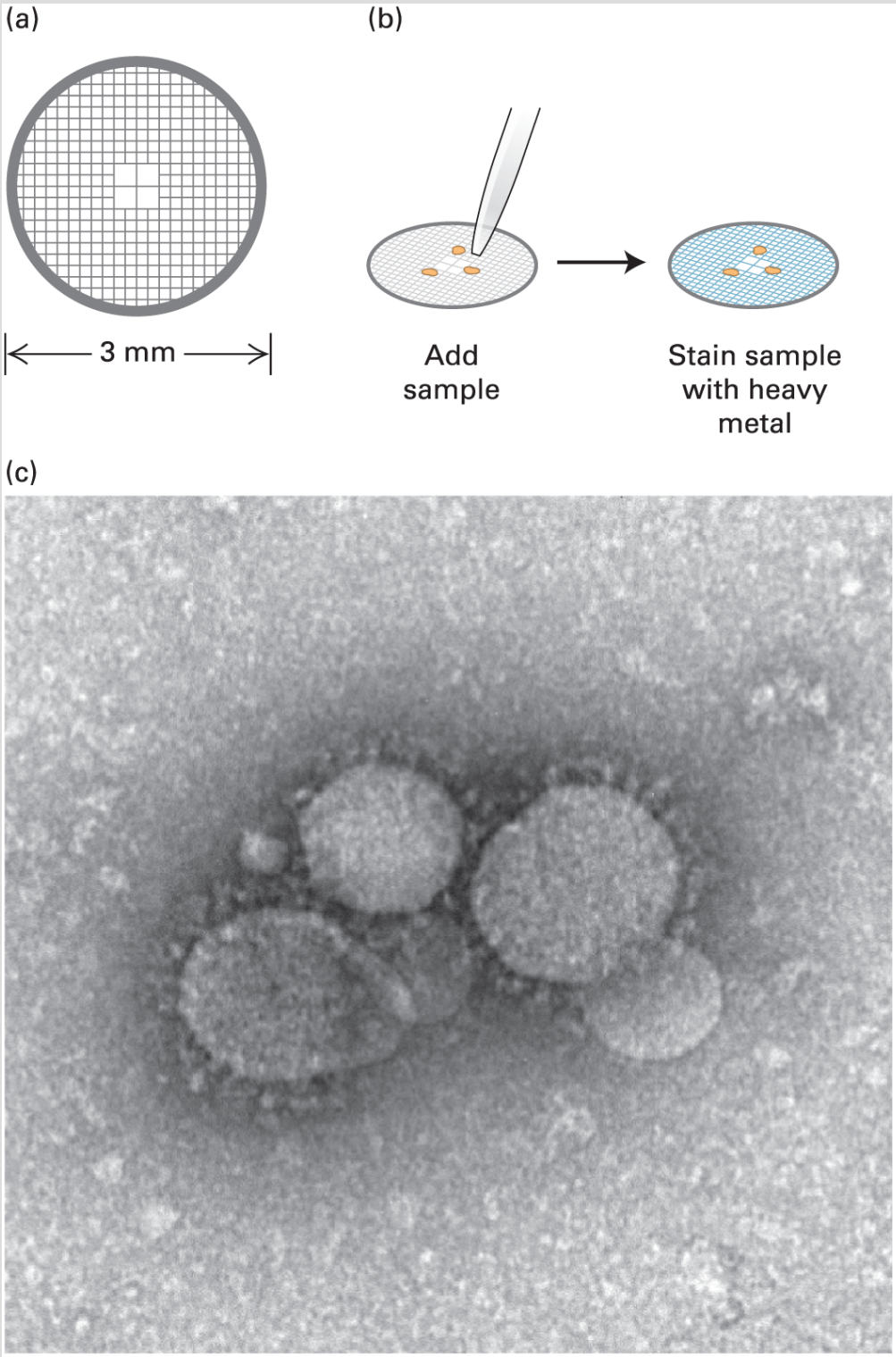
transmission electron microscope (TEM, *left*), electrons are extracted from a heated filament, accelerated by an electric field, and focused on the specimen by a magnetic condenser lens. Electrons that pass through the specimen are focused by a series of magnetic objective and projector lenses to form a magnified image of the specimen on a detector, which may be a fluorescent viewing screen, a photographic film, or a charged-couple-device (CCD) camera. In a scanning electron microscope (SEM, *right*), electrons are focused by condenser and objective lenses on a metal-coated specimen. Scanning coils move the beam across the specimen, and electrons scattered from the metal are collected by a photomultiplier tube detector. In both types of microscopes, because electrons are easily scattered by air molecules, the entire column is maintained at a very high vacuum.

In this section, we describe various approaches to viewing biological material by electron microscopy. The most widely used instrument is the transmission electron microscope, but also in common use is the *scanning electron microscope* (*SEM*), which provides complementary information, as we discuss at the end of this section (Figure 4-28, *right*).

Single Molecules or Structures Can Be Imaged Using a Negative Stain or Metal Shadowing

It is common in biology to explore the detailed shapes of single macromolecules, such as proteins or nucleic acids, or of structures, such as viruses and the filaments that make up the cytoskeleton. It is

relatively easy to view these objects in the transmission electron microscope, provided they are stained with a heavy metal that scatters the incident electrons. A sample is prepared first by adsorbence to a 3-mm *electron microscope grid* (Figure 4-29a), which is coated with a thin film of plastic and carbon. The sample is then bathed in a solution of a heavy metal, such as uranyl acetate, and excess solution is removed (Figure 4-29b). As a result of this procedure, the uranyl acetate coats the grid, but is excluded from the regions where the sample has adhered. When we view the sample in the TEM, we see where the stain has been excluded, so the sample is said to be *negatively stained*. Because the stain can precisely reveal the topology of the sample, a high-resolution image can be obtained (Figure 4-29c).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Part (c) CDC/Maureen Metcalfe; Azaibi Tamin, Ph.D.

FIGURE 4-29 Transmission electron microscopy of negatively stained samples reveals fine features. (a) Samples for transmission electron microscopy (TEM) are usually mounted on a small copper or gold grid. The grid is usually covered

with a very thin film of plastic and carbon to which a sample can adhere. (b) The specimen is then incubated in a heavy metal, such as uranyl acetate, and excess stain is removed. (c) The sample excludes the stain, so when it is observed in the TEM, it is seen in negative outline. The example in (c) is a negative stain of a coronavirus, very similar to the one that caused the COVID-19 pandemic.

Samples can also be prepared by *metal shadowing* (Figure 4-30). In this technique, the sample is adsorbed to a small piece of mica, then coated with a thin film of platinum by evaporation of the metal, and dissolved with acid or bleach, leaving the platinum coating (known as a replica). When the replica is transferred to a grid and examined in the TEM, it provides information about the three-dimensional topology of the sample.

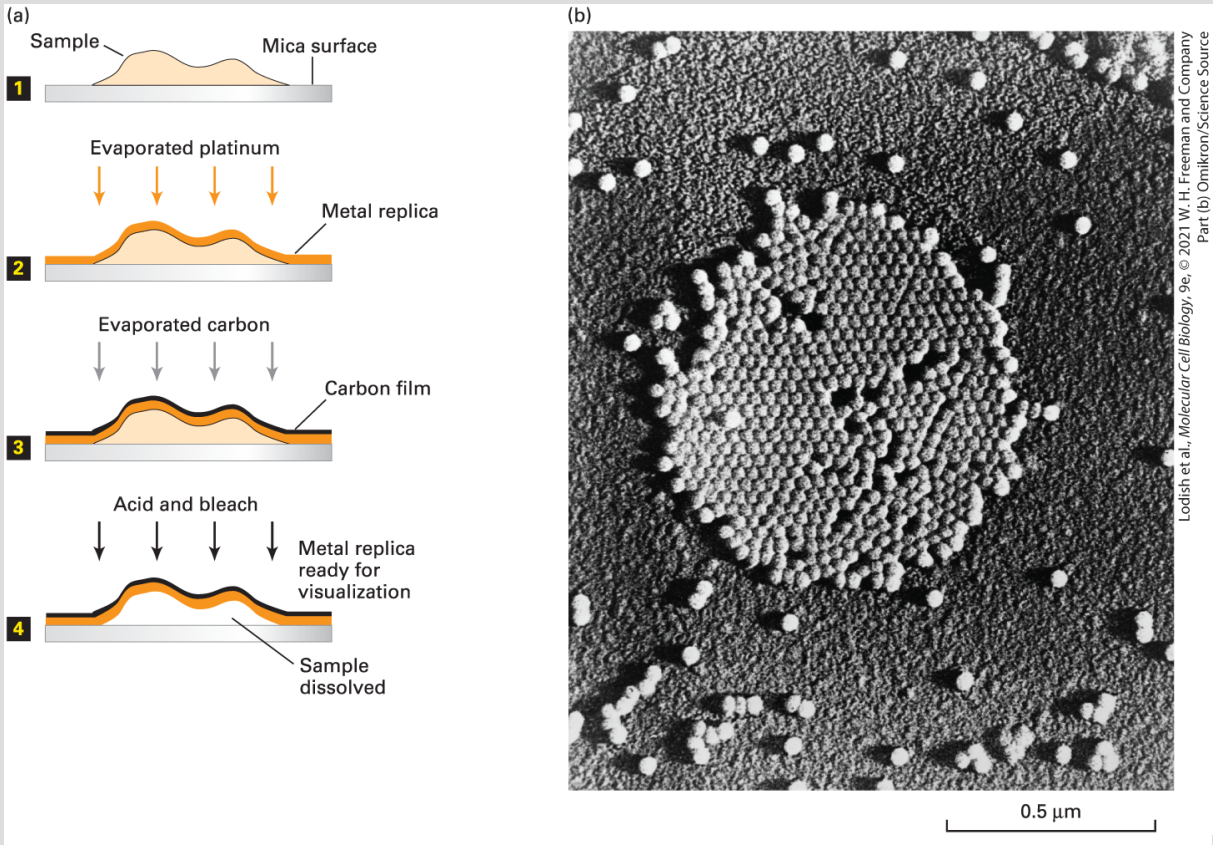
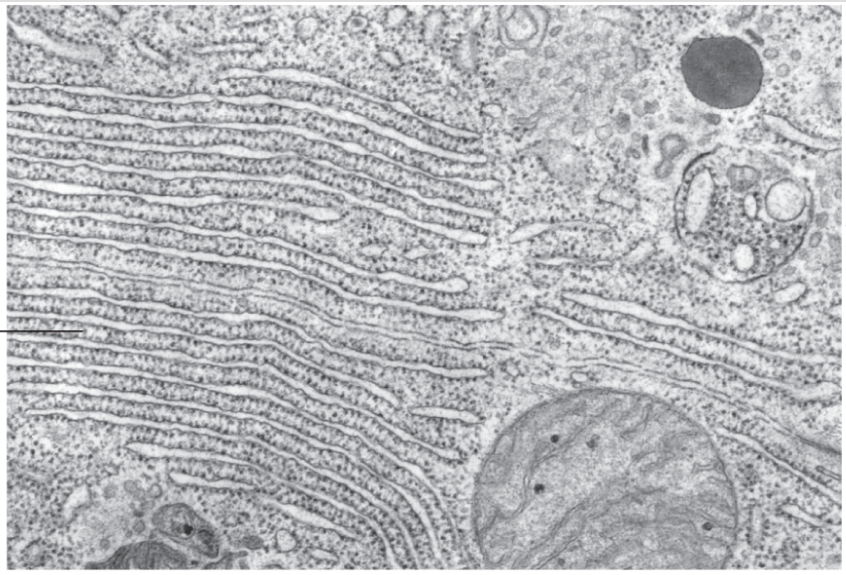


FIGURE 4-30 Metal shadowing makes surface details on very small objects visible by transmission electron microscopy. (a) The sample is spread on a mica surface and then dried in a vacuum evaporator (step 1). The sample grid is coated with a thin film of a heavy metal, such as platinum or gold, evaporated from an electrically heated metal filament (step 2). To stabilize the replica, the specimen is then coated with a carbon film evaporated from an overhead electrode (step 3). The biological material is then dissolved by acid and bleach (step 4), and the remaining metal replica is viewed in a TEM. (b) A platinum-shadowed replica of poliovirus particles.

Cells and Tissues Are Cut into Thin Sections for Viewing by Electron Microscopy

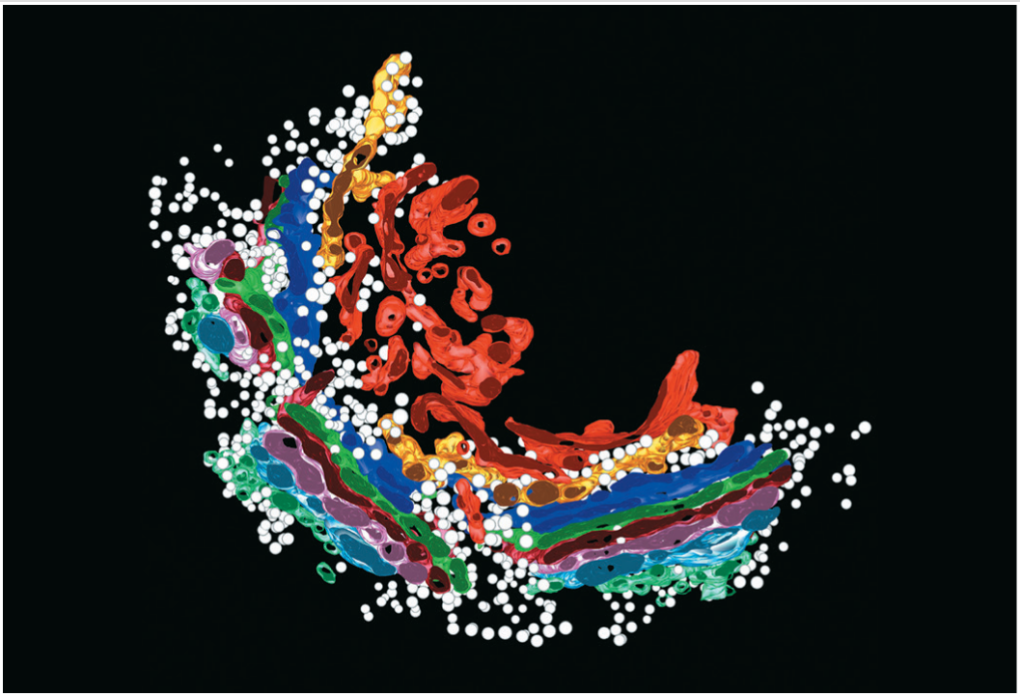
Single cells and pieces of tissue are too thick to be viewed directly in the standard transmission electron microscope. To overcome this problem, methods were developed to prepare and cut *thin sections* of cells and tissues. When these sections were examined in the electron microscope, the organization, beauty, and complexity of the cell interior was revealed and led to a revolution in cell biology — for the first time, new organelles and the first glimpses of the cytoskeleton were seen.

To prepare thin sections, it is necessary to chemically fix the sample, dehydrate it, impregnate it with a liquid plastic that hardens (similar to Plexiglas), and then cut sections of about 5 to 100 nm in thickness. For structures to be seen, the sample has to be stained with heavy metals such as uranium and lead salts, which can be done either before embedding in the plastic or after sections are cut. Examples of cells and tissues viewed by thin-section electron microscopy appear here (Figure 4-31) and throughout this book. It is important to realize that the images obtained represent just a thin slice through a cell, so to get a three-dimensional view, it is necessary to cut *serial sections* through the sample and reconstruct the sample from a series of sequential images (Figure 4-32).



Keith R. Porter Archive, University of Maryland, Baltimore County.

FIGURE 4-31 Example of a thin section viewed by transmission electron microscopy. Section through a pancreatic cell showing the extensive rough endoplasmic reticulum involved in the synthesis and secretion of digestive enzymes.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

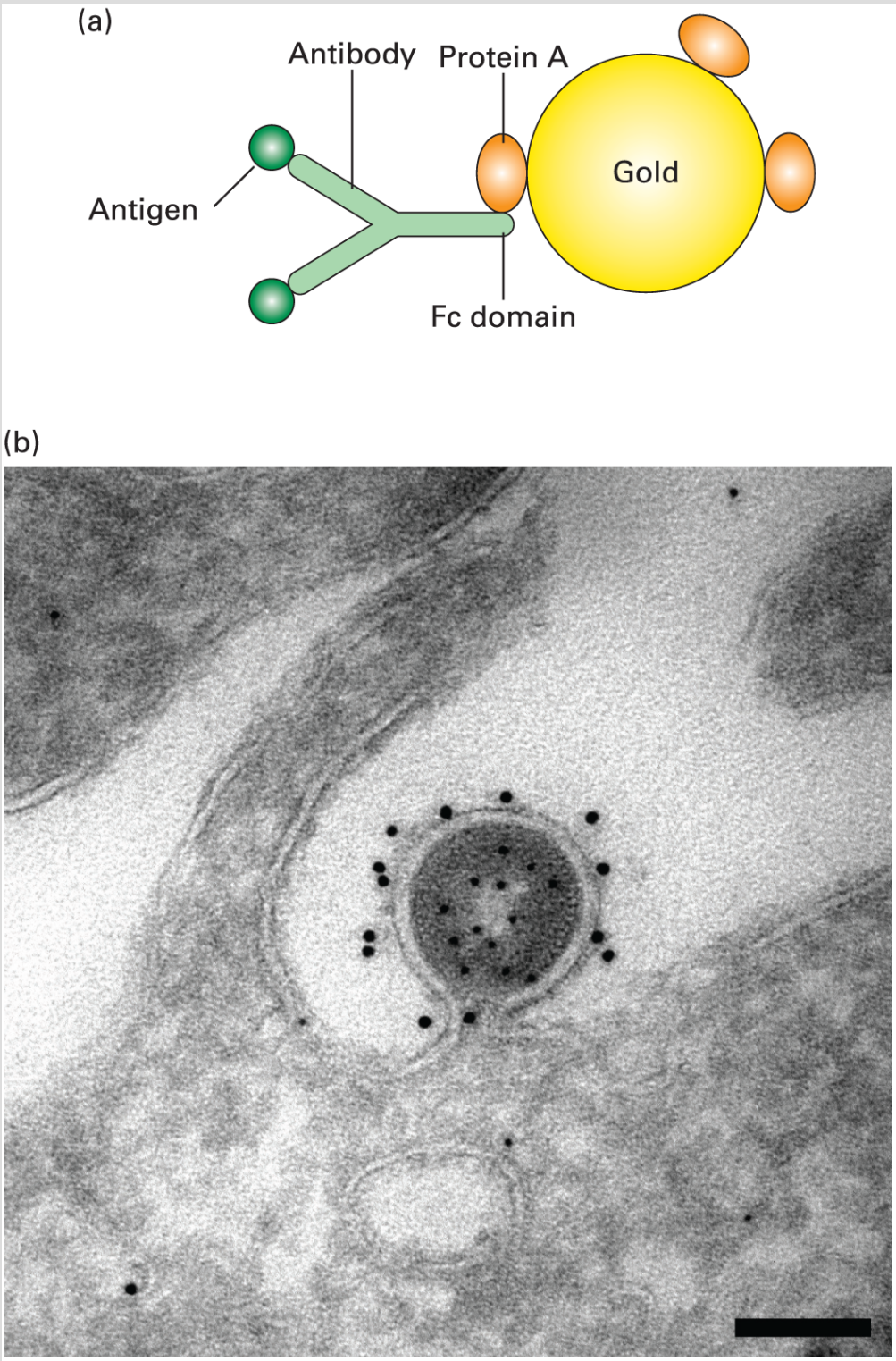
FIGURE 4-32 Model of the Golgi complex based on three-dimensional reconstruction of electron microscopy images. Transport vesicles (white spheres) that have budded off the rough ER fuse with the *cis* membranes (light blue) of the Golgi complex. By mechanisms described in Chapter 14, proteins move from the *cis* region to the *medial* region and finally to the *trans* region of the Golgi complex. Eventually, vesicles that bud off the *trans*-Golgi membranes (orange and red) move either to the cell surface or to lysosomes. The Golgi complex, like the rough endoplasmic reticulum, is especially prominent in secretory cells.

[Reprinted by permission from Nature Publishing Group, from: Marsh, B.J. & Howell, K.E., "The mammalian Golgi–complex debates," *Nat. Rev. Mol. Cell Biol*., 2002, **3**:789–785; permission conveyed through the Copyright Clearance Center, Inc.]

Immunoelectron Microscopy Localizes Proteins at the Ultrastructural Level

Just as immunofluorescence microscopy is used for localizing proteins at the light-microscope level, methods have been developed to use antibodies to localize proteins in thin sections at the electron microscope level. However, the harsh procedures used to prepare traditional thin sections — chemical fixation and embedding in plastic — can denature or modify the antigens so that they are no longer recognized by specific antibodies. Gentler methods have been developed, such as light fixation, sectioning of material after freezing at the temperature of liquid nitrogen, and finally incubation with antibody at room temperature. To make the antibody visible in the electron microscope, it must be attached to

an electron-dense marker. One way to do this is to use electron-dense gold particles coated with protein A, a bacterial protein that binds the Fc segment of all antibody molecules (Figure 4-33). Because the gold particles diffract incident electrons, they appear as dark spots.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Annegret Pelchen-Matthews and Mark Marsh, MRC-Laboratory for Molecular Cell Biology, University College London.

FIGURE 4-33 Gold particles coated with protein A are used to detect an antibody-bound protein by transmission electron microscopy. (a) First, antibodies are allowed to interact with their specific antigen in a section of

fixed tissue. Then the section is treated with electron-dense gold particles coated with protein A from the bacterium *Staphylococcus aureus*. Binding of the bound protein A to the Fc domains of the antibody molecules makes the location of the target protein visible in the electron microscope. (b) HIV particle budding from an infected HeLa cell. A cryosection of the specimen was prepared and first incubated with an antibody to capsid protein, then with protein A–coated 5-nm gold particles to localize the internal capsid protein. The unoccupied sites in the protein A were inactivated, and the specimen was incubated with antibody to the membrane-bound Env protein, followed by protein A–coated 10-nm gold particles. The distinct localization of the 5-nm gold labeling the capsid protein and the 10-nm gold labeling the Env protein can be seen. Scale bar is 100 nm.

Cryoelectron Microscopy Allows Visualization of Specimens Without Fixation or Staining

Standard transmission electron microscopy cannot be used to study live cells, and the absence of water in samples causes macromolecules to become denatured and nonfunctional. However, hydrated, unfixed, and unstained biological specimens can be viewed directly in a transmission electron microscope if the samples are frozen. In **cryoelectron microscopy**, an aqueous suspension of a sample is applied to a grid in an extremely thin film, frozen in liquid nitrogen, and maintained in this state by means of a special mount. The frozen sample is then placed in the electron microscope. The very low temperature ($-196\ ^\circ\mathrm{C}$) keeps water from evaporating, even in a vacuum. Thus the sample can be observed in

detail in its native, hydrated state without fixing or heavy metal staining. By computer-based averaging of hundreds of images, a three-dimensional model can be generated almost to atomic resolution. For example, this method was used to generate models of ribosomes and the muscle calcium pump discussed in Chapter 11. Over the last few years, techniques have been developed to use cryoelectron microscopy to determine the atomic structures of large proteins and protein complexes. This technique, described in detail in Chapter 3, involves taking up to a million images of different orientations of the molecules embedded in ice and then using computational analysis to generate a detailed, three-dimensional structure. This approach has revolutionized structural biology because it does not require the formation of protein crystals, as is required for x-ray crystallography.

An extension of this technique, *cryoelectron tomography*, allows researchers to determine the three-dimensional architecture of organelles or even whole cells embedded in ice, that is, in a state close to life. A single picture is a two-dimensional representation of a structure that lacks depth information. However, looking at the same structure from different angles gives us a three-dimensional perspective. In cryoelectron tomography, the specimen holder is tilted in small increments around the axis perpendicular to the electron beam; thus images of the object viewed from different directions are obtained (Figure 4-34a, b). The images are then merged computationally into a three-dimensional reconstruction

termed a *tomogram* (Figure 4-34c, d). A disadvantage of cryoelectron tomography is that the samples must be relatively thin, about 200 nm; this is much thinner than the samples (200-μm thick) that can be studied by confocal light microscopy.

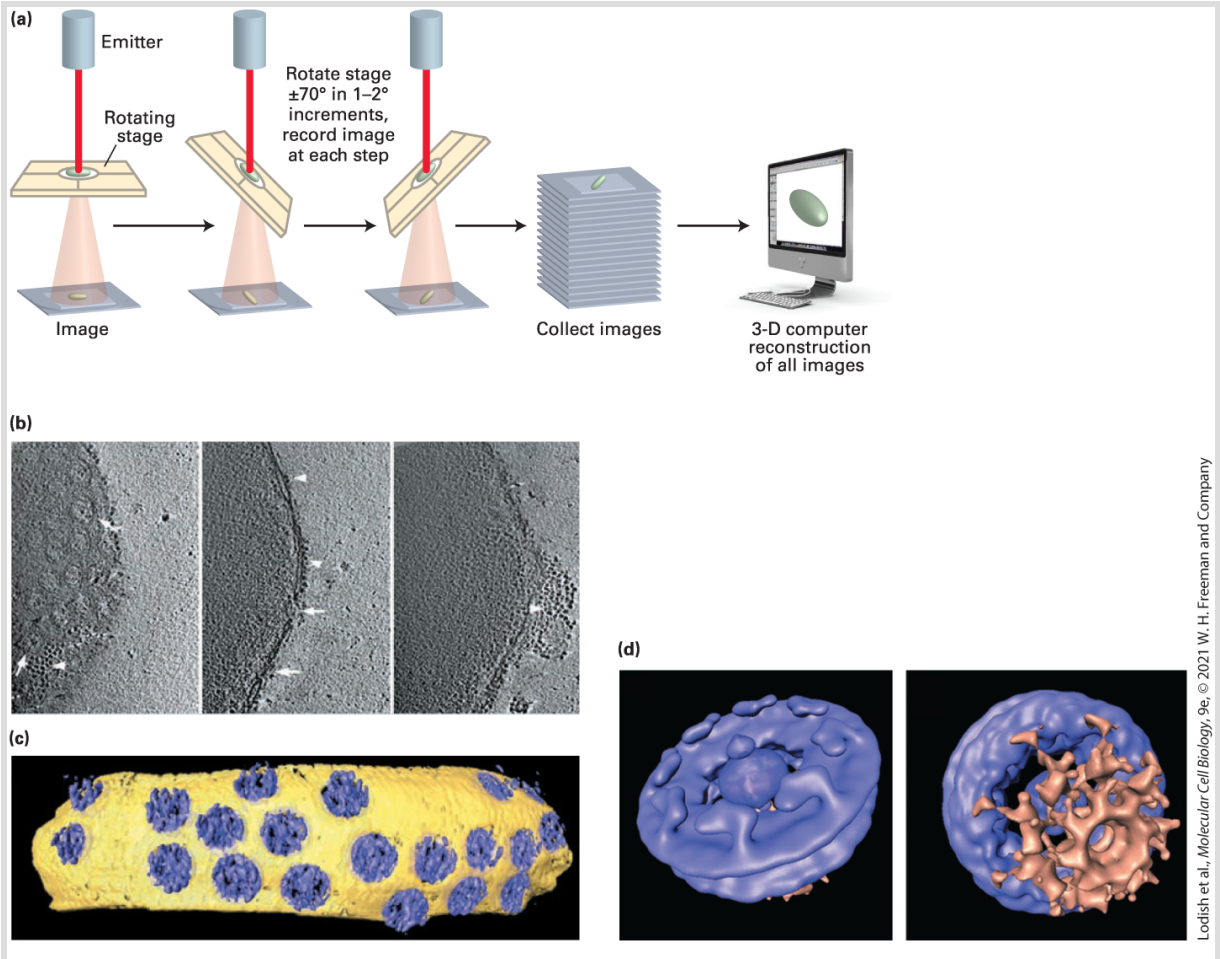


FIGURE 4-34 Structure of the nuclear pore complex (NPC) imaged by cryoelectron tomography. (a) In cryoelectron tomography, a semicircular series of two-dimensional projection images is recorded from the three-dimensional specimen that is located at the center; the specimen is tilted while the electron optics and detector remain stationary. The three-dimensional structure is then computed from the collected two-dimensional images. (b) Isolated nuclei from the cellular slime mold *Dictyostelium discoideum* were quick-frozen in liquid nitrogen and maintained in this state as the sample was observed in the electron microscope. The panel shows three sequential, tilted images. Different orientations of NPCs

(arrows) are shown in top view (*left* and *center*) and side view (*right*). Ribosomes connected to the outer nuclear membrane are visible, as is a patch of rough ER (arrowheads). (c) Computer-generated surface-rendered representation of a segment of the nuclear envelope membrane (yellow) studded with NPCs (blue). (d) By averaging the images of multiple nuclear pores, much more detail can be discerned. See S. Nickell et al., 2006, *Nat. Rev. Mol. Cell Biol.* **7**:225.

[Republished with permission of American Association for the Advancement of Science, from M. Beck. et. al. 2004. "Nuclear pore complex structure and dynamics revealed by cryoelectron tomography," *Science* **306**: 1387–1390; permission conveyed through the Copyright Clearance Center, Inc.]

Scanning Electron Microscopy of Metal-Coated Specimens Reveals Surface Features

Scanning electron microscopy (SEM) allows investigators to view the surfaces of unsectioned metal-coated specimens. An intense electron beam inside the microscope scans rapidly over the sample. Molecules in the coating are excited and release secondary electrons that are focused onto a scintillation detector; the resulting signal is displayed on a cathode-ray tube much like a conventional television (see Figure 4-28, *right*). The resulting scanning electron micrograph has a three-dimensional appearance because the number of secondary electrons produced by any one point on the sample depends on the angle of the electron beam in relation to the surface (Figure 4-35). The resolving power of scanning electron microscopes, which is limited by the thickness of the metal coating,

is only about 10 nm, much less than that of transmission instruments.

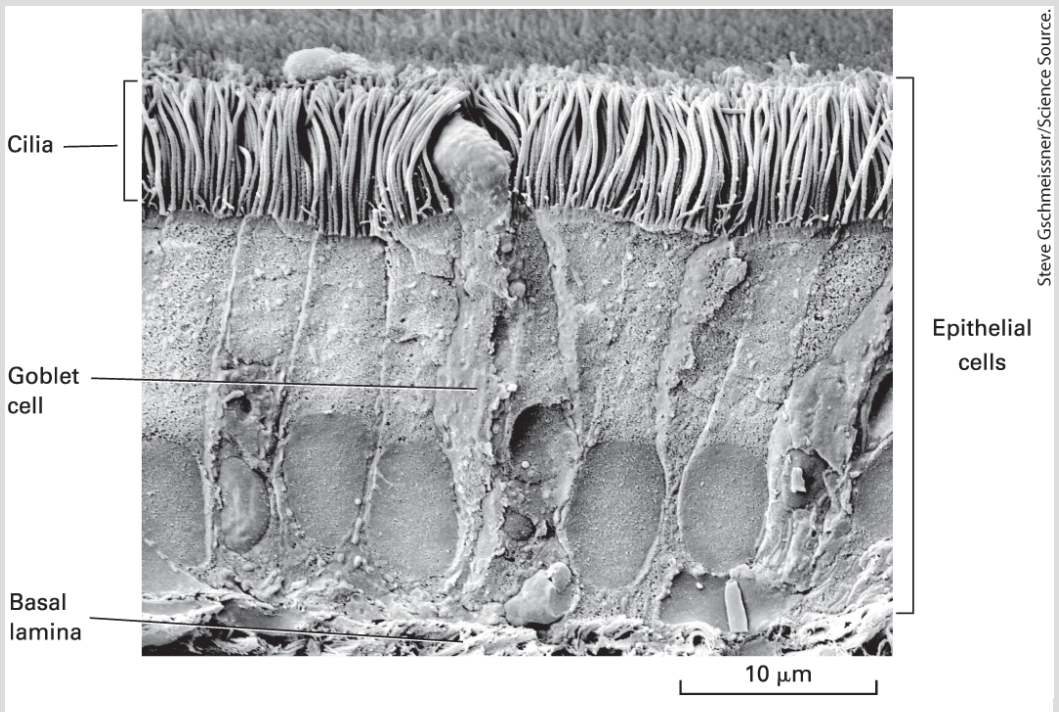


FIGURE 4-35 Scanning electron microscopy (SEM) produces a three-dimensional image of the surface of an unsectioned specimen. Seen here is an SEM image of cells of the trachea. In the middle is a goblet cell, which secretes mucus. On either side of the goblet cell are epithelial cells with abundant cilia on their apical surfaces.

KEY CONCEPTS OF SECTION 4.3

Electron Microscopy: High-Resolution Imaging

- Electron microscopy provides very high-resolution images because of the short wavelength of the high-energy electrons used to image the sample.
- Simple specimens, such as proteins or viruses, can be negatively stained or shadowed with heavy metals for examination in a transmission electron microscope (TEM).

- Thicker sections generally must be fixed, dehydrated, embedded in plastic, sectioned, and then stained with electron-dense heavy metals before viewing by TEM.
- Specific proteins can be localized by TEM by employing specific antibodies associated with a heavy metal marker, such as small gold particles.
- Cryoelectron microscopy allows examination of hydrated, unfixed, and unstained biological specimens in the TEM by maintaining them at very low temperatures.
- Scanning electron microscopy (SEM) of metal-shadowed material reveals the surface features of specimens.

4.4 Isolation of Cell Organelles

The examination of cells by light and electron microscopy led to the appreciation that eukaryotic cells contain a common set of organelles, introduced in Chapter 1 (see Figure 1-13a). However, observing organelles and documenting their detailed structure by microscopy does not clearly reveal the roles they play and how they work. For this, it is necessary to isolate organelles in their native state and identify and dissect the function of each component. For this reason, methods to isolate and characterize organelles were developed in parallel with advances in microscopy. Lysosomes, for example, are organelles in which biological molecules are degraded, as described in Chapter 1. Lysosomes had been seen by microscopy, but their function was discovered only after a method was developed to isolate them. When a method is developed to purify a type of organelle, it is possible to begin to catalogue all of its components and probe the function of each one. In another example, electron microscopy revealed that some ribosomes are associated with the endoplasmic reticulum, suggesting that the ER is a site of protein synthesis. We now know that proteins to be secreted from the cell are made by these ribosomes and that the nascent secretory protein is translocated across the membrane into the lumen of the endoplasmic reticulum. As we describe in Chapter 13, understanding the mechanism by which this occurs depended on the isolation of the endoplasmic reticulum and the development of in vitro assays for the synthesis and translocation of secretory

proteins. Thus before one can begin to fully understand organelles, biochemical assays need to be established to probe the functions of each component, with the eventual goal of reconstituting functional organelles from purified components.

Most organelles are enclosed in a lipid bilayer and perform a specific function. Each type of organelle has a recognizable structure and contains a specific set of proteins to perform its function. Cell biologists use this fact to identify specific organelles. For example, as discussed in Chapter 12, most of the ATP in a cell is made by ATP synthase, which converts ADP to ATP in mitochondria, so ATP synthase is a good marker for mitochondria. As we will discuss below, the availability of specific markers for organelles has helped in the development of organelle purification.

In this section, we discuss methods that are used to open up cells for the purification of organelles. We end with recent advances in proteomics aimed at defining the complete protein inventories of organelles.

Disruption of Cells Releases Their Organelles and Other Contents

The initial step in purifying subcellular structures is to release the cell's contents by rupturing the plasma membrane and the cell wall, if present. To do this, the cells are suspended in a solution of

appropriate pH and salt content, usually isotonic sucrose (0.25 M) or a combination of salts similar in composition to those in the cell's interior. Many cells can then be broken by stirring the cell suspension in a high-speed blender or by exposing it to ultrahigh-frequency sound (*sonication*). Alternatively, plasma membranes can be sheared by special pressurized tissue homogenizers in which cells are forced through a very narrow space between a plunger and the wall of the vessel; the pressure of being forced between the vessel wall and the plunger ruptures the cell.

Recall that water flows into cells when they are placed in a hypotonic solution; that is, one with a lower concentration of ions and small molecules than is found inside the cell. This osmotic flow can be used to cause cells to swell, weakening the plasma membrane and facilitating its rupture. Generally the cell solution is best kept at 0 °C to preserve enzymes and other constituents after their release from the stabilizing forces of the cell.

Disrupting the cell produces a mix of suspended cellular components, the *homogenate*, from which the desired organelles can be retrieved. Because rat liver contains an abundance of a single cell type, this tissue has been used in many classic studies of cell organelles. However, the same isolation principles apply to virtually all cells and tissues, and modifications of these cell-fractionation techniques can be used to separate and purify any desired components.

Centrifugation Can Separate Many Types of Organelles

In Chapter 3, we considered the principles of centrifugation and the uses of centrifugation techniques for separating proteins and nucleic acids. Since different organelles have characteristic densities because of different membrane, protein, and nucleic acid composition, similar approaches can be used for separating them. Organelles differ in both size and density and thus undergo sedimentation at different rates.

Most cell-fractionation procedures begin with *differential centrifugation* of a filtered cell homogenate at increasingly higher speeds (Figure 4-36). After centrifugation at each speed for an appropriate time, the liquid that remains at the top of the vessel, called the *supernatant*, is poured off and centrifuged at higher speed. The pelleted fractions obtained by differential centrifugation generally contain a mixture of organelles, although nuclei and viral particles can sometimes be purified completely by this procedure.

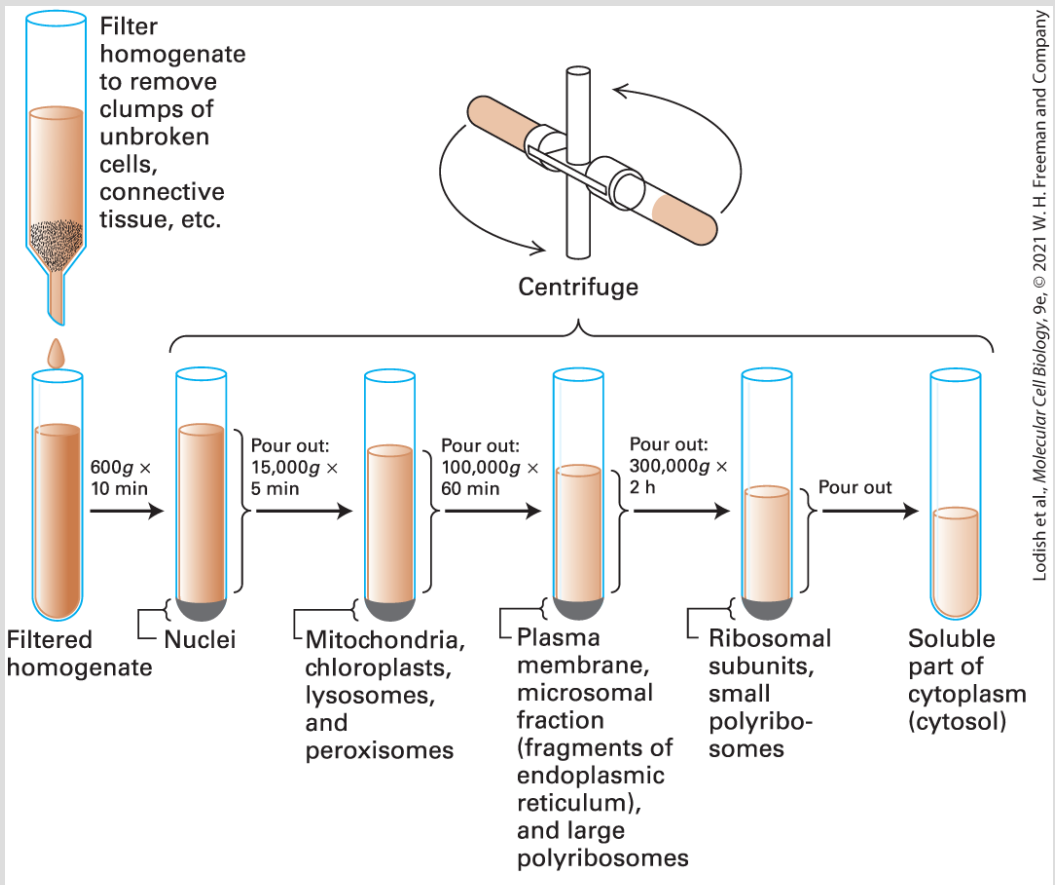


FIGURE 4-36 Differential centrifugation is a common first step in fractionating a cell homogenate. The homogenate that results from disrupting cells is usually filtered to remove unbroken cells and then centrifuged at a fairly low speed to selectively pellet the nuclei — the largest organelle. The undeposited material (the supernatant) is next centrifuged at a higher speed to sediment the mitochondria, chloroplasts, lysosomes, and peroxisomes. Subsequent centrifugation in an ultracentrifuge at 100,000*g* for 60 minutes results in deposition of the plasma membrane, fragments of the endoplasmic reticulum, and large polyribosomes. The recovery of ribosomal subunits, small polyribosomes, and particles such as complexes of enzymes requires additional centrifugation at still higher speeds. Only the cytosol — the soluble aqueous part of the cytoplasm — remains in the supernatant after centrifugation at 300,000*g* for 2 hours.

An impure organelle fraction obtained by differential centrifugation can be further purified by *equilibrium density-gradient centrifugation*, which separates cellular components according to their density. After the fraction is resuspended, it is layered on top of a solution that contains a gradient of a dense non-ionic substance (e.g., sucrose or glycerol). The tube is centrifuged at a high speed (about 40,000 rpm) for several hours, allowing each particle to migrate to an equilibrium position where the density of the surrounding liquid is equal to the density of the particle (Figure 4-37). The different layers of the gradient are then recovered by pumping out the contents of the centrifuge tube through a narrow piece of tubing and collecting the fractions (see Classic Experiment 4-1).

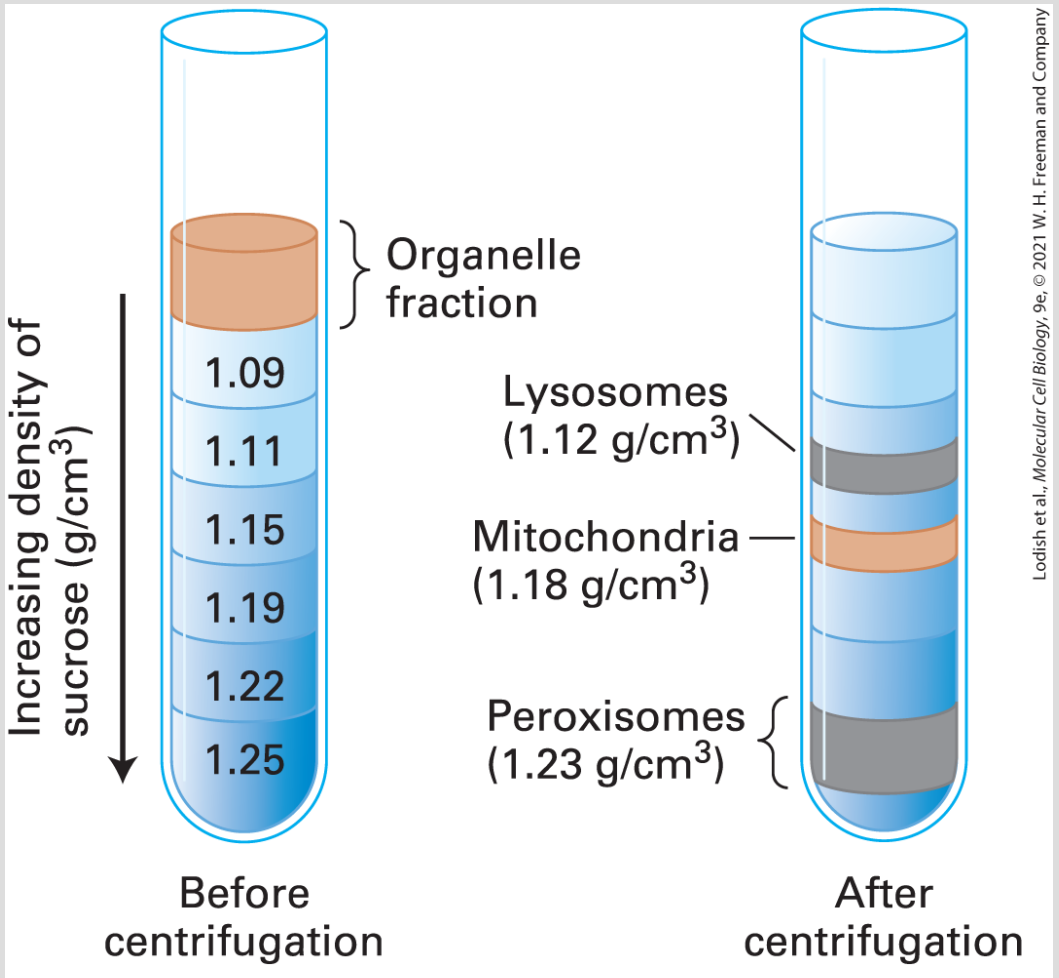


FIGURE 4-37 A mixed-organelle fraction can be further separated by equilibrium density-gradient centrifugation. In this example, using rat liver, material in the pellet from centrifugation at 15,000*g* (see Figure 4-36) is resuspended and layered on a gradient of increasingly dense sucrose solutions in a centrifuge tube. During centrifugation for several hours, each organelle migrates to its appropriate equilibrium density and remains there. To obtain a good separation of lysosomes from mitochondria, the liver is perfused with a solution containing a small amount of detergent before the tissue is disrupted. During this perfusion period, detergent is taken into the cells by endocytosis and transferred to the lysosomes, making them less dense than they would normally be and permitting a clean separation of lysosomes from mitochondria.

Because each organelle has unique morphological features, the purity of organelle preparations can be assessed by examination in an electron microscope. Alternatively, organelle-specific marker molecules can be quantified. For example, the protein cytochrome *c* is present only in mitochondria, so the presence of this protein in a fraction of lysosomes would indicate its contamination by mitochondria. Similarly, catalase is present only in peroxisomes; acid phosphatase, only in lysosomes; and ribosomes, only in the rough endoplasmic reticulum or the cytosol.

Organelle-Specific Antibodies Are Useful in Preparing Highly Purified Organelles

Cell fractions remaining after differential and equilibrium density-gradient centrifugation usually contain more than one type of organelle. Monoclonal antibodies to various organelle-specific membrane proteins are a powerful tool for further purifying such fractions. One example is the purification of vesicles whose outer surface is covered with the protein **clathrin**; these coated vesicles are derived from coated pits at the plasma membrane during receptor-mediated endocytosis, a topic we will discuss in detail in Chapter 14. An antibody to clathrin, bound to a dead bacterial cell that expressed protein A on its surface, can selectively bind these vesicles in a crude preparation of membranes, and the whole

antibody complex can then be isolated by low-speed centrifugation (Figure 4-38). A related technique uses tiny metallic beads coated with specific antibodies. Organelles that bind to the antibodies, and are thus linked to the metallic beads, are recovered from the preparation by adhesion to a small magnet on the side of the test tube.

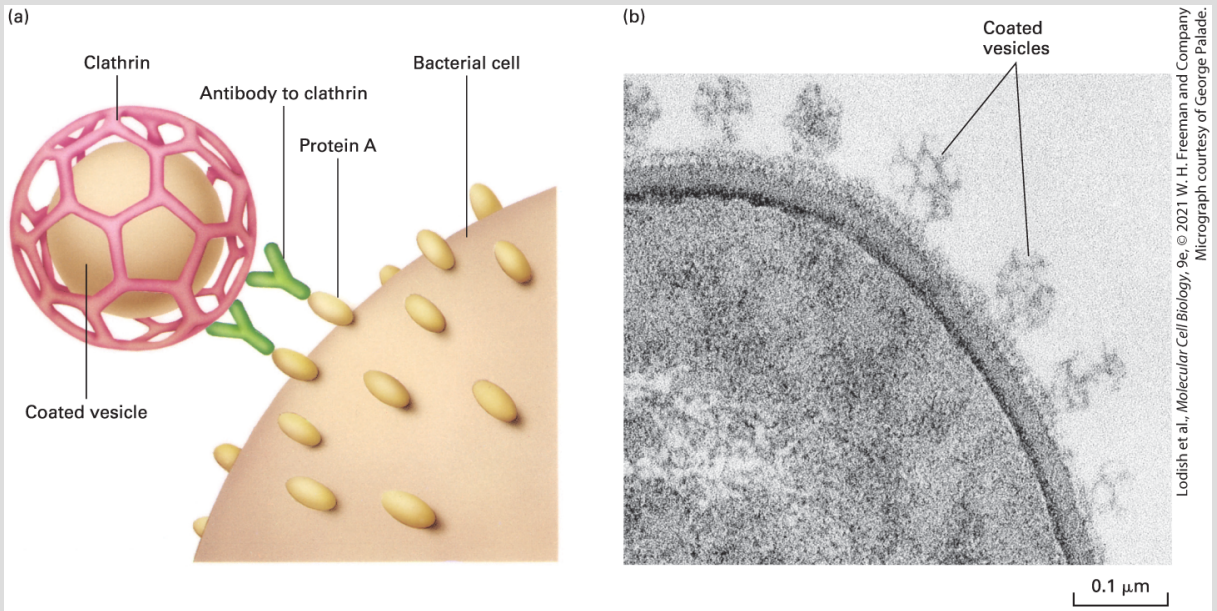


FIGURE 4-38 Coated vesicles can be purified by binding to an antibody specific for a vesicle surface protein and linkage to bacterial cells. In this example, a suspension of membranes from rat liver is incubated with an antibody specific for clathrin, a protein that coats the outer surface of certain cytosolic vesicles. To this mixture is added a suspension of killed *Staphylococcus aureus* bacteria, whose surface membrane contains protein A, which binds to the constant (Fc) region of antibodies. (a) Interaction of protein A with antibodies bound to clathrin-coated vesicles links the vesicles to the bacterial cells. The vesicle-bacteria complexes can then be recovered by low-speed centrifugation. (b) A thin-section electron micrograph reveals clathrin-coated vesicles bound to an *S. aureus* cell. See E. Merisko et al., 1982, *J. Cell Biol.* **93**:846.

All cells contain a dozen or more different types of small, membrane-limited vesicles of about the same size (50–100 nm in diameter) and density, which makes them difficult to separate from one another by centrifugation techniques. Immunological techniques are particularly useful for purifying specific classes of such vesicles. Fat and muscle cells, for instance, contain a particular glucose transporter (GLUT4) that is localized to the membrane of one of these vesicle types. When insulin is added to the cells, these vesicles fuse with the plasma membrane and increase the number of glucose transporters able to take up glucose from the blood. As we will see in Chapter 15, this process is critical for maintaining the appropriate concentration of sugar in the blood. The GLUT4-containing vesicles can be purified by using an antibody that binds to a segment of the GLUT4 protein that faces the cytosol. Likewise, the various transport vesicles discussed in Chapter 14 are characterized by unique surface proteins that permit their separation with the aid of specific antibodies.

A variation of this technique is employed when no antibody specific for the organelle under study is available. A gene encoding an organelle-specific membrane protein is modified by the addition of a segment encoding an epitope tag; the tag is placed on a segment of the protein that faces the cytosol. Following stable expression of the recombinant protein in the cell under study, an anti-epitope monoclonal antibody (described above) can be used to rapidly purify the organelle following disruption of the cell.

Proteomics Reveals the Protein Composition of Organelles

We introduced this section by emphasizing how important it is to isolate organelles to identify their components. For many years, the technology was not sufficiently sophisticated to generate a complete inventory of the proteins in each organelle, but recent advances in genomics and mass spectrometry have now made it possible. This approach combines organelle isolation with the proteomics techniques discussed in Chapter 3.

Identifying all the proteins in an organelle requires three steps. First, one has to be able to isolate the organelle to a high degree of purity. Second, one has to have a way to identify all the sequences of the proteins in the organelle. This identification is generally done by digesting all the proteins with a protease such as trypsin, which cleaves all polypeptides at lysine and arginine residues, and then determining the mass and sequence of all the resulting peptides by mass spectrometry. Third, one has to have a genomic sequence to identify the proteins from which all the peptides came. In this way, the proteomes of many organelles have been determined. As one example, a recent proteomic study on mitochondria purified from mouse brain, heart, kidney, and liver revealed 591 mitochondrial proteins, including 163 proteins not previously known to be associated with this organelle. Several proteins were found in mitochondria only from specific cell types. Determining the

functions associated with these newly identified mitochondrial proteins is a major objective of current research on this organelle.

KEY CONCEPTS OF SECTION 4.4

Isolation of Cell Organelles

- Microscopy has revealed a common set of organelles that are present in eukaryotic cells.
- Disruption of cells by vigorous homogenization, sonication, or other techniques releases their organelles. Swelling of cells in a hypotonic solution weakens the plasma membrane, making it easier to rupture.
- Sequential differential centrifugation of a cell homogenate yields fractions of partly purified organelles that differ in mass and density.
- Equilibrium density-gradient centrifugation, which separates cellular components according to their densities, can further purify cell fractions obtained by differential centrifugation.
- Immunological techniques using antibodies against organelle-specific membrane proteins are particularly useful in purifying organelles and vesicles of similar sizes and densities.
- Proteomic analysis can identify all the protein components in a purified organelle, and differences in organelle composition between tissues.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Analyze the Data
- Classic Experiment 4-1: Separating Organelles
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

cell-adhesion molecules (CAMs)
cell line
cell strain
clathrin
clone
cryoelectron microscopy
epithelium
epitope
fibroblast
fluorescence-activated cell sorter (FACS)
fluorescent staining

Förster resonance energy transfer (FRET)
hybridoma
monoclonal antibody
polyclonal antibody
resolution
scanning electron microscopy (SEM)
transmission electron microscopy (TEM)

Review the Concepts

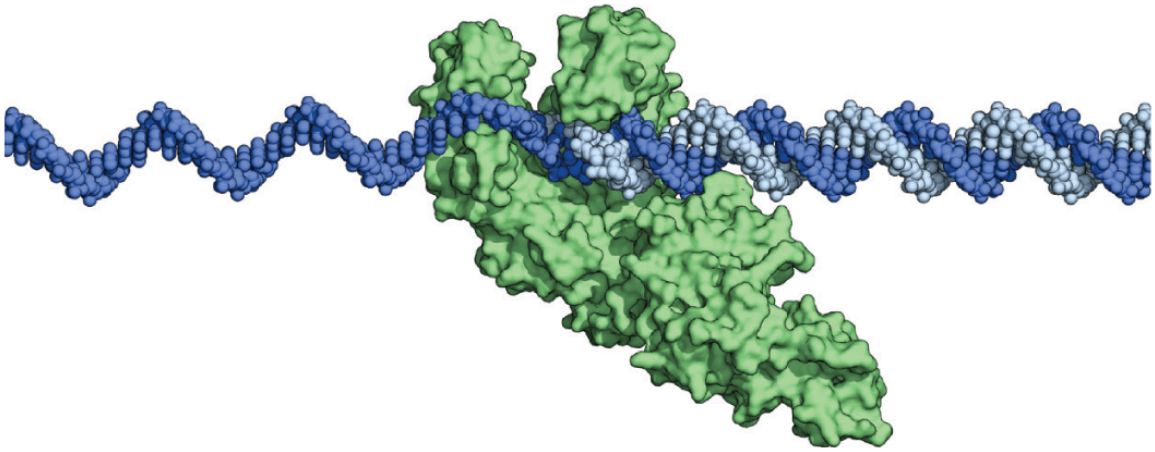
1. Both light and electron microscopy are commonly used to visualize cells, cell structures, and the location of specific molecules. Explain why a scientist may choose one or the other microscopy technique for use in research.
2. The magnification possible with any type of microscope is an important property, but its resolution, the ability to distinguish between two very closely apposed objects, is even more critical. Describe why the resolving power of a microscope is more important for seeing finer details than its magnification. What is the formula used to describe the resolution of a microscope lens, and what are the limitations placed on the values in the formula?
3. Why are chemical stains required for visualizing cells and tissues with the basic light microscope? What advantage do fluorescent dyes and fluorescence microscopy provide in comparison to the chemical dyes used to stain specimens for light microscopy? What advantages do confocal and

deconvolution microscopy provide in comparison to conventional fluorescence microscopy?

4. In certain electron microscopy methods, the specimen is not directly imaged. How do these methods provide information about cellular structure, and what types of structures do they visualize? What limitation applies to most forms of electron microscopy?
5. What is the difference between a cell strain, a cell line, and a clone?
6. Explain why the process of cell fusion is necessary to produce monoclonal antibodies used for research.
7. Much of what we know about cell function depends on experiments using specific cells and specific parts (e.g., organelles) of cells. What techniques do scientists commonly use to isolate cells and organelles from complex mixtures, and how do these techniques work?
8. Hoechst 33258 is a chemical dye that binds specifically to DNA in live cells, and when excited by UV light, it fluoresces in the visible spectrum. Name and describe one specific method, employing Hoechst 33258, an investigator would use to isolate fibroblasts in the G_2 phase of the cell cycle from those fibroblasts in interphase.

CHAPTER 5

Fundamental Molecular Genetic Mechanisms



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Model of DNA Polymerase I from *Escherichia coli* in the process of replicating DNA. The universal mechanism for copying information in DNA is replication by DNA polymerase, a remarkable copying and proofreading machine. The *E. coli* DNA polymerase I shown here can synthesize a complementary copy of the template strand at a rate of about 500 bases per second, making a mistake only once in every 10^5 bases.

OUTLINE

5.1 The Double-Helical Structure of DNA
5.2 DNA Replication
5.3 DNA Repair and Recombination
5.4 Transcription of Protein-Coding Genes and Formation of mRNA
5.5 The Decoding of mRNA by tRNAs
5.6 Stepwise Synthesis of Proteins on Ribosomes
5.7 Viruses: Parasites of the Cellular Genetic System

One of the foundational principles of cell biology is that the cells that make up a multicellular organism are autonomous — each cell carries its own set of genetic instructions for making all of the specific components of cellular structures, cellular catalysts, and molecular machines that make up a functioning cell of a given type. As was described in Chapter 3, the vast majority of these active cellular components are proteins. The specific function of each individual protein is dictated by its sequence of amino acids, the instructions for which are ultimately encoded in the sequence of nucleotides in **deoxyribonucleic acid (DNA)**, the biological information-carrying molecule. One copy of DNA of the human genome encompassing all 23 chromosomes is about 3.3×10^9 nucleotides long and encodes about 20,000 different proteins. In physical terms, one copy of the genome adds up to a thread of DNA that is about one meter in total length, but that is so fine that two complete copies are compacted into the nucleus of *each* of the approximately 4×10^{13} cells in our body. This information, under the control of a complex regulatory network, is then read out to express the appropriate sets of proteins for each of the specific cell types in the body. Thus the basic principle of autonomy of cellular information implies a system of stored and expressed genetic information that is so vast as to defy comprehension. Yet the general rules, as they are now well understood, for how biological information is managed are in fact quite straightforward and are the same in virtually all forms of life. In this chapter, we describe the basic principles for storage, replication, and expression of biological

information. A full understanding of this subject not only requires an understanding of the structure and biochemical properties of the molecules involved, but also requires the application of biological information theory.

Let us first see how information theory can be used to compute the information-carrying capacity of DNA and to compare this to the memory capacity of a computer. Whether in a digital computer or in a biological system, the smallest unit of information is a *bit*, which corresponds to a unit that can have one of two values — in a digital computer these values correspond to 1 and 0. DNA is made up of two complementary strands that can have one of four possible bases (A, G, C, or T) at each position. The information in n bits can specify 2^n possible values; thus each base can be specified by two bits and is therefore said to contain two bits of information. The two strands are base paired with one another according to the principles of base pairing, such that A always pairs with T and G always pairs with C. Because of the fixed rules for base pairing, the sequence of the second complementary strand of DNA is determined by the first and does not add to the information carrying capacity of DNA, but it does add redundancy of information which, as we will see, allows DNA to be readily copied and for damage in DNA to be repaired. Since the human genome contains a total of 3.3×10^9 base pairs, the total information capacity is 6.6×10^9 bits, which corresponds to about 8×10^8 *bytes* of information ($8 \text{ bits} = 1 \text{ byte}$, which is the unit that is used to measure computer memory). A single copy of the

human genome weighs only 3.6×10^{-12} grams and thus on a weight basis, DNA has enormous capacity to store information — in theory, the same information contained in a million one-terabyte laptop computer hard drives could be stored in a mass of DNA the size of a grain of sand.

When cells divide, each daughter cell must receive its own exact copy of the information contained in DNA. DNA is copied or replicated by separation of the two complementary strands of a DNA molecule and then enzymes known as **DNA polymerases** use each strand as a template for the synthesis of two identical copies of the parent molecule by following the rules of base pairing. From the perspective of an information-copying reaction, we are interested in both the rate at which the genome can be copied and the *fidelity* of the reaction, which is related to how frequently errors are produced during copying. From fundamental chemical principles, fidelity is dictated by the difference in energy between a correct and an incorrect base pair in the replication reaction. A typical DNA polymerase incorporates one nucleotide every 2 milliseconds. In this timeframe the theoretical rate of errors in replication would be one mistake in every 10^5 bases, which is the measured fidelity of replication catalyzed by the simplest DNA polymerases. However, if the human genome were replicated with such a low-fidelity polymerase, each parent would pass on about one million new mutations to each of their children. The measured fidelity of DNA replication in humans corresponds to one new mutation in 10^8 bases

per generation. In this chapter, we will see how elaborate proofreading and repair processes have evolved to achieve this 1000-fold enhancement in the fidelity in the replication and transmission of information in DNA.

Gene expression is the mechanism by which the information in the DNA sequence of a gene is used as the instructions for production of a functional protein or RNA molecule. For protein-coding genes, the flow of information within cells follows the outline DNA $\rightarrow$ RNA $\rightarrow$ protein, which is known as the *central dogma* of molecular biology. Before any of the details of how information flows from DNA to proteins were understood, molecular biologists used the basic principles of information theory to deduce how a DNA sequence comprised of four letters (A, G, C, and T) could encode a protein sequence made up of 20 different amino acids. In theory, at least five bits of information (which could give 32 different possibilities) would be required to specify one of 20 amino acids. For a simple code in which blocks of nucleotides in DNA would code for each amino acid, a block size of a least three nucleotides would be required. We now know that the nucleotide sequences of genes are in fact read three nucleotides at a time in blocks known as **codons**. Working out the universal **genetic code** for which amino acid is coded by each of the 64 possible codons was a monumental achievement of twentieth-century science. The genetic code table is the central component of all of the algorithms that are used to

analyze new genomic sequences to deduce the protein sequences that they encode.

The information in the nucleotide sequence of DNA is not translated directly into an amino acid sequence; instead, through the process of **transcription**, the information stored in DNA is copied into a **messenger ribonucleic acid (mRNA)**. Like DNA, RNA carries information in the form of a sequence of four nucleotides, and the process of transcription is similar to replication in that the rules of base pairing are used to synthesize an mRNA of complementary sequence from a DNA template. RNA is generally single stranded and RNA is chemically less stable than DNA. Thus mRNA serves as a temporary copy of the information in a gene that allows the expression of each gene sequence to be controlled independently.

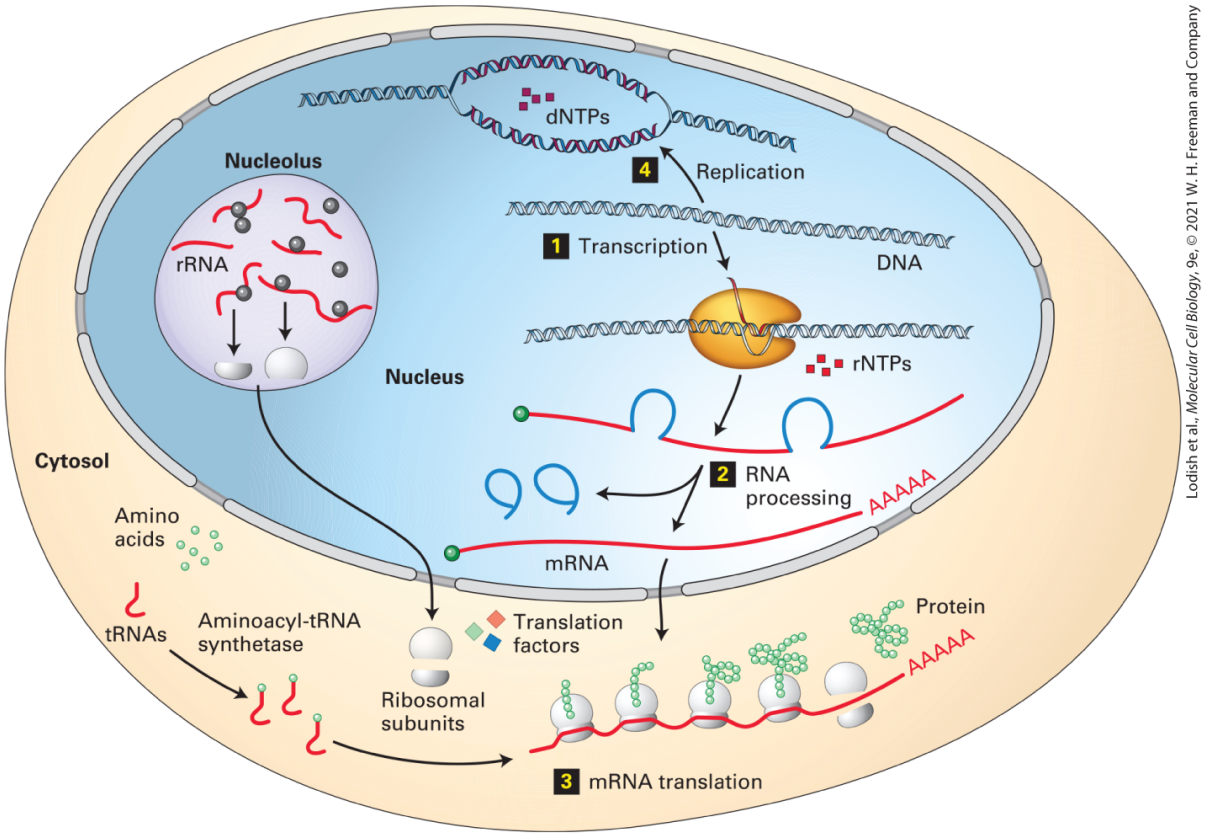
The nucleotide sequence of an mRNA molecule contains information that specifies the correct order of amino acids during the synthesis of a protein. In this process, known as **translation**, the nucleotide sequence of an mRNA molecule is *read* by a second type of RNA called **transfer RNA (tRNA)**. The reading occurs by the base pairing of each of the three nucleotide codons of the mRNA with a complementary anti-codon triplet in the tRNA. The accurate, stepwise assembly of amino acids into proteins takes place in the context of a remarkable macromolecular machine, the **ribosome**, which is composed of a third type of RNA, **ribosomal RNA (rRNA)**, and associated proteins. As the correct amino acids are brought into

sequence by tRNAs, they are linked by peptide bonds to make proteins.

Another type of information that is encoded in DNA is represented by the control sequences that regulate gene expression and determine which genes are on and which are off in each of the cell types in the body. Typically, gene expression is regulated at three possible control points: transcription initiation, mRNA stability, or initiation of translation. Regulation of transcription initiation by the control sequences at a gene promoter and the protein *transcription factors* that interact with them are discussed in Chapter 9. Other types of control sequences that govern the rate at which a given mRNA will be translated or the rate at which the mRNA will be degraded are discussed in Chapter 10.

In this chapter, we first review the chemical and structural properties of DNA that make this molecule ideally suited as the carrier of genetic information in the cell. We next consider the molecular problems involved in DNA replication and the complex cellular machinery that ensures accurate copying of the genetic material. The next section describes how damage to DNA is repaired and how regions of different DNA molecules are exchanged in the process of recombination to generate new combinations of traits in the individual organisms of a species.

In the next several sections, we discuss the basic processes summarized in Figure 5-1: transcription of DNA into mRNA precursors and processing of these precursors to make functional mRNA molecules. After outlining the functions of mRNA, tRNA, and rRNA in protein synthesis, we present a detailed description of these and other components in the biochemical steps of translation.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 5-1 Overview of four basic molecular genetic processes. In this chapter, we cover the three processes that lead to production of proteins (1–3) and the process for replicating DNA (4). During transcription of a protein-coding gene by RNA polymerase (1), the four-base DNA code specifying the amino acid sequence of a protein is copied, or *transcribed*, into a precursor messenger RNA (pre-mRNA) by the polymerization of ribonucleoside triphosphate monomers (rNTPs). Removal of noncoding sequences and other modifications to the pre-mRNA (2), collectively known as *RNA processing*, produce a functional mRNA, which is transported to the cytoplasm. During *translation* (3), the four-

base code of the mRNA is decoded into the 20–amino acid language of proteins. Transfer tRNAs (tRNAs) are the adaptors that read a codon of three nucleotides on the mRNA. In preparation for translation, tRNAs are charged with the correct amino acid by enzymes known as aminoacyl-tRNA synthetases. Ribosomes, the macromolecular machines that translate the mRNA code, are composed of two subunits assembled in the nucleolus from ribosomal RNAs (rRNAs) and multiple proteins (*left*). After transport to the cytoplasm, ribosomal subunits associate with an mRNA and carry out protein synthesis with the help of charged tRNAs and translation factor proteins. During DNA replication (**4**), which occurs only in cells preparing to divide, deoxyribonucleoside triphosphate monomers (dNTPs) are polymerized to yield two identical copies of each chromosomal DNA molecule. Each daughter cell receives one of the identical copies.

The final section of the chapter presents basic information about viruses — parasites that exploit the cellular machinery for DNA replication, transcription, and protein synthesis. In addition to being significant pathogens, viruses are important models for studying these cellular mechanisms of macromolecular synthesis and other cellular processes. Viruses continue to teach us important lessons in molecular cell biology today and have been adapted as experimental tools for introducing genes into cells, tools that are currently being tested for their effectiveness in human gene therapy.

5.1 The Double-Helical Structure of DNA

DNA and RNA have very similar primary structures: both are linear **polymers** composed of only four different nucleotides. Recall from Chapter 2 that all nucleotides consist of an organic base linked to a five-carbon sugar that has a phosphate group attached to the 5′ carbon. Though chemically similar, DNA and RNA polymers in cells have quite different three-dimensional structures. Cellular DNA exists as two long, perfectly complementary, base-paired strands that can be as long as several hundred million nucleotides. Cellular RNAs usually exist as much shorter molecules (RNAs range in length from about 20 to thousands of nucleotides); they are typically single-stranded molecules that may have limited regions that form self-complementary base pairs. As we will see, the relative chemical stability of DNA and its unique redundant complementary structure make DNA an ideal molecule for the stable storage of genetic information.

Native DNA Is a Double Helix of Complementary Antiparallel Strands

The four nucleotides of DNA contain the **purine** bases *adenine* (A) and *guanine* (G) and the **pyrimidine** bases *cytosine* (C) and *thymine* (T) (see Figure 2-17). A single DNA strand has a *backbone* composed

of repeating pentose-phosphate units from which the purine and pyrimidine bases extend as side groups. (Note that the single-letter abbreviations for these bases are also commonly used to denote all the nucleotides in nucleic acid polymers.) Like a polypeptide, a nucleic acid strand has an end-to-end chemical orientation: the $5'$ *end* has a hydroxyl or phosphate group on the $5'$ carbon of its terminal sugar; the $3'$ *end* usually has a hydroxyl group on the $3'$ carbon of its terminal sugar (Figure 5-2). This directionality, plus the fact that synthesis always proceeds $5'$ to $3'$, has given rise to the convention that polynucleotide sequences are written and read in the $5' \rightarrow 3'$ direction (from left to right); for example, the sequence ATG is assumed to be $(5')\mathrm{ATG}(3')$. The chemical linkage between adjacent nucleotides, commonly called a **phosphodiester bond**, actually consists of two phosphoester bonds, one on the $5'$ side of the phosphate and another on the $3'$ side.

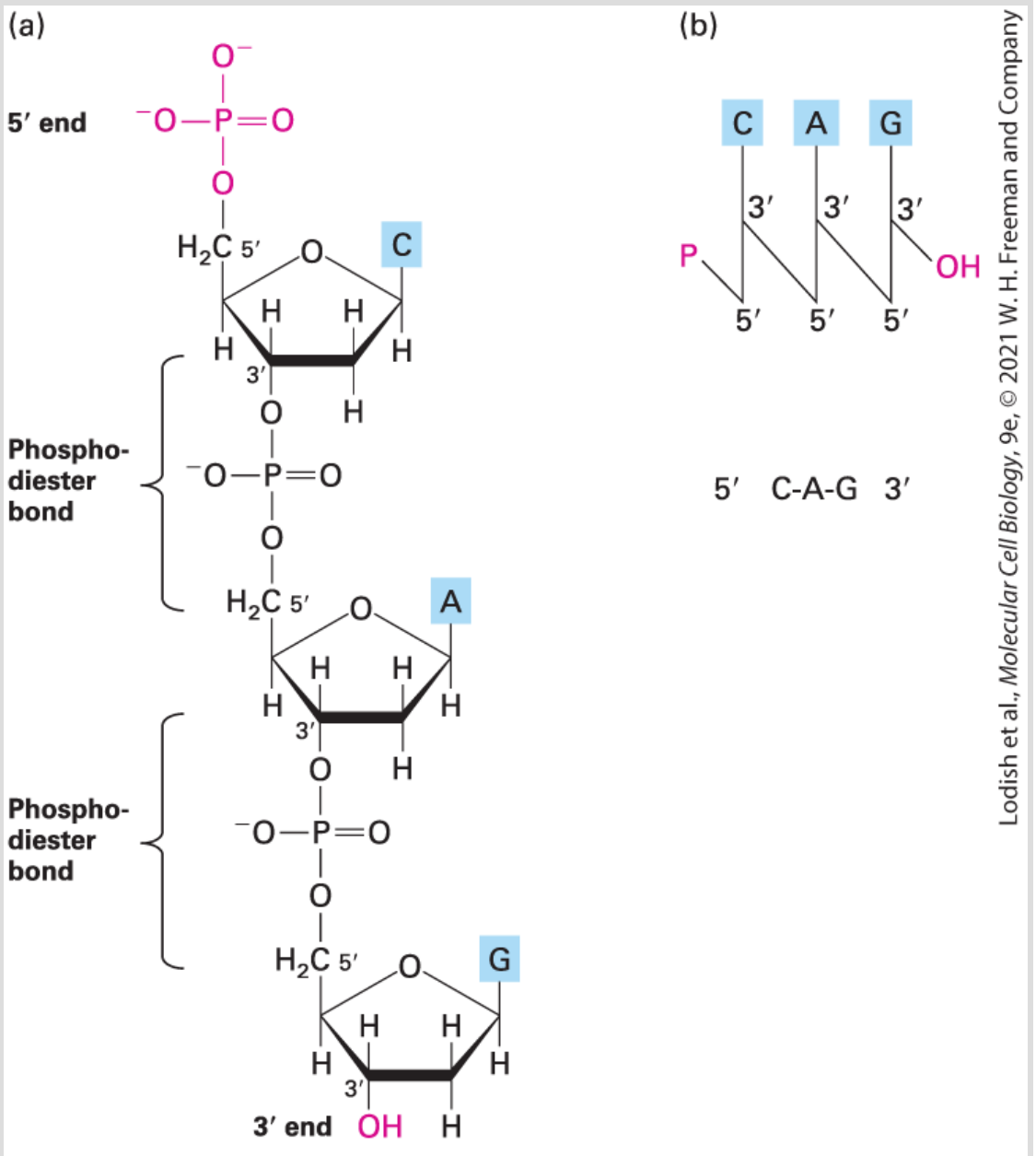


FIGURE 5-2 Chemical directionality of a nucleic acid strand. Shown here are alternative representations of a single strand of DNA containing only three bases: cytosine (C), adenine (A), and guanine (G). (a) The chemical structure shows a hydroxyl group at the 3′ end and a phosphate group at the 5′ end. Note also that two phosphoester bonds link adjacent nucleotides; this two-bond linkage is commonly referred to as a *phosphodiester bond*. (b) In the stick diagram (*top*), the sugars are indicated as vertical lines and the phosphodiester bonds as slanting lines; the bases are denoted by their single-letter abbreviations. In the simplest representation

(*bottom*), only the bases are indicated. By convention, a polynucleotide sequence is always written in the $5' \rightarrow 3'$ direction (left to right) unless otherwise indicated.

An understanding of how genetic information might be stored in the nucleotide sequence of DNA began to take shape in 1953 when James D. Watson and Francis H. C. Crick proposed that DNA has a double-helical structure. Their proposal was based on analysis of x-ray diffraction patterns of DNA fibers generated by Rosalind Franklin and Maurice Wilkins, which showed that the structure was helical, and analyses of the base composition of DNA from multiple organisms by Erwin Chargaff and colleagues. Chargaff's studies revealed that while the base composition of DNA (percentages of A, T, G, and C) varies greatly between distantly related organisms, the percentage of A always equals the percentage of T, and the percentage of G always equals the percentage of C, in all organisms. Based on these discoveries and the structures of the four nucleotides, Watson and Crick performed careful molecular model building, proposing a **double helix**, with A always hydrogen-bonded to T and G always hydrogen-bonded to C along the central axis of the double helix, as the structure of DNA.

The most telling feature of the Watson and Crick model was that each of the four possible base pairs fit within the helical backbones in precisely the same manner. Thus any arbitrary sequence of nucleotides can be accommodated by the same double-helical structure. This immediately lead to the understanding that DNA functions as an informational molecule by carrying genetic

information as a sequence of the four letters A, G, C, and T, just as a digital computer memory contains information encoded as a sequence of 1s and 0s. The extraordinary capacity of DNA to carry information stems from the fact that each nucleotide base is only about 50 atoms yet in the context of a DNA sequence can carry two bits of information — while even the most advanced computer memories require many thousands of atoms to store one bit of information.

DNA consists of two associated polynucleotide strands that wind together to form a double helix. The two sugar-phosphate backbones are on the outside of the double helix, and the bases project into the interior. The adjoining bases in each strand stack on top of one another in parallel planes (Figure 5-3a). The orientation of the two strands is *antiparallel;* that is, their $5' \rightarrow 3'$ directions are opposite. The strands are held in precise register by formation of **base pairs** between the two strands: A is paired with T through two hydrogen bonds; G is paired with C through three hydrogen bonds (Figure 5-3b). This base-pair complementarity is a consequence of the size, shape, and chemical composition of the bases. The presence of thousands of such hydrogen bonds in a DNA molecule contributes greatly to the stability of the double helix. Hydrophobic and van der Waals interactions between the stacked adjacent base pairs further stabilize the double-helical structure.

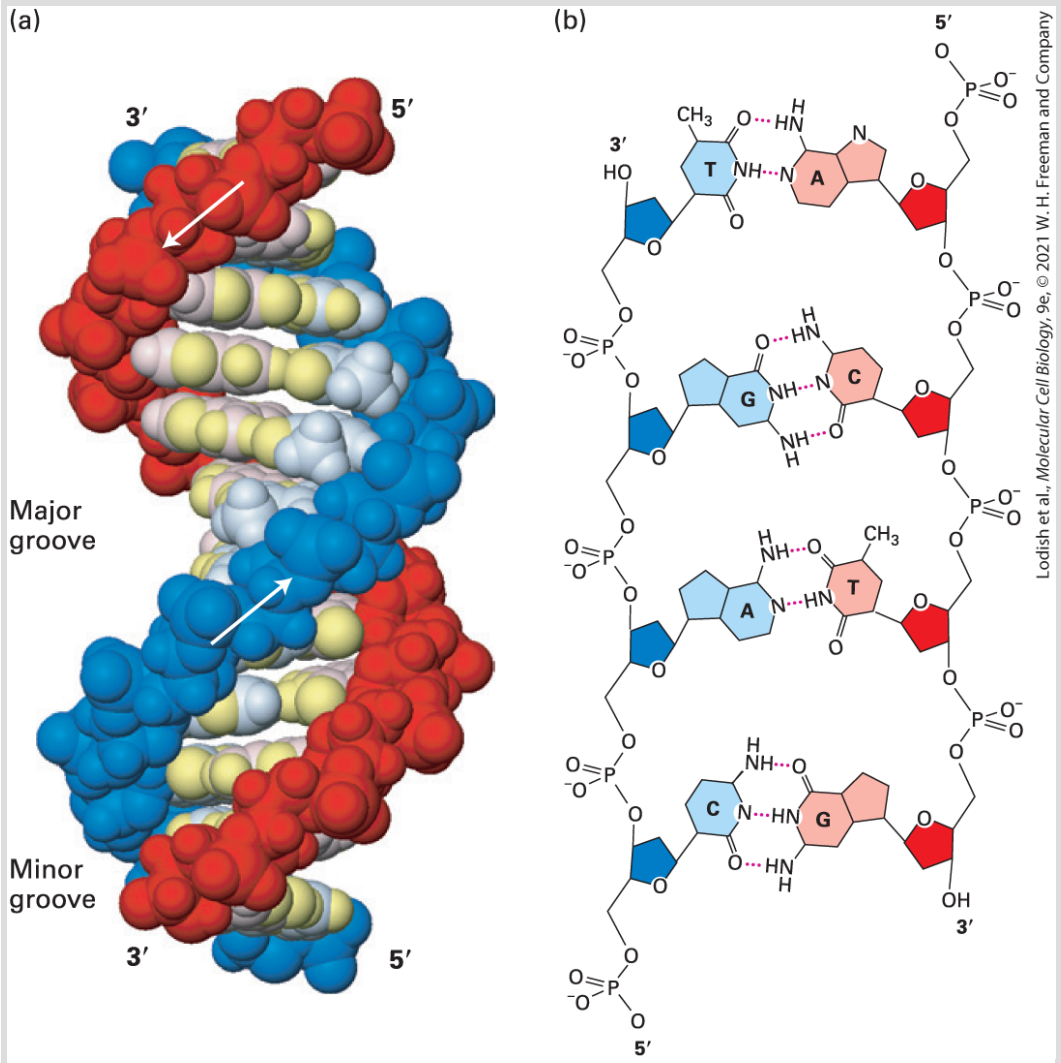


FIGURE 5-3 The DNA double helix. (a) Space-filling model of B DNA, the most common form of DNA in cells. The bases (light shades) project inward from the sugar-phosphate backbones (dark red and blue) of each strand, but their edges are accessible through major and minor grooves. Arrows indicate the $5' \rightarrow 3'$ direction of each strand. Hydrogen bonds between the bases are in the center of the structure. The major and minor grooves are lined by potential hydrogen bond donors and acceptors (highlighted in yellow). (b) Chemical structure of DNA double helix. This extended schematic shows the two sugar-phosphate backbones and hydrogen bonding between the Watson-Crick base pairs, A·T and G·C. See R. E. Dickerson, 1983, *Sci. Am*. **249**:94.

[Part (a) Data from R. Wing et al., 1980, *Nature* **287**:755, PDB ID 1bna.]

In natural DNA, A always hydrogen-bonds with T and G with C, forming A·T and G·C base pairs as shown in Figure 5-4. These associations, always between a larger purine and a smaller pyrimidine, are often called *Watson-Crick base pairs.* Two polynucleotide strands, or regions thereof, in which all the nucleotides form such base pairs are said to be **complementary**. However, in theory and in synthetic DNAs, other base pairs can form. For example, guanine (a purine) could theoretically form hydrogen bonds with thymine (a pyrimidine) within the space available in the helix. Similarly, pairing between adenosine and cytosine could also be accommodated. However, as shown in Figure 5-4, nonstandard G·T and A·C base pairs deviate sufficiently from the precise geometry of natural base pairs to be excluded from double-stranded DNA by the DNA-copying enzyme, as described later in this chapter.

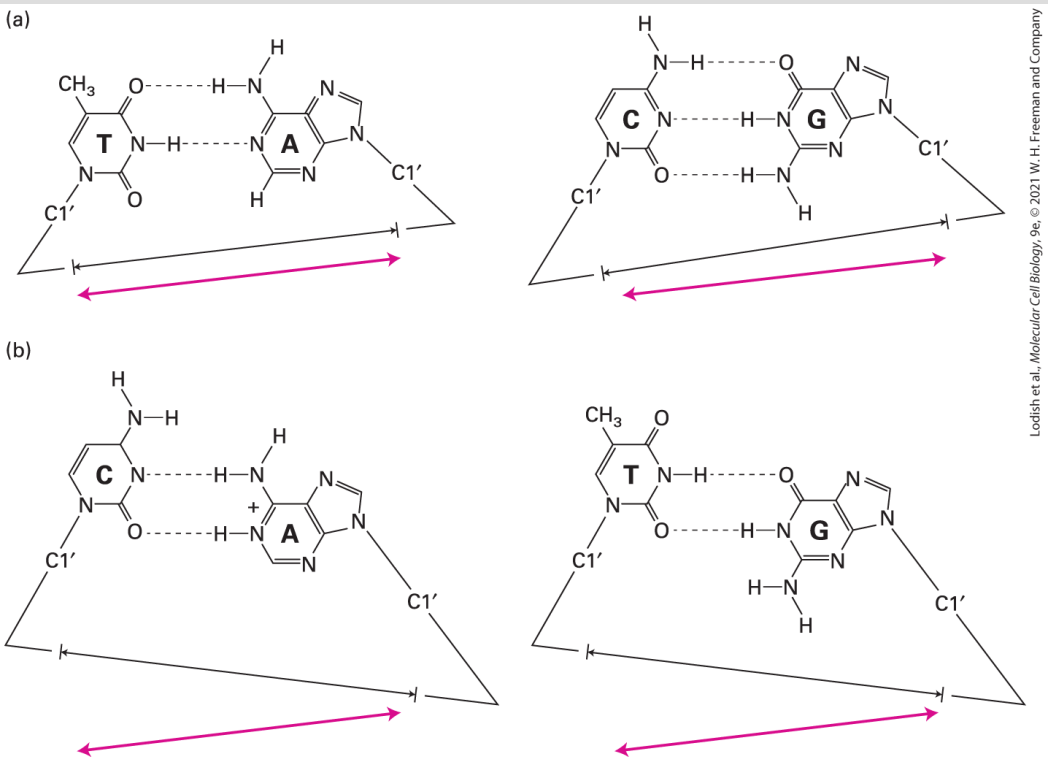


FIGURE 5-4 (a) The structure of the standard A·T and G·C base pairs showing the angle and distance between the bonds connecting the bases to the ribose-phosphate backbone. Because the geometry of the base pairs is almost identical, both of these base pairs are accommodated in the double-helical structure of DNA. (b) Nonstandard base pairs, such as G·T and A·C, have very different bond geometries and are thus excluded from natural DNA by the copying mechanism of DNA polymerase.

Almost all DNA in cells takes the form of a *right-handed* helix. The x-ray diffraction pattern of DNA indicates that the stacked bases are regularly spaced 0.34-nm apart along the helix axis. The helix makes a complete turn every 3.4 to 3.6 nm, depending on the sequence; thus there are about 10–10.5 base pairs per turn. This helical form, referred to as the *B form* of DNA, is the normal form present in most DNA stretches in cells. On the outside of the helix, the spaces

between the intertwined strands form two helical grooves of different widths, described as the *major groove* and the *minor groove* (see Figure 5-3a). As a consequence, the atoms on the edges of each base within these grooves are accessible from outside the helix, forming two types of binding surfaces. DNA-binding proteins can read the sequence of bases in duplex DNA by contacting atoms in either the major or the minor grooves.

Although the atoms facing the surfaces of the major and minor grooves differ between the four bases, the overall structure of the intertwined strands of DNA is uniform regardless of the base sequence. Any irregularities in the DNA caused by chemical damage or mispairing of bases will disrupt this uniform structure and, as we will see in Section 5.3, will allow DNA repair enzymes to identify and act on the site of the damage.

Important modifications in the structure of standard B-form DNA come about as a result of protein binding to specific DNA sequences. Although the multitude of hydrogen and hydrophobic bonds between the bases provides stability to DNA, the double helix is flexible about its long axis. Unlike the α helix in proteins (see Figure 3-4), it has no hydrogen bonds parallel to the axis of the helix. This property allows DNA to bend when complexed with a DNA-binding protein, such as the transcription factor TBP (Figure 5-5). Bending of DNA is also critical to the dense packing of DNA in chromatin, the

protein-DNA complex in which nuclear DNA occurs in eukaryotic cells (see Chapter 8).

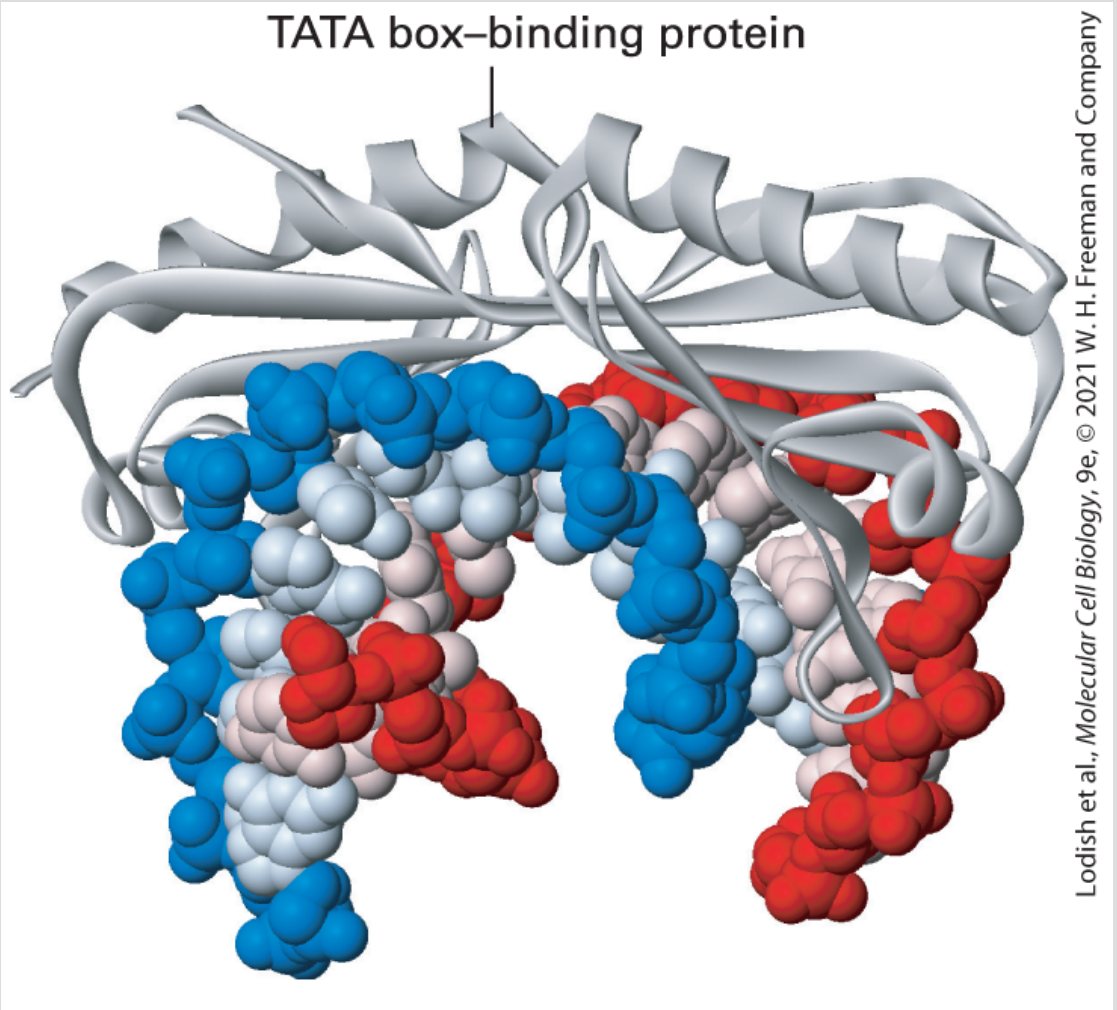


FIGURE 5-5 Interaction with a protein such as TBP can bend DNA. The conserved C-terminal domain of the TATA box–binding protein (TBP) binds to the minor groove of specific DNA sequences rich in A and T, untwisting and sharply bending the double helix. Transcription of most eukaryotic genes requires participation of TBP.

[Data from D. B. Nikolov and S. K. Burley, 1997, *Proc. Nat'l. Acad. Sci. USA* **94**:15, PDB ID 1cdw.]

Why did DNA, rather than RNA, evolve to be the carrier of genetic information in cells? The hydrogen at the 2′ position in the deoxyribose of DNA makes it a far more stable molecule than RNA, which instead has a hydroxyl group at the 2′ position of ribose (see Figure 2-16). The 2′-hydroxyl groups in RNA participate in the slow, OH^--catalyzed hydrolysis of phosphodiester bonds at neutral pH (as shown in Figure 5-6). The absence of 2′-hydroxyl groups in DNA prevents this process. Therefore, the presence of deoxyribose in DNA makes it a more stable molecule — a characteristic that is critical to its function in the long-term storage of genetic information.

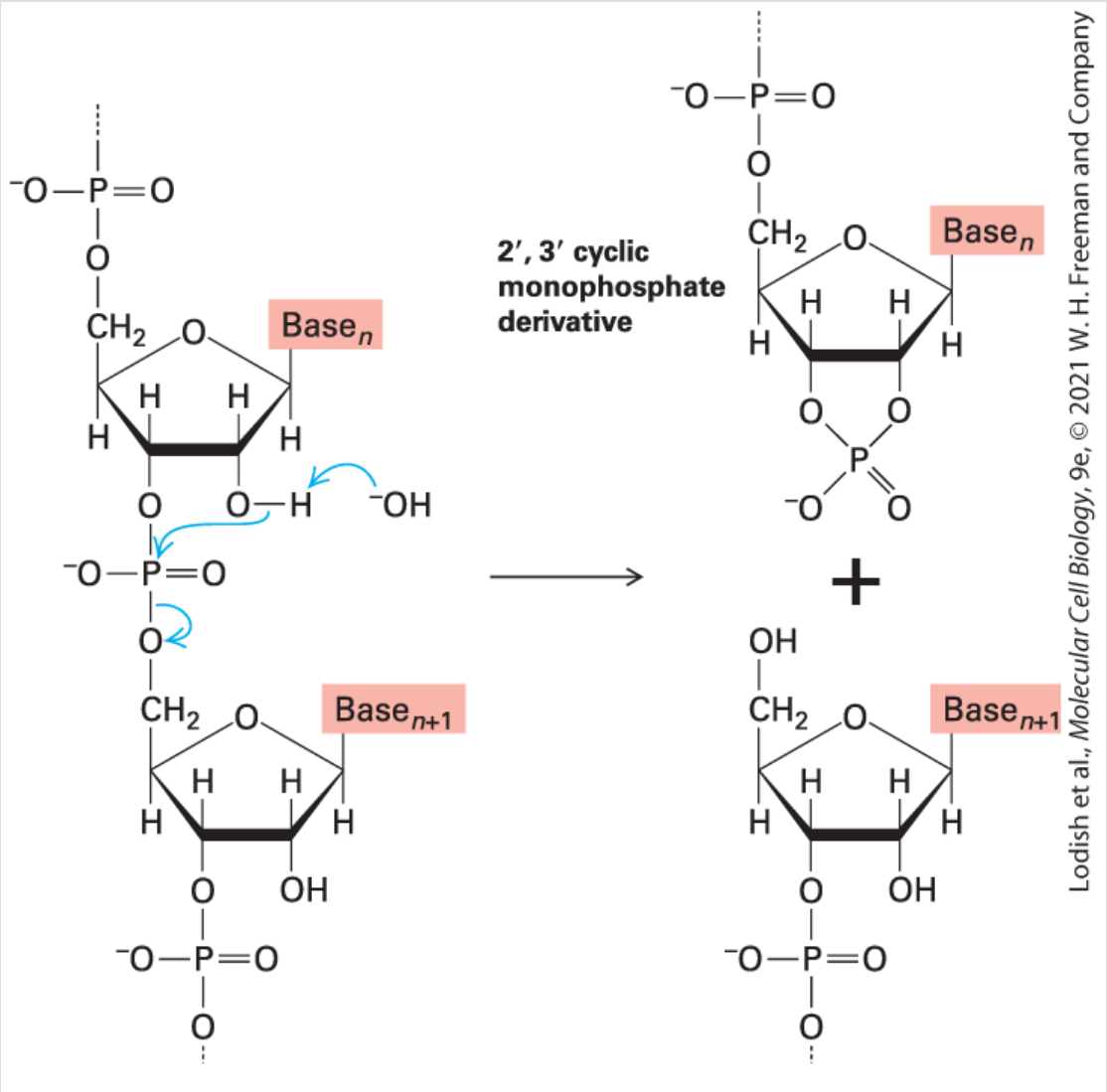
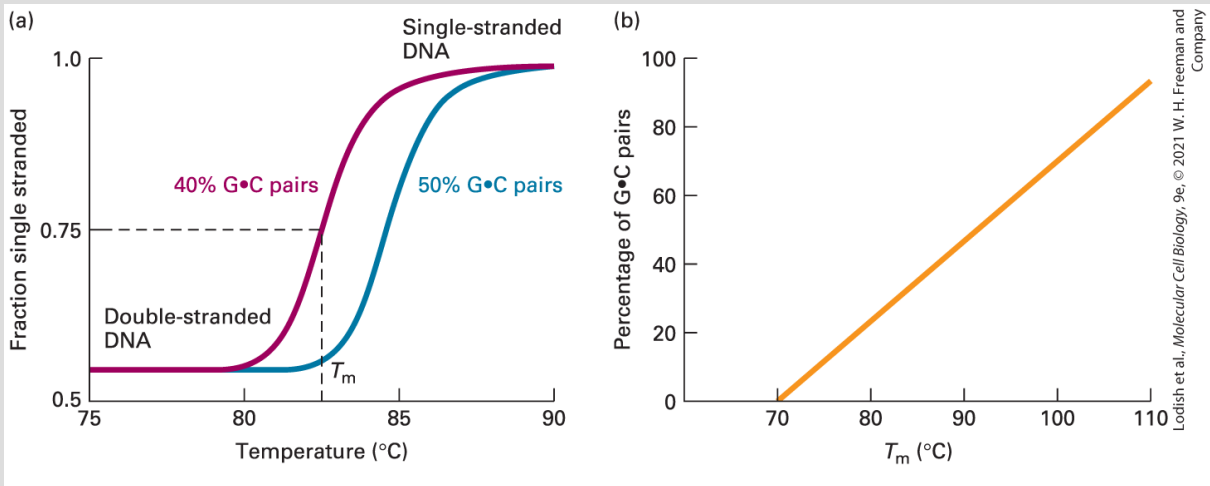


FIGURE 5-6 Spontaneous hydrolysis of RNA catalyzed by the 2′-hydroxyl group. The 2′-hydroxyl group in RNA can act as a nucleophile, attacking the phosphodiester bond. The 2′,3′ cyclic monophosphate derivative is further hydrolyzed to a mixture of 2′ and 3′ monophosphates. Importantly, this facile mechanism of phosphodiester bond hydrolysis cannot occur in DNA, which lacks 2′-hydroxyl groups.

DNA Can Undergo Reversible Strand Separation

During replication and transcription of DNA, the strands of the double helix must separate to allow the internal edges of the bases to pair with the bases of the nucleotides being polymerized into new complementary polynucleotide chains. In later sections, we describe the cellular mechanisms that separate and subsequently reassociate DNA strands during replication and transcription. Here we discuss the fundamental factors that influence the separation and reassociation of DNA strands. These properties of DNA were elucidated by in vitro experiments.

The unwinding and separation of DNA strands, referred to as **denaturation,** or **melting**, can be induced experimentally by increasing the temperature of a solution of DNA. As the thermal energy increases, the resulting increase in molecular motion eventually breaks the hydrogen bonds and other forces that stabilize the double helix. The strands then separate, driven apart by the electrostatic repulsion of the negatively charged deoxyribose-phosphate backbones of the two strands. Near the denaturation temperature, a small increase in temperature causes a rapid, nearly simultaneous loss of the multiple weak interactions holding the strands together along the entire length of the DNA molecules (Figure 5-7a). The stacked base pairs in duplex DNA absorb less ultraviolet (UV) light than the unstacked bases in single-stranded DNA, and this change leads to an abrupt increase in the absorption of UV light. This phenomenon, known as *hyperchromicity*, is useful for monitoring DNA denaturation.



EXPERIMENTAL FIGURE 5-7 G·C content of DNA affects melting temperature. The temperature at which DNA denatures increases with the proportion of G·C pairs. (a) Melting of double-stranded DNA can be monitored by its absorption of UV light at $260\ \mathrm{nm}$. As regions of double-stranded DNA unpair, the absorption of light by those regions increases almost twofold. The temperature at which half the bases in a double-stranded DNA sample have denatured is denoted T_m (for "temperature of melting"). (b) The T_m is a function of the G·C content of the DNA; the greater the $\mathrm{G} + \mathrm{C}$ percentage, the higher the T_m.

The *melting temperature* (T_m) at which DNA strands separate depends on several factors. Molecules that contain a greater proportion of G·C pairs require higher temperatures to denature because the three hydrogen bonds in G·C pairs make these base pairs more stable than A·T pairs, which have only two hydrogen bonds. Indeed, the percentage of G·C base pairs in a DNA sample can be estimated from its T_m (Figure 5-7b). The ion concentration of the solution also influences the T_m because the negatively charged phosphate groups in the two strands are shielded by positively charged ions. When the ion concentration is low, this shielding is decreased, thus increasing the repulsive forces between the strands and reducing the T_m.

Agents that destabilize hydrogen bonds, such as formamide or urea, also lower the T_m. Finally, extremes of pH denature DNA at low temperatures. At low (acid) pH, the bases become protonated and thus positively charged, repelling each other. At high (alkaline) pH, the bases lose protons and become negatively charged, again repelling each other because of their similar charges. In cells, pH and temperature are, for the most part, maintained at a constant level. These features of DNA denaturation are most useful for manipulating DNA in a laboratory setting.

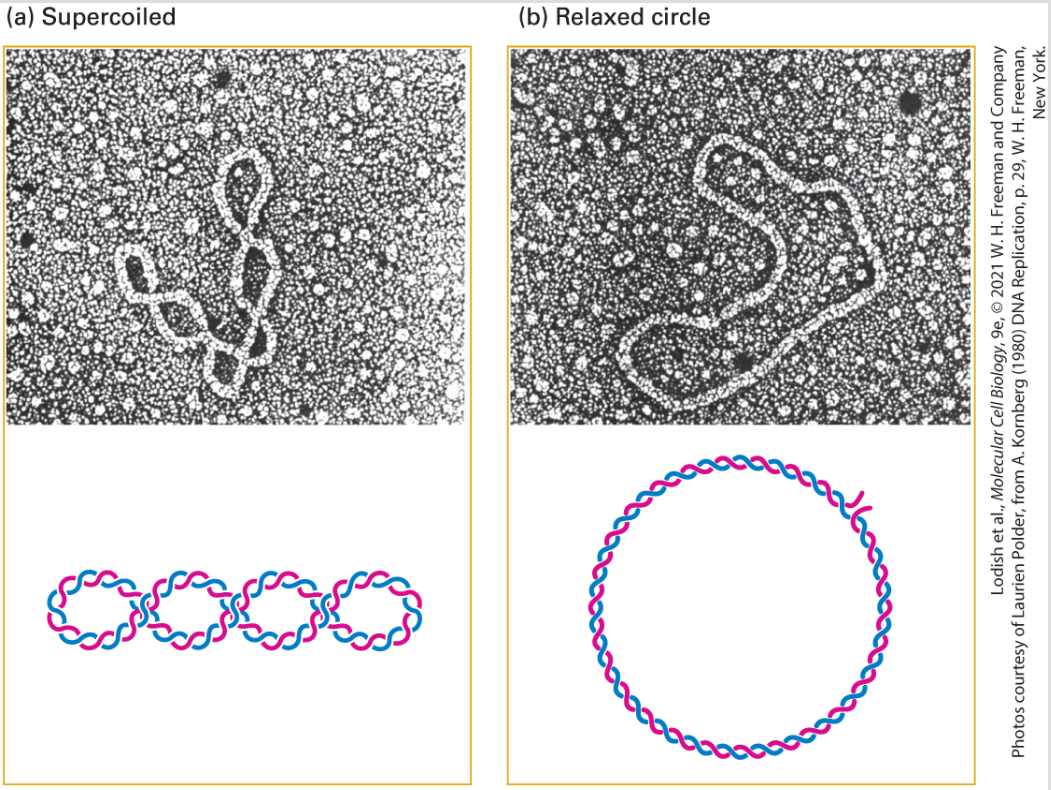
The single-stranded DNA molecules that result from denaturation form random coils without an organized structure. Lowering the temperature, increasing the ion concentration, or neutralizing the pH causes two complementary strands to reassociate into a perfect double helix. The extent of such *renaturation* is dependent on time, the DNA concentration, and the ion concentration. Two DNA strands that are not related in sequence will remain as random coils and will not renature, but they will not inhibit complementary DNA partner strands from finding each other and renaturing. Denaturation and renaturation of DNA are the basis of nucleic acid **hybridization**, a powerful technique used to study the relatedness of two DNA samples and to detect and isolate specific DNA molecules in a mixture containing numerous different DNA sequences (see Chapter 6).

DNA Molecules Can Acquire Torsional Stress

Many bacterial genomic DNAs and many viral DNAs are circular molecules. Circular DNA molecules also occur in mitochondria, which are present in almost all eukaryotic cells, and in chloroplasts, which are present in plants and some unicellular eukaryotes. Although eukaryotic nuclear DNA is linear, long loops of DNA are fixed in place within chromosomes (see Chapter 8). Each of the two strands in a circular DNA molecule or in a fixed loop of a eukaryotic chromosome forms a closed structure without free ends and is therefore subject to torsional stress.

When DNA is either overwound or underwound, meaning that it has more or fewer helical turns than B-form linear DNA of the same length, the torsional stress thus produced will be relieved by the DNA molecule twisting back on itself forming *supercoils* (Figure 5-8a). A similar topological interconversion between turns and supercoils can be demonstrated by the twists that form in a rubber band after being rolled between your fingertips. As we will see in the next section, when DNA replicates, the two strands must unwind; the remaining portion of the molecule would gain one supercoil for the unwinding of every complete turn of the strands in the double helix, which occurs every 10 base pairs. Since a typical replication fork moves at a rate of about 500 bases per second, supercoils would be introduced into the DNA ahead of the moving replication fork at a

rate of 50 per second. To relieve the torsional stress that would be caused by supercoils developed during replication, all cells contain *topoisomerase I*. This enzyme binds to DNA at random sites and breaks a phosphodiester bond in one strand. Such a one-strand break in DNA is called a *nick*. The broken end then winds around the uncut strand, leading to loss of supercoils (Figure 5-8b). Finally, the same enzyme joins (ligates) the two ends of the broken strand. Another type of enzyme, *topoisomerase II*, relieves torsional stress in DNA by making breaks in both strands of a double-stranded DNA and then religating them.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Photos courtesy of Laurien Polder, from A. Kornberg (1980) DNA Replication, p. 29, W. H. Freeman, New York.

EXPERIMENTAL FIGURE 5-8 Topoisomerase I relieves torsional stress on DNA. (a) Electron micrograph of SV40 viral DNA. When the circular DNA of the SV40 virus is isolated and separated from its associated protein, the DNA duplex is

underwound and assumes the supercoiled configuration. (b) If a supercoiled DNA is nicked (i.e., one strand cleaved), the strands can rewind, leading to loss of a supercoil. Topoisomerase I catalyzes this reaction and also *reseals* the broken ends. All the supercoils in isolated SV40 DNA can be removed by the sequential action of this enzyme, producing the relaxed-circle conformation. For clarity, the shapes of the molecules at the bottom have been simplified.

KEY CONCEPTS OF SECTION 5.1

The Double-Helical Structure of DNA

- Deoxyribonucleic acid (DNA), the genetic material, is made up of two long, unbranched polymers of nucleotides. A nucleotide consists of a phosphorylated pentose linked to either a purine-adenine (A) or guanine (G), or a pyrimidine-cytosine (C) or thymine (T). Adjacent nucleotides in a polynucleotide are linked by phosphodiester bonds. The entire strand has a chemical directionality with $5'$ and $3'$ ends (see Figure 5-2).
- Natural DNA contains two complementary antiparallel polynucleotide strands wound together into a regular right-handed double helix with the bases on the inside and the two sugar-phosphate backbones on the outside (see Figure 5-3). The double-stranded structure of DNA is significantly stabilized by noncovalent bonds consisting of hydrogen bonds between pairs of bases on opposite strands and by hydrophobic interactions between adjacent base pairs stacked perpendicular to the helix axis.
- The standard Watson-Crick base pairs — G·C and A·T — have nearly identical dimensions and are the only base pairs accommodated by the native double-helical structure.
- Binding of protein to DNA can deform its helical structure, causing local bending or unwinding of the DNA molecule.
- Heat causes the DNA strands to separate (denature). The melting temperature (T_{m}) of DNA increases with the percentage of G·C base pairs. Under suitable conditions, separated complementary nucleic acid strands will renature.
- Overwinding or underwinding of DNA causes long DNA molecules to twist on themselves, forming supercoils (see Figure 5-7). Unwinding of chromosomal DNA during replication would generate a large number of supercoils except that enzymes called *topoisomerases* relieve this torsional stress by removing supercoils from DNA molecules.

5.2 DNA Replication

In a human cell, the genetic information encoded in DNA resides in 23 pairs of chromosomes, with each chromosome consisting of a double-stranded helical molecule hundreds of millions of nucleotides long. When cells divide, each of the two daughter cells must contain exactly the same genetic information as the parent cell. Thus every cell division must be accompanied by the exact copying of the DNA sequence of each chromosome by a process known as DNA replication. A key feature of the double-helical DNA structure that was immediately apparent to Watson and Crick was that the invariant rules for base pairing (A with T and G with C) mean that each strand of a DNA polymer carries the same information in complementary form. The implication was that DNA replication could be achieved by separating the two strands of the DNA polymer and then using each strand as a template in the synthesis of new, complementary strands, thus producing two duplicate, double-stranded copies of the original DNA molecule.

An understanding of the biochemical mechanism for copying of DNA came from the isolation of the first template-directed DNA polymerase enzyme from extracts of *Escherichia coli* cells by Arthur Kornberg and colleagues in 1956. Although we now know that the isolated enzyme known as DNA polymerase I is involved in DNA repair and not the replication of the *E. coli* chromosome, all template-directed DNA polymerases operate by the same

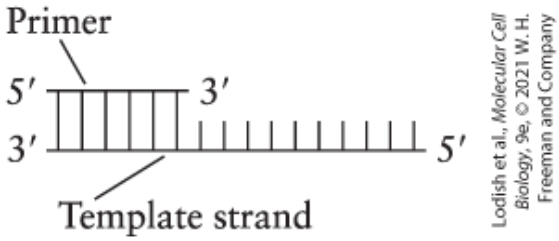
fundamental rules. The lessons learned from studying DNA polymerase I have proved to be universal.

DNA Polymerases Require a Template and a Primer to Replicate DNA

All DNA polymerases have the same requirements for synthesis of a new DNA strand:

i. A single-stranded DNA template.

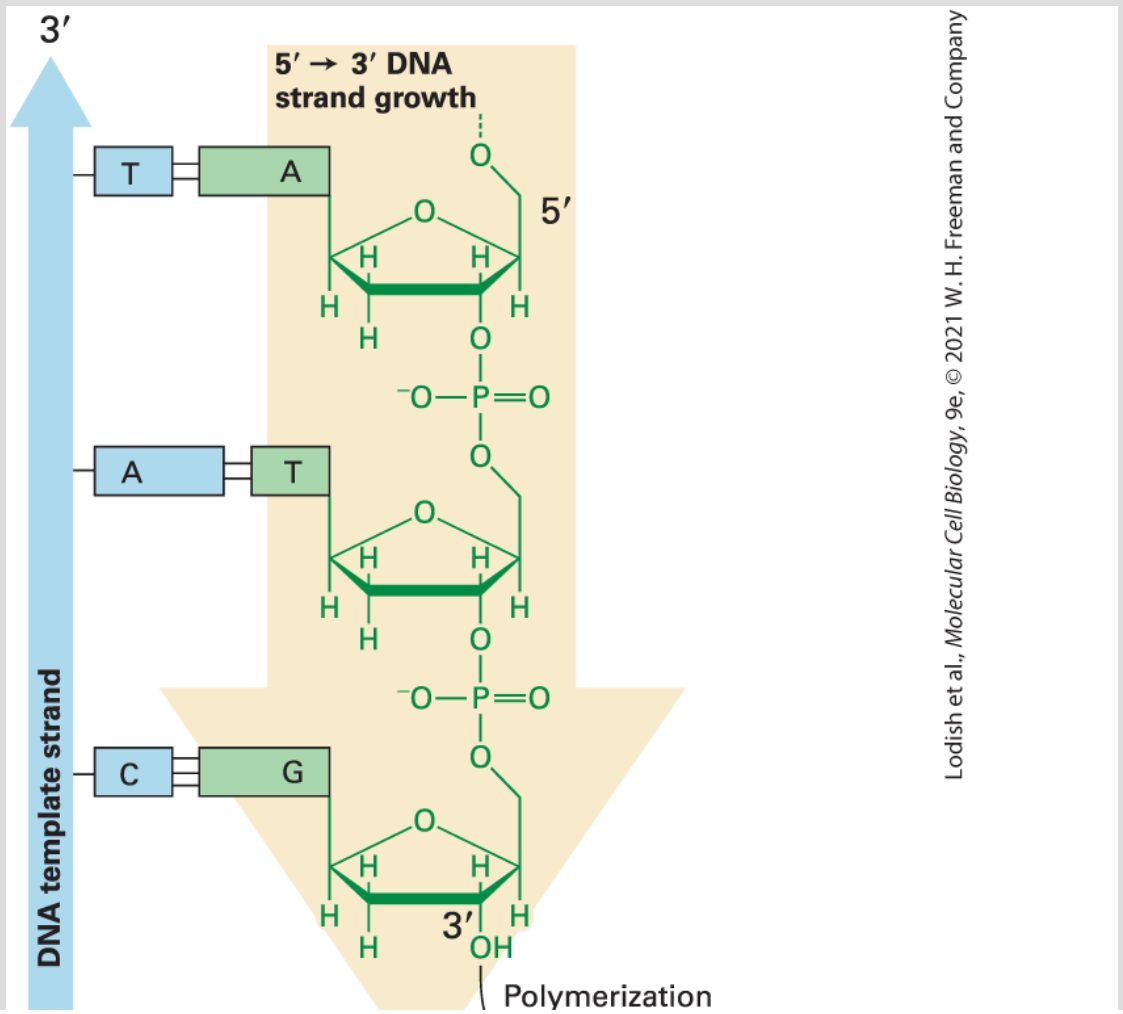
ii. A DNA **primer** base paired with the template, and with a free hydroxyl group at the 3′ end of the primer to accept a new nucleotide.



iii. A source of deoxyribonucleoside 5′-triphosphate (dNTPs) precursors.

With a primer base-paired to the template strand, a DNA polymerase will add deoxyribonucleotides to the free hydroxyl group at the 3′ end of the primer, as directed by the sequence of the template

strand. The energy for this reaction comes from the release of diphosphate when a phosphoester bond forms between the 3′ oxygen and the α phosphate of the appropriate dNTP. The equilibrium for the reaction is driven further toward chain elongation by pyrophosphatase, an enzyme that catalyzes cleavage of the released PP_i into two molecules of inorganic phosphate (Figure 5-9). After the addition of a nucleotide, a new primer 3′ end will be available to accept the next nucleotide. Polymerization thus proceeds in the 5′ → 3′ direction and will continue until the growing end of the primer reaches the end of the template strand.



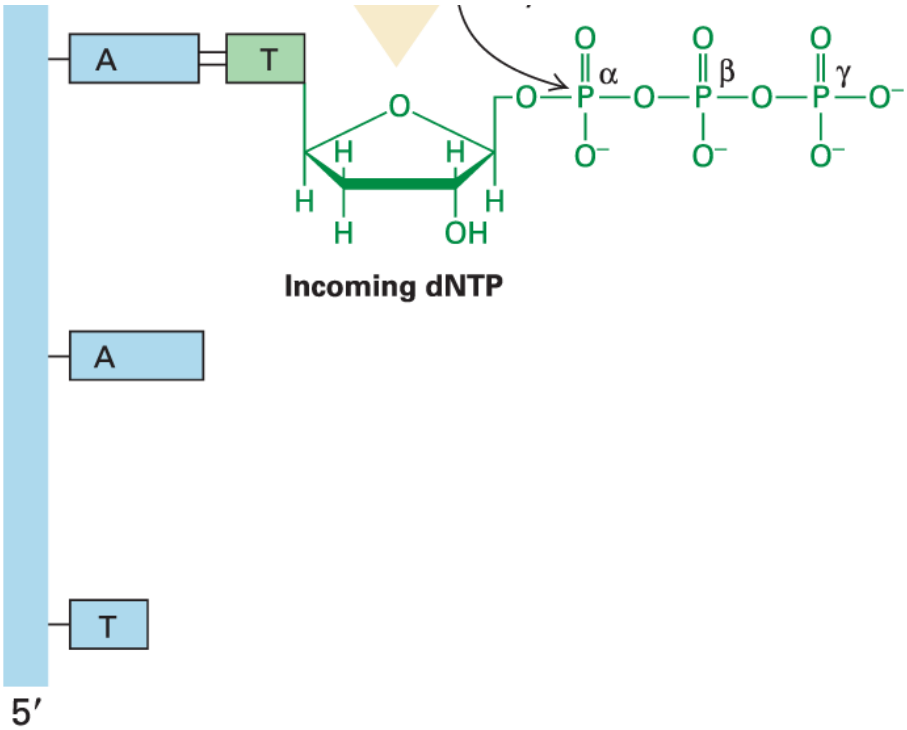


FIGURE 5-9 DNA is synthesized $5' \rightarrow 3'$ from dNTP precursors. The three requirements for DNA synthesis are a primer strand with a free $3'$ terminus, a template strand that is base paired to the primer, and a source of deoxyribonucleoside triphosphates (dNTPs). DNA polymerase will act on this substrate to add a new dNTP at the $3'$ end of the primer strand as specified by base pairing between the added base and the template DNA strand. DNA polymerase catalyzes formation of a phosphodiester bond between the $3'$ oxygen of the primer strand and the α phosphate of a correctly base-paired dNTP. New DNA strands are always synthesized in the $5' \rightarrow 3'$ direction and are opposite in polarity to their template DNA strands.

The faithful replication of chromosomal DNA is crucial for life, and DNA polymerases have evolved to replicate DNA with extraordinary speed and fidelity. A typical DNA polymerase can add a single nucleotide to a growing chain in about 2 milliseconds, but this does not give much time for the polymerase to discriminate between a correctly and an incorrectly base-paired nucleotide. If the fidelity of

polymerization was based only on the difference in energy between the correctly and incorrectly hydrogen-bonded nucleotide pair, the polymerase would make at least one mistake every 50 nucleotides. A much higher degree of fidelity is achieved by the active site of the polymerase only accommodating the base pairs with the exact geometry of a normal Watson-Crick base pair while rejecting nonstandard base-pair geometries even if they are only slightly different (see Figure 5-4). By using this type of geometric selection, most polymerases can achieve a fidelity of initial nucleotide incorporation that would produce error rates of about 1 incorrect nucleotide per 10^4 (10,000) polymerized nucleotides.

The fidelity of DNA replication in most organisms is about 1 mistake in 10^9 (one billion) nucleotides incorporated into a growing strand, which is much greater than can be explained by the maximum selectivity of base pairing. This remarkable accuracy is largely due to *proofreading* by DNA polymerases. Proofreading depends on the *$3' \rightarrow 5'$ exonuclease activity* of some DNA polymerases. When an incorrect base is incorporated during DNA synthesis, base pairing between the 3′ nucleotide of the nascent strand and the template strand does not occur. As a result, the polymerase pauses, then transfers the 3′ end of the growing chain to its exonuclease site, where the incorrect mispaired base is removed (Figure 5-10). The 3′ end is then transferred back to the polymerase site, where this region is copied correctly. Most DNA polymerases require a high degree of fidelity and thus have proofreading. For example, all three

E. coli DNA polymerases have proofreading $3' \rightarrow 5'$ exonuclease activity.

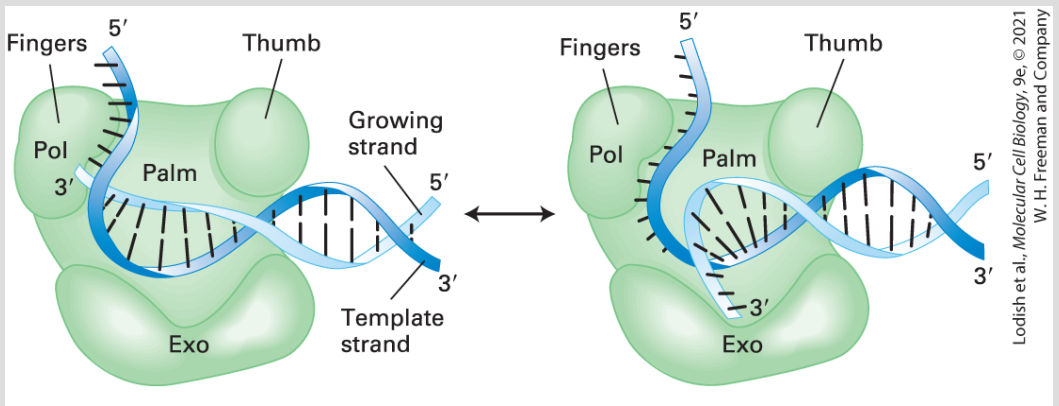


FIGURE 5-10 Proofreading by DNA polymerase. All DNA polymerases have a similar three-dimensional structure, which resembles a half-opened right hand. The "fingers" bind the single-stranded segment of the template strand, and the polymerase catalytic activity (Pol) lies in the junction between the fingers and "palm." As long as the correct nucleotides are added to the $3'$ end of the growing strand, it remains in the polymerase site. Incorporation of an incorrect base at the $3'$ end results in a substrate that does not satisfy the requirements for polymerization of the next nucleotide and the polymerase pauses. Because the $3'$ end of the primer strand is not stably base-paired to the template, it is free to move to the $3' \rightarrow 5'$ exonuclease site (Exo) about $3\ \mathrm{nm}$ away, where the mispaired base is removed. Subsequently, the new $3'$ end, now freed of the mispaired nucleotide, can base pair properly with the template and polymerization can resume. See C. M. Joyce and T. T. Steitz, 1995, *J. Bacteriol.* **177**:6321, and S. Bell and T. Baker, 1998, *Cell* **92**:295.

Duplex DNA Is Unwound, and Daughter Strands Are Formed at the DNA Replication Fork

In order for duplex DNA to function as a template during replication, the two intertwined strands must be unwound, or melted, to make their bases available for pairing with the bases of the dNTPs that are polymerized into the newly synthesized daughter strands. This unwinding of the parent DNA strands is performed by enzymes called **helicases**. Unwinding begins at segments in a DNA molecule called *replication origins*, or simply *origins*. The nucleotide sequences of origins from different organisms vary greatly, although they usually contain AT-rich sequences. DNA polymerases cannot initiate replication de novo because, as we have seen, all of these polymerases require a free 3′ primer terminus. However, RNA polymerases that have a similar mechanism to DNA polymerases do have the capability of initiating new strand synthesis from a template without a preexisting primer (see Figure 5-23a). Once helicases have unwound the parent DNA at an origin, a specialized RNA polymerase called **primase** forms a short (~12-nucleotide) RNA primer complementary to the unwound template strands. The primer, still base-paired to its complementary DNA strand, is then elongated by *DNA polymerase α* for another 25 nucleotides or so, forming a primer made of RNA at the 5′ end and DNA at the 3′ end. This primer is further extended by DNA polymerase δ, thereby forming a new daughter strand.

The DNA region at which all these proteins come together to carry out the synthesis of daughter strands is called the **replication fork**. As replication proceeds, the replication fork and the associated

proteins move away from the origin. As noted earlier, local unwinding of duplex DNA produces torsional stress, which is relieved by topoisomerase I. In order for DNA polymerases to move along and copy a duplex DNA, helicase must sequentially unwind the duplex and topoisomerase must remove the supercoils that form.

A major complication in the operation of a DNA replication fork arises from two properties of DNA: the two strands of the parent DNA duplex are antiparallel, and DNA polymerases (like RNA polymerases) can add nucleotides to the growing daughter strands only in the $5' \rightarrow 3'$ direction. Synthesis of one daughter strand, called the **leading strand**, can proceed continuously from a single RNA primer in the $5' \rightarrow 3'$ direction, *the same direction as the movement of the replication fork* (Figure 5-11). The problem comes in the synthesis of the other daughter strand, called the **lagging strand**.

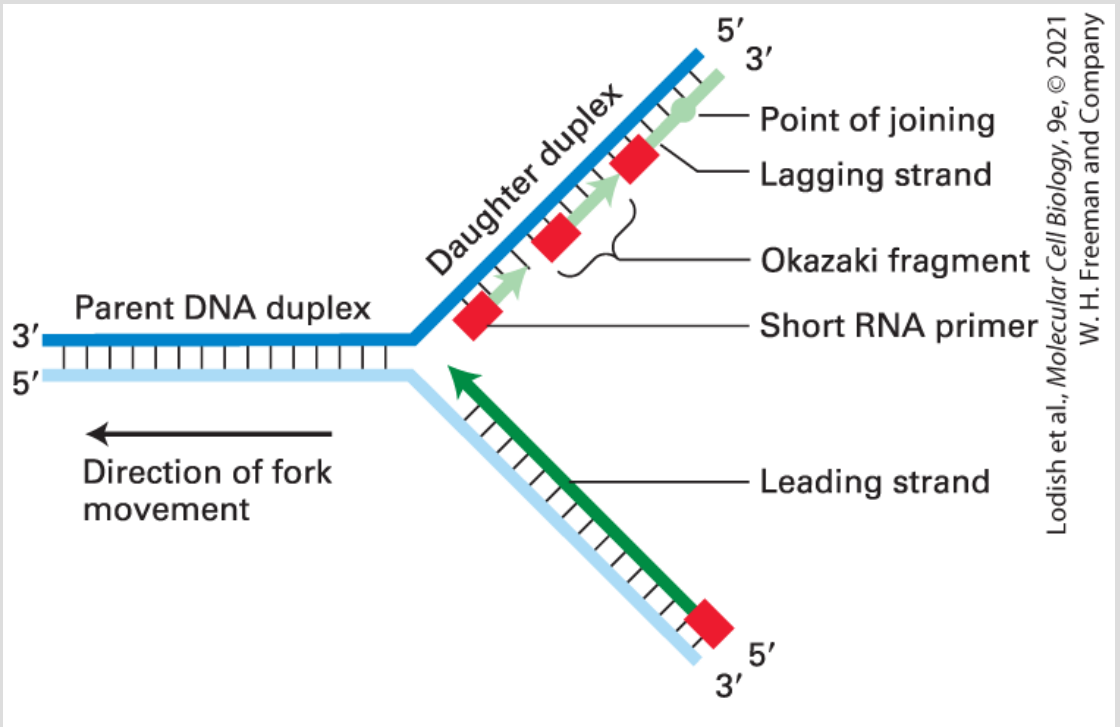


FIGURE 5-11 Leading-strand and lagging-strand DNA synthesis. Nucleotides are added by a DNA polymerase to each growing daughter strand in the $5' \rightarrow 3'$ direction (indicated by arrowheads). The leading strand is synthesized continuously from a single RNA primer (red) at its $5'$ end. The lagging strand is synthesized discontinuously from multiple RNA primers that are formed periodically as each new region of the parent duplex is unwound. Elongation of these primers initially produces Okazaki fragments. As each growing fragment approaches the previous primer, that primer is removed and the fragments are ligated. Repetition of this process eventually results in synthesis of the entire lagging strand.

Because growth of the lagging strand must occur in the $5' \rightarrow 3'$ direction, copying of this strand must proceed in the *opposite* direction from the movement of the replication fork. A cell accomplishes this feat by synthesizing a new primer every 100 to 200 nucleotides on that template strand as more of the strand is exposed by unwinding. Each of these primers, base-paired to the template

strand, is elongated in the $5' \rightarrow 3'$ direction, forming discontinuous segments named **Okazaki fragments** after their discoverer, Reiji Okazaki (see Figure 5-11). The RNA primer of each Okazaki fragment is removed and replaced by DNA chain growth from the neighboring Okazaki fragment; and an enzyme called *DNA ligase* joins the adjacent fragments.

A DNA Replication Fork Advances by Cooperation of Multiple Proteins

Detailed understanding of the eukaryotic proteins that participate in DNA replication initially came from studies with small viral DNAs, particularly the circular genome of SV40, a virus which infects monkeys. Virus-infected cells replicate large numbers of the simple viral genome in a short period of time, which makes them an ideal model system for studying basic aspects of DNA replication. Because simple viruses such as SV40 depend largely on the DNA replication machinery of their host cells (in this case monkey cells), they offer a unique opportunity to study the replication of multiple identical, small DNA molecules by cellular proteins. Figure 5-12 depicts the multiple proteins that coordinate the copying of SV40 DNA at a replication fork. The assembled proteins at a replication fork further illustrate the concept of molecular machines introduced in Chapter 3. These multicomponent complexes permit the cell to

carry out an ordered sequence of events that accomplishes essential cell functions.

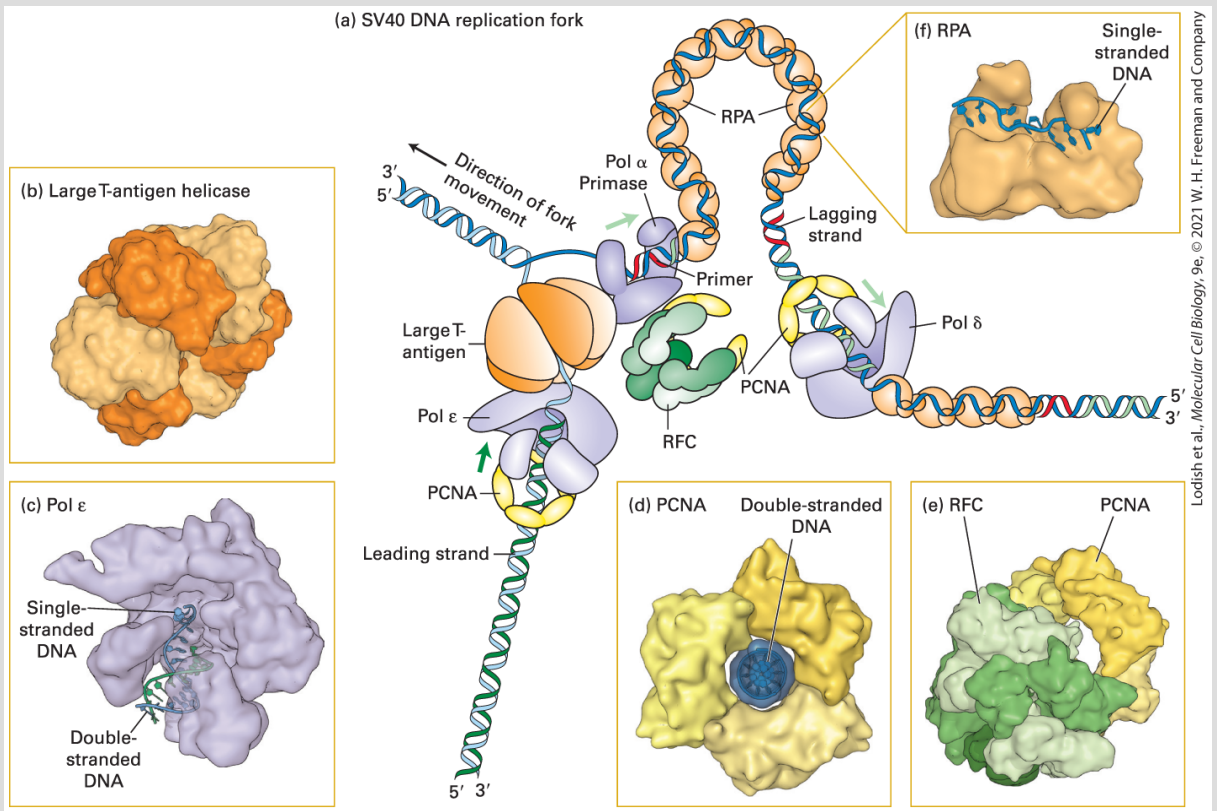


FIGURE 5-12 Model of an SV40 DNA replication fork. (a) A hexamer of large T-antigen, a viral protein, functions as a helicase to unwind the parent DNA strands. The leading strand is extended by DNA polymerase ε (Pol ε) up to the replication fork. Pol ε is bound to a ring of PCNA that surrounds the daughter double-stranded DNA so that the Pol ε–PCNA complex remains stably associated with the replication fork. The single-stranded region of the lagging-strand template generated by T-antigen helicase is bound by multiple copies of the heterotrimeric protein RPA. Primers for lagging-strand synthesis (red, RNA; light green, DNA) are synthesized by a complex of primase and DNA polymerase α (Pol α). The 3′ end of each primer synthesized by Pol α–primase is then bound by a PCNA–Pol δ complex, which extends the primer and synthesizes most of each Okazaki fragment. (b) The helicase domain of SV40 T-antigen forms a hexameric replicative helicase. Subunits are shown in alternating light and dark orange. (c) Model of DNA polymerase ε extending the 3′-end of the leading strand. (d) The three subunits of PCNA, shown in different shades of yellow, form a circular structure with a central hole through which daughter double-stranded DNA passes. (e) RFC,

the pentameric “clamp-loader” (monomers shown in different shades of green) is shown bound to a circular trimer PCNA before the PCNA *clamp* is opened. (f) The large subunit of RPA contains two domains that bind single-stranded DNA. Note that the single DNA strand is extended, with the bases oriented in an optimal conformation for replication by Pol δ. See M. O’Donnell, L. Langston, and B. Stillman, 2013, *Cold Spring Harbor Perspect. Biol.* **5**:a010108.

[Part (b) Data from D. Li et al., 2003, *Nature* **423**:512, PDB ID 1n25. Part (c) Data from M. Hogg et al., 2014, *Nat. Struct. Mol. Biol.*, **21**:49, PDB ID 4m8o. Part (d) Data from J. M. Gulbis et al., 1996, *Cell* **87**:297, PDB ID 1axc. Part (e) Data from G. D. Bowman, M. O’Donnell, and J. Kuriyan, 2004, *Nature* **429**:724, PDB ID 1sxj. Part (f) Data from A. Bochkarev et al., 1997, *Nature* **385**:176, PDB ID 1jmc.]

The molecular machine that replicates SV40 DNA contains only one viral protein; all other proteins involved in SV40 DNA replication are provided by the host cell. This viral protein, *large T-antigen*, forms a hexameric *replicative helicase*, a protein that uses energy from ATP hydrolysis to unwind the parent strands at a replication fork. Primers for the leading and lagging daughter strands are synthesized by a complex of *primase*, which synthesizes a short RNA primer (~12 nucleotides), and DNA polymerase α (Pol α), which extends the RNA primer with deoxyribonucleotides for another 25 nucleotides or so, forming a mixed RNA-DNA primer.

The short RNA-DNA primer is extended by the high fidelity *DNA polymerase δ* (Pol δ), which possess a proofreading mechanism based on $3' \rightarrow 5'$ exonuclease activity. During the replication of cellular DNA, Pol δ synthesizes lagging-strand DNA, while a second high-fidelity polymerase, *DNA polymerase* ε (Pol ε), synthesizes most of the

length of the leading strand. Pol δ and Pol ε each form a complex with *PCNA* (*p*roliferating *c*ell *n*uclear *a*ntigen), which displaces the primase–Pol α complex following primer synthesis. As illustrated in Figure 5-12d, PCNA is a homotrimeric protein that has a central hole through which the daughter duplex DNA passes, thereby preventing the PCNA–Pol δ and PCNA–Pol ε complexes from dissociating from the template. As such, PCNA is known as a *sliding clamp* that enables Pol δ and Pol ε to remain stably associated with a single template strand for thousands of nucleotides. A pentameric protein called *RFC* (*r*eplication *f*actor *C*) functions to open the PCNA ring so that it can encircle the short region of double-stranded DNA synthesized by Pol α. Consequently, RFC is called a *clamp loader*.

After parent DNA is separated into single-stranded templates at the replication fork, the leading strand is extended by Pol ε, which can extend the growing strand up to the replication fork. The single-stranded template for lagging-strand synthesis is bound by multiple copies of RPA (*r*eplication *p*rotein *A*), a heterotrimeric protein (Figure 5-12c). Binding of RPA maintains the template in a uniform conformation that is optimal for copying by Pol δ. Bound RPA proteins are dislodged from the parent strand by Pol δ as it synthesizes the complementary strand base-paired with the parent strand.

Several other eukaryotic proteins that function in DNA replication are not depicted in Figure 5-12. For example, topoisomerase I

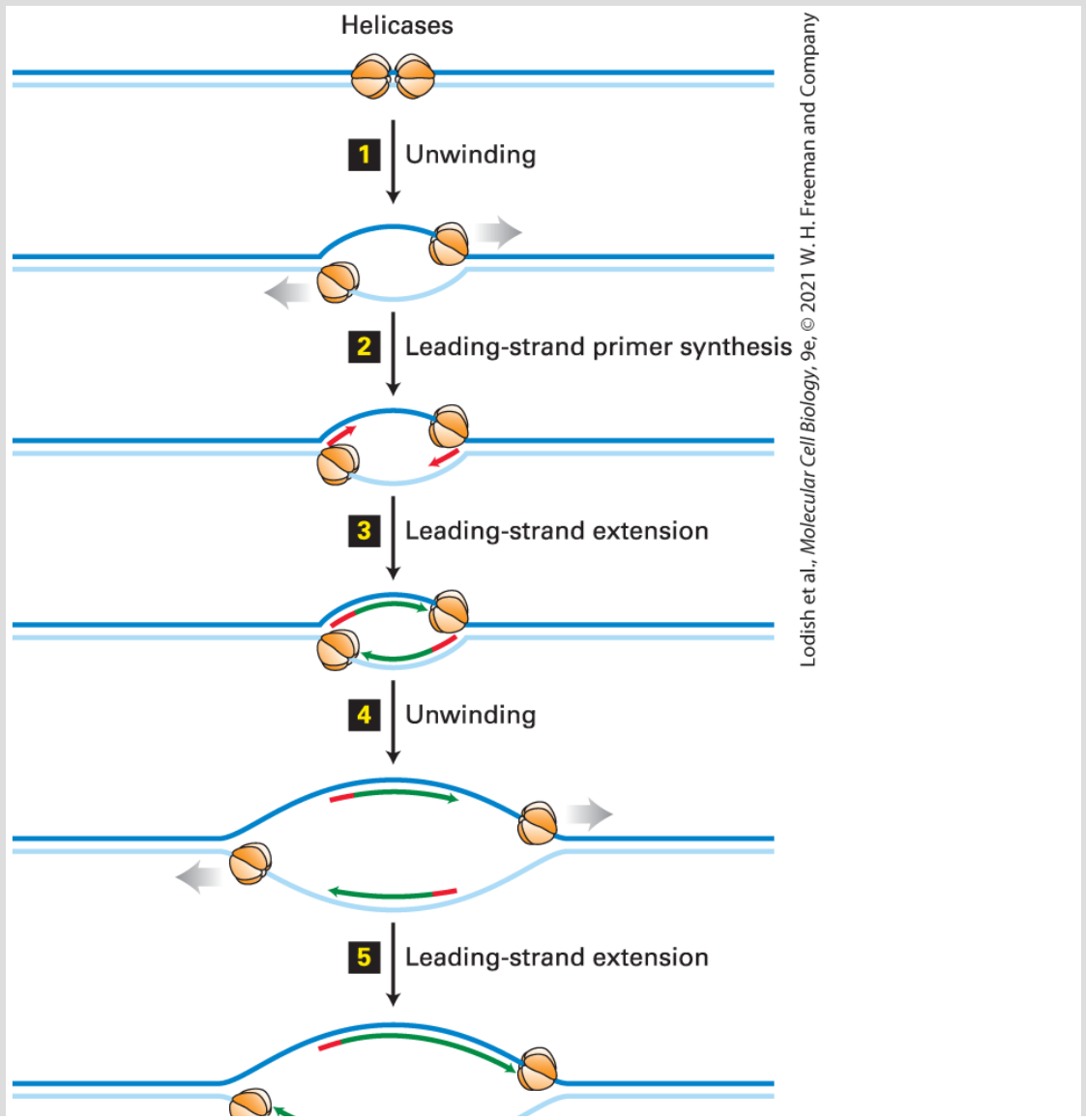
associates with the parental double-stranded DNA ahead of the replicative helicase (i.e., to the left of large T-antigen in Figure 5-12) to remove torsional stress introduced by the unwinding of the parent strands (see Figure 5-7a). Ribonuclease H and FEN I remove the ribonucleotides at the 5′ ends of Okazaki fragments; these ribonucleotides are replaced by deoxyribonucleotides added by Pol δ as it extends the upstream Okazaki fragment. Successive Okazaki fragments are coupled by DNA ligase through standard $5' \rightarrow 3'$ phosphoester bonds. Other specialized DNA polymerases are involved in the repair of mismatches and lesions in DNA (see Section 5.3).

DNA Replication Occurs Bidirectionally from Each Origin

As indicated in Figures 5-11 and 5-12, both parent DNA strands that are exposed by local unwinding at a replication fork are copied into daughter strands. In theory, DNA replication from a single origin could involve one replication fork that moves in one direction. Alternatively, two replication forks might assemble at a single origin and then move in opposite directions, leading to *bidirectional growth* of both daughter strands.

We now know that all bacterial, archaeal, and eukaryotic cells employ a bidirectional mechanism of DNA replication. In the case of

SV40 DNA, replication is initiated by the binding of two large T-antigen hexameric helicases to the single SV40 origin and the assembly of other proteins to form two replication forks. These forks then move away from the SV40 origin in opposite directions, and leading- and lagging-strand synthesis occurs at both forks. As shown in Figure 5-13, the left replication fork extends DNA synthesis in the leftward direction; similarly, the right replication fork extends DNA synthesis in the rightward direction.



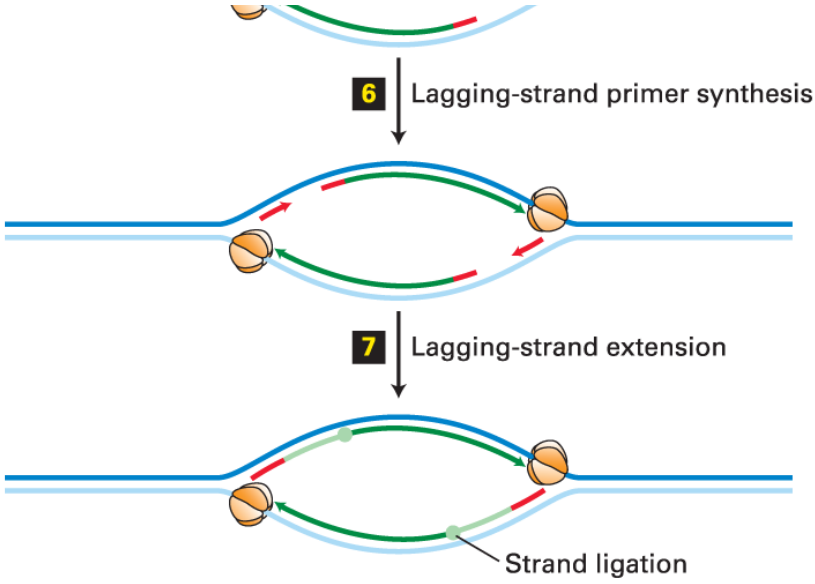


FIGURE 5-13 Bidirectional mechanism of DNA replication. The left replication fork here is comparable to the replication fork diagrammed in Figure 5-29 (although that figure also shows proteins other than large T-antigen, which are not shown here). *Top:* Two large T-antigen hexameric helicases first bind at the replication origin in opposite orientations. Step **1**: Using energy provided by ATP hydrolysis, the helicases move in opposite directions, unwinding the parent DNA and generating single-stranded templates, which are bound by RPA proteins (not shown). Step **2**: Primase–Pol α complexes synthesize short primers (red) base-paired to each of the separated parent strands. Step **3**: PCNA-Rfc–Pol ε complexes replace the primase–Pol α complexes and extend the short primers, generating the leading strands (dark green) at each replication fork. Step **4**: The helicases further unwind the parent strands, and RPA proteins bind to the newly exposed single-stranded regions. Step **5**: PCNA-Rfc–Pol ε complexes extend the leading strands farther. Step **6**: Primase–Pol α complexes synthesize primers for lagging-strand synthesis at each replication fork. Step **7**: PCNA-Rfc–Pol δ complexes displace the primase–Pol α complexes and extend the lagging-strand Okazaki fragments (light green), which are eventually ligated to the 5′ ends of the leading strands. The position where ligation occurs is represented by a circle. Replication continues by further unwinding of the parent strands and synthesis of leading and lagging strands as in steps **4**–**7**. (Although depicted here as individual steps for clarity, unwinding and synthesis of leading and lagging strands occur concurrently.)

Unlike SV40 DNA, eukaryotic chromosomal DNA molecules contain multiple replication origins separated by tens to hundreds of kilobases. A six-subunit protein called *ORC*, for *o*rigin *r*ecognition *c*omplex, binds to each origin and associates with other proteins required to load cellular hexameric helicases composed of six homologous *MCM* (*m*ini*c*hromosome *m*aintenance) proteins. Two MCM helicases, oriented in opposite directions, separate the parent strands at an origin, and RPA proteins bind to the resulting single-stranded DNA. Synthesis of primers and subsequent steps in the replication of cellular DNA are thought to be analogous to those in SV40 DNA replication (see Figure 5-13).

Cell division begins with the duplication of chromosomes; therefore, the initiation of DNA replication is usually the first step in the cell division cycle. Cellular DNA replication is initiated by activation of MCM helicase by a specific protein kinase (*DDK*), which in turn is regulated by S-phase *cyclin-dependent kinases*. Other cyclin-dependent kinases regulate additional aspects of cell proliferation, including the complex process of mitosis by which a eukaryotic cell divides into two daughter cells. Mitosis and another specialized type of cell division called *meiosis,* which generates haploid sperm and egg cells, is discussed in Chapter 6. We discuss the various regulatory mechanisms that determine the rate of cell division in Chapter 19.

KEY CONCEPTS OF SECTION 5.2

DNA Replication

- DNA polymerases require, as a substrate, a template strand that is base paired to a primer strand with an available $3'$ hydroxyl group. Newly synthesized DNA extends from the primer in the $5' \rightarrow 3'$ direction.
- The high fidelity of DNA replication comes from rigorous geometric constraints for only A·T and G·C base pairs at the polymerase active site, and by a proofreading exonuclease that removes incorrect base pairs.
- Replication begins at a sequence called an *origin*. Each eukaryotic chromosomal DNA molecule contains multiple replication origins. Replication initiates at an origin by unwinding of the DNA and initiation of new synthesis by an RNA polymerases known as primase.
- After initiation replication proceeds from a replication fork, at which one daughter strand (the leading strand) is elongated continuously. The other daughter strand (the lagging strand) is formed discontinuously (see Figure 5-11).
- Discontinuous replication of the lagging strand proceeds by formation of Okazaki fragments synthesized every 100 to 200 nucleotides. The ribonucleotide primers at the $5'$ end of each Okazaki fragment are removed and replaced by elongation from the $3'$ end of the next Okazaki fragment. Finally, adjacent Okazaki fragments are joined by DNA ligase.
- Helicases use energy from ATP hydrolysis to separate the parent (template) DNA strands, which are initially bound by multiple copies of a single-stranded DNA-binding protein, RPA. Primase synthesizes a short RNA primer, which remains base-paired to the template DNA. This primer is initially extended at the $3'$ end by DNA polymerase α (Pol α), resulting in a short daughter strand made up of RNA and DNA.
- Most of the DNA in eukaryotic cells is synthesized by Pol δ and Pol ε, which take over from Pol α and continue elongation of the daughter strands in the $5' \rightarrow 3'$ direction. Pol δ synthesizes most of the length of the lagging strand, while Pol ε synthesizes the leading strand. Pol δ and Pol ε remain stably associated with the template by binding to PCNA, a trimeric protein that encircles the daughter duplex DNA, functioning as a sliding clamp.
- DNA replication generally occurs by a bidirectional mechanism in which two replication forks form at an origin and move in opposite directions, and both template strands are copied at each fork.
- MCM helicases initiate eukaryotic DNA replication in vivo at multiple origins spaced along chromosomal DNA. Synthesis of eukaryotic DNA is regulated by controls on the binding and activity of these helicases.

5.3 DNA Repair and Recombination

Damage to DNA is unavoidable and arises in many ways. DNA damage can be caused by spontaneous cleavage of chemical bonds in DNA, by environmental agents such as ultraviolet and ionizing radiation, and by reaction with genotoxic chemicals that are by-products of normal cellular metabolism or occur in the environment. A change in the normal DNA sequence, called a **mutation**, can occur during replication when a DNA polymerase inserts the wrong nucleotide as it reads a damaged template. Mutations also occur at a low frequency as the result of copying errors introduced by DNA polymerases when they replicate an undamaged template. In humans, mutations that compromise the integrity of the genome can have multiple undesirable consequences. Mutations in the DNA in germ cells will be passed on to offspring, usually with deleterious effects. By contrast, mutations in somatic cells that affect genes that are normally involved in the careful regulation of cell division can cause uncontrolled cell division, eventually leading to tumor formation and cancer. It is thus not surprising that cells have evolved elaborate mechanisms to repair different kinds of damage to the integrity of DNA before the damage becomes fixed as a mutation that is passed to daughter cells.

DNA repair mechanisms take advantage of the inherent redundancy of information contained in complementary strands of DNA and correct the damaged strand using information from the undamaged

strand that is still intact. To accomplish this, most DNA repair processes follow the same general outline. All repair processes must first identify the site of the damage. This is usually done by scanning for disruptions to the regular structure of the DNA helix. Once a lesion is identified, most repair processes then selectively remove a portion of the damaged strand and copy the missing information from the intact complementary strand. In this section, we will review the mechanisms for repair of mismatches between normal bases, repair of chemically altered bases, and repair of broken ends of DNA.

Chemical and Radiation Damage to DNA Can Lead to Mutations

DNA is continually subjected to a barrage of damaging chemical reactions; estimates of the number of DNA-damage events in a single human cell range from 10^4 to 10^6 per day. Even if DNA were not exposed to damaging chemicals, certain aspects of DNA structure are inherently unstable. For example, the bond connecting a purine base to deoxyribose is prone to hydrolysis at a low rate under physiological conditions, leaving a sugar without an attached base. Thus coding information is lost on that strand, and this loss can lead to a mutation during DNA replication. Normal cellular reactions, including the movement of electrons along the electron-transport chain in mitochondria and lipid oxidation in peroxisomes

(see Chapter 12), produce several chemicals that react with and damage DNA, including hydroxyl radicals and superoxide (O_2^-). These chemicals can also cause chemical damage to DNA that can lead to mutations, including mutations that lead to cancers.

Many spontaneous mutations are **point mutations**, which involve a change in a single base pair in the DNA sequence. If a point mutation falls within the protein-coding region of a gene, the mutation may change one codon to another, thus changing the amino acid sequence of the encoded protein. Another possibility is that the mutation may introduce a stop codon in the protein sequence, leading to early termination. Point mutations outside of the coding sequence can also interfere with gene function by, for example, altering the regulation of a gene's transcription, as discussed in Chapter 8. One of the most frequent causes of point mutations is *deamination* of a cytosine (C) base, which converts it into a uracil (U) base. Another is deamination of the common modified base 5-methylcytosine, which forms thymine when it is deaminated. If these alterations are not corrected before the DNA is replicated, the cell will use the strand containing U or T as a template to form a U·A or T·A base pair, thus creating a permanent change in the DNA sequence that is a mutation (Figure 5-14).

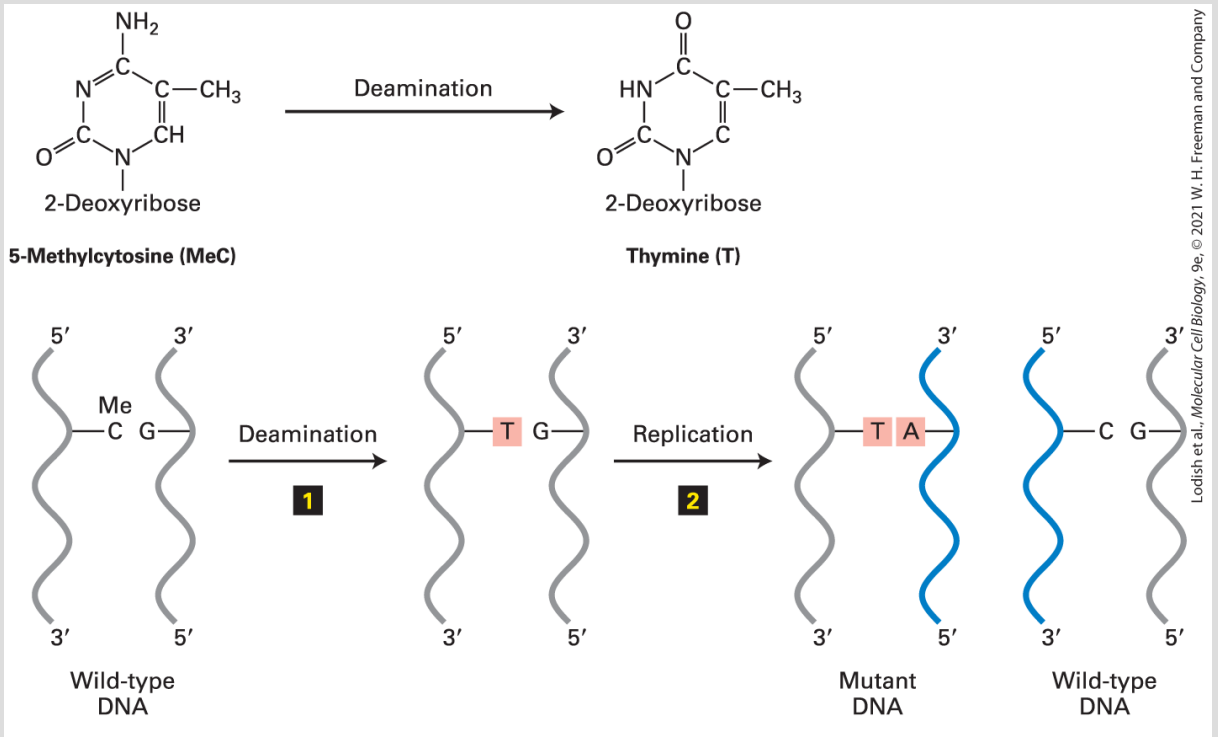


FIGURE 5-14 Deamination leads to point mutations. A spontaneous point mutation can be caused by deamination of 5-methylcytosine (C) to form thymine (T). If the resulting T·G base pair is not restored to the normal C·G base pair by base excision-repair mechanisms (step 1), it will lead to a permanent change in sequence (i.e., a mutation) following DNA replication (step 2). After one round of replication, one daughter DNA molecule will have the mutant T·A base pair and the other will have the wild-type C·G base pair.

High-Fidelity DNA Excision-Repair Systems Recognize and Repair Damage

DNA excision-repair systems were first elucidated through a combination of genetic and biochemical studies in *E. coli*. Homologs of the key bacterial proteins exist in eukaryotes from yeasts to

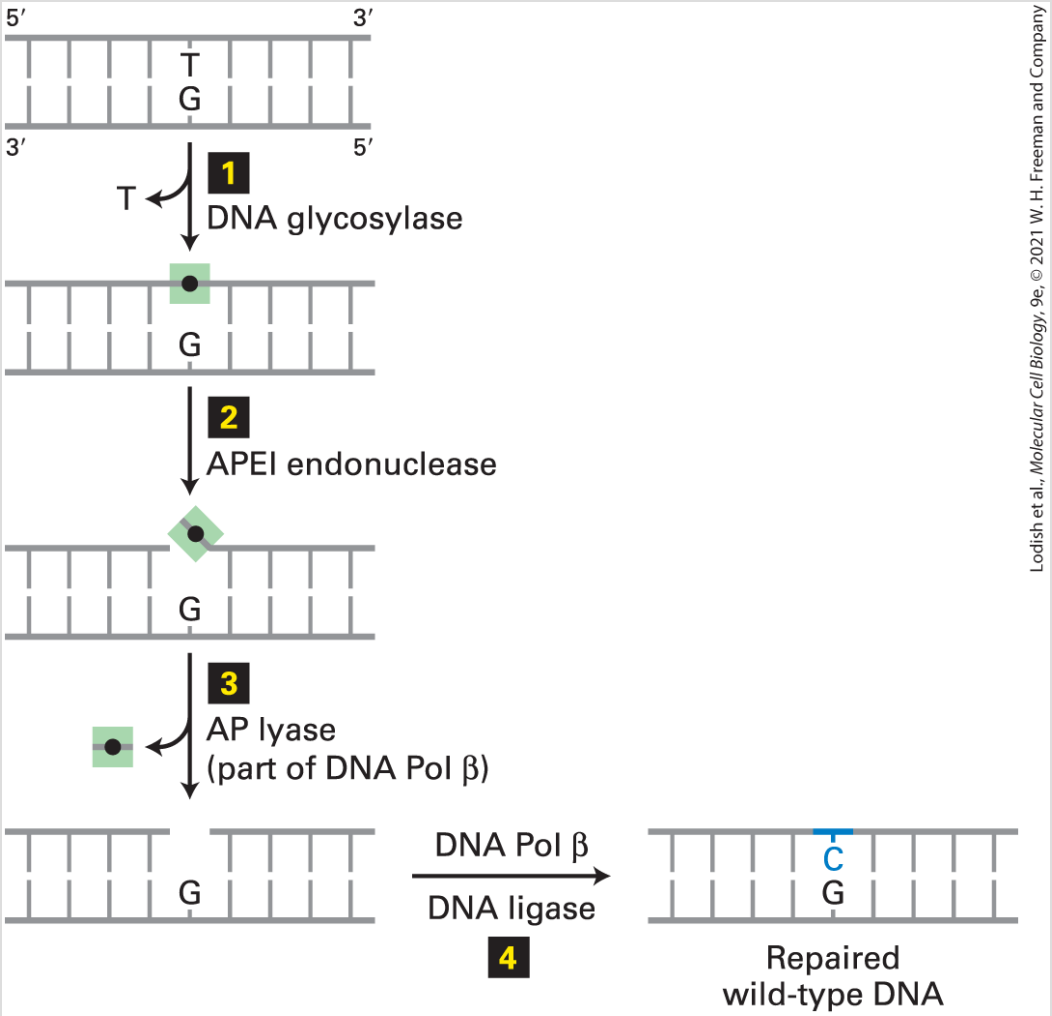
humans, indicating that these error-free mechanisms arose early in evolution to protect DNA integrity. Each of these systems functions in a similar manner: the damaged DNA strand is recognized, a segment of the damaged DNA strand is excised, and the gap is filled by DNA polymerase and ligase using the complementary DNA strand as a template.

Base Excision Repairs T-G Mismatches and Damaged Bases

In humans, the most common type of point mutation is a change from a C to a T, which is caused by deamination of 5-methyl C (see Figure 5-14). The conceptual problem with correcting a lesion in DNA caused by chemically normal bases that are mismatched is in determining which one is the normal and which one is the mutant DNA strand. But since a G·T mismatch is almost invariably caused by chemical conversion of C to U or 5-methyl C to T, the repair system evolved to selectively remove the T and replace it with a C.

The G·T mismatch is recognized by a DNA glycosylase that flips the thymine base out of the helix and then hydrolyzes the bond that connects it to the sugar-phosphate DNA backbone. Following this initial incision, an endonuclease, APE1, cuts the DNA strand near the now abasic site. The deoxyribose phosphate lacking the base is

then removed and replaced with a C by a specialized DNA polymerase that reads the G in the template strand (Figure 5-15).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 5-15 Base excision repair of a T·G mismatch. A DNA glycosylase specific for G·T mismatches, which are usually formed by deamination of 5-methyl C (see Figure 5-34), flips the thymine base out of the helix and then cuts it away from the sugar-phosphate DNA backbone (step 1), leaving just the deoxyribose phosphate (black dot). An endonuclease specific for the resultant abasic site (apurinic endonuclease I, APE1) then cuts the DNA backbone (step 2), and the deoxyribose phosphate is removed by an endonuclease, apurinic lyase (AP lyase) associated with DNA polymerase β, a specialized DNA polymerase used in repair (step 3). The gap is then filled in by DNA Pol β and

sealed by DNA ligase (step **4**, restoring the original G·C base pair. See O. Schärer, 2003, *Angewandte Chemie* **42**:2946.

As mentioned earlier, this repair must take place prior to DNA replication because the incorrect base in this pair, T, occurs naturally in normal DNA. Consequently, it would be able to engage in normal Watson-Crick base pairing during replication, generating a stable point mutation that would no longer be recognized by repair mechanisms (see Figure 5-14, step **2**).

Human cells contain a battery of glycosylases, each of which is specific for a different set of chemically modified DNA bases. For example, one glycosylase removes 8-oxyguanine, an oxidized form of guanine, allowing its replacement by an undamaged G, and others remove bases modified by alkylating agents. The resulting abasic nucleotide is then replaced by the repair mechanism described above. A similar mechanism functions in the repair of lesions resulting from *depurination*, the loss of a guanine or adenine base from DNA caused by hydrolysis of the glycosylic bond between deoxyribose and the base. Depurination occurs spontaneously and is fairly common in mammals and birds because of their warm body temperatures. The resulting abasic sites, if left unrepaired, generate mutations during DNA replication because they cannot specify the appropriate paired base.

Mismatch Excision Repairs Other Mismatches and Small Insertions and Deletions

Another process, also conserved in organisms from bacteria to humans, principally eliminates base-pair mismatches and insertions or deletions of one or a few nucleotides that are accidentally introduced by DNA polymerases during replication. As with base excision repair of a T in a T-G mismatch, the conceptual problem with *mismatch excision repair* is determining which is the normal and which is the mutant DNA strand. In *E. coli*, this discrimination is achieved by preferentially correcting the base in the most recently synthesized strand of DNA which is the most likely to be incorrect. Sensing of the relative age of the two DNA strands takes advantage of a methylation reaction that occurs in the context of double-stranded DNA. Because methylation lags after DNA synthesis, the newly synthesized strand can be identified by it having relatively less methylation. Human cells are also able to discriminate the newly synthesized strand from the template strand immediately after DNA replication. However, this discrimination does not appear to be based on methylation of the DNA, but instead probably involves recognition of the 3′ end of the newly synthesized strand. Once identified, the mispaired segment of the daughter strand — the one with the replication error — is excised and repaired to produce an exact complement of the template strand (Figure 5-

16). In contrast to base excision repair, mismatch excision repair occurs after DNA replication.

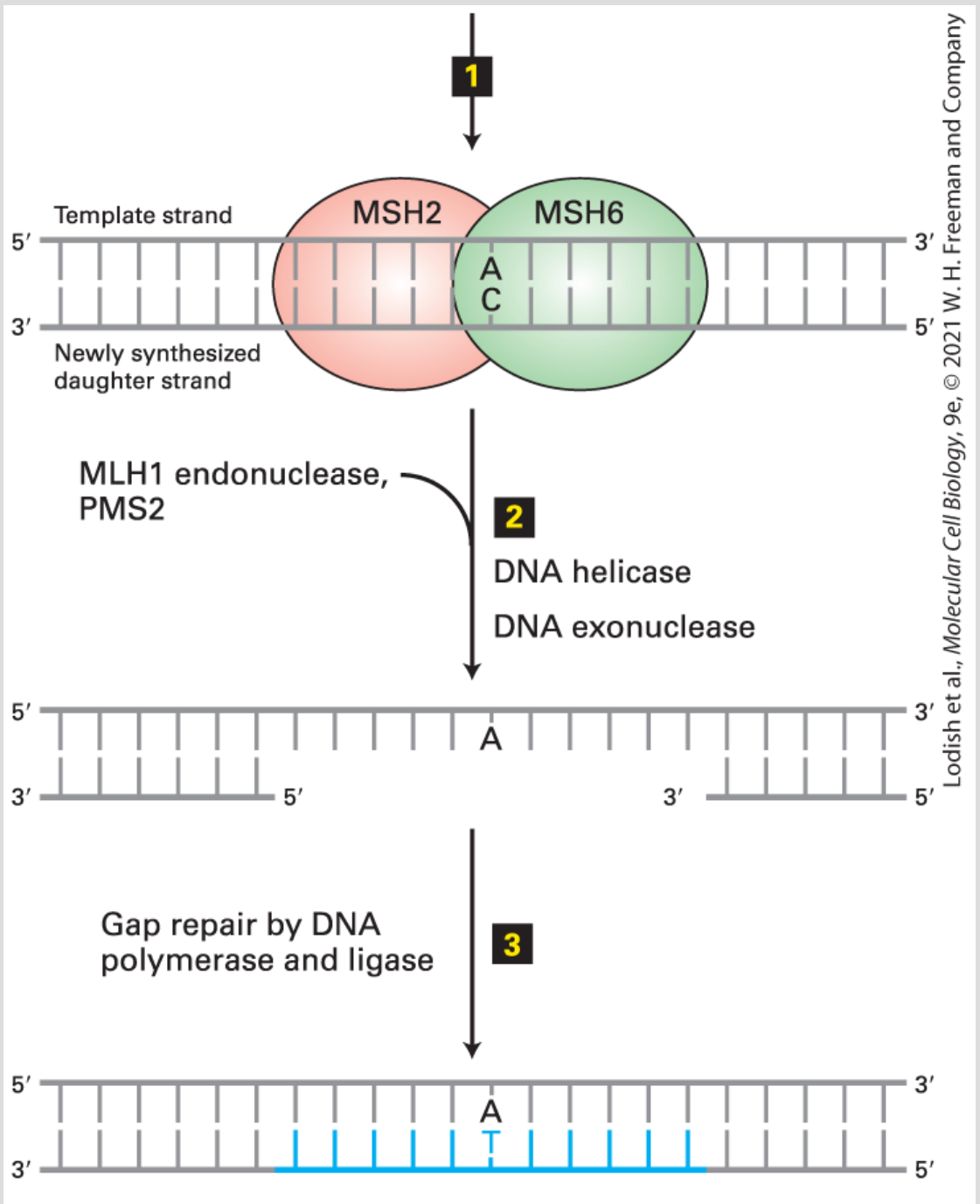


FIGURE 5-16 Mismatch excision repair in human cells. The mismatch excision-repair pathway corrects errors introduced during replication. A complex of the MSH2

and MSH6 proteins (bacterial *MutS homologs* 1 and 6) binds to a mispaired segment of DNA in such a way as to distinguish between the template and the newly synthesized daughter strand (step 1). This binding triggers binding of MLH1 and PMS2 (both homologs of bacterial MutL). The resulting DNA-protein complex then binds an endonuclease that cuts the newly synthesized daughter strand. Next a DNA helicase unwinds the helix, and an exonuclease removes several nucleotides from the cut end of the daughter strand, including the mismatched base (step 2). Finally, as with base excision repair, the gap is filled in by a DNA polymerase (Pol δ, in this case) and sealed by DNA ligase (step 3).

Predisposition to a colon cancer known as hereditary nonpolyposis colorectal cancer results from an inherited loss-of-function mutation in one copy of either the *MLH1* or the *MSH2* gene. The MSH1 and MLH2 proteins are essential for DNA mismatch repair (see Figure 5-16). Cells with at least one functional copy of each of these genes exhibit normal mismatch repair. However, tumor cells frequently arise from individual cells that have experienced a random mutation in the second copy; when both copies of one gene have been inactivated, the mismatch repair system is no longer functional. Inactivating mutations in these genes are also common in noninherited forms of colon cancer. ■

Nucleotide Excision Repairs Chemical Adducts That Distort Normal DNA Shape

Cells use *nucleotide excision repair* to fix DNA regions containing chemically modified bases, often called *chemical adducts*, that distort the normal shape of DNA locally. A key to this type of repair is the ability of certain proteins to slide along the surface of a double-stranded DNA molecule looking for bulges or other irregularities in the shape of the double helix. For example, this mechanism repairs *thymine-thymine dimers*, a common type of chemical modification caused by UV light (Figure 5-17); these dimers interfere with both replication and transcription of DNA. Figure 5-18 illustrates how the nucleotide excision-repair system repairs damaged DNA.

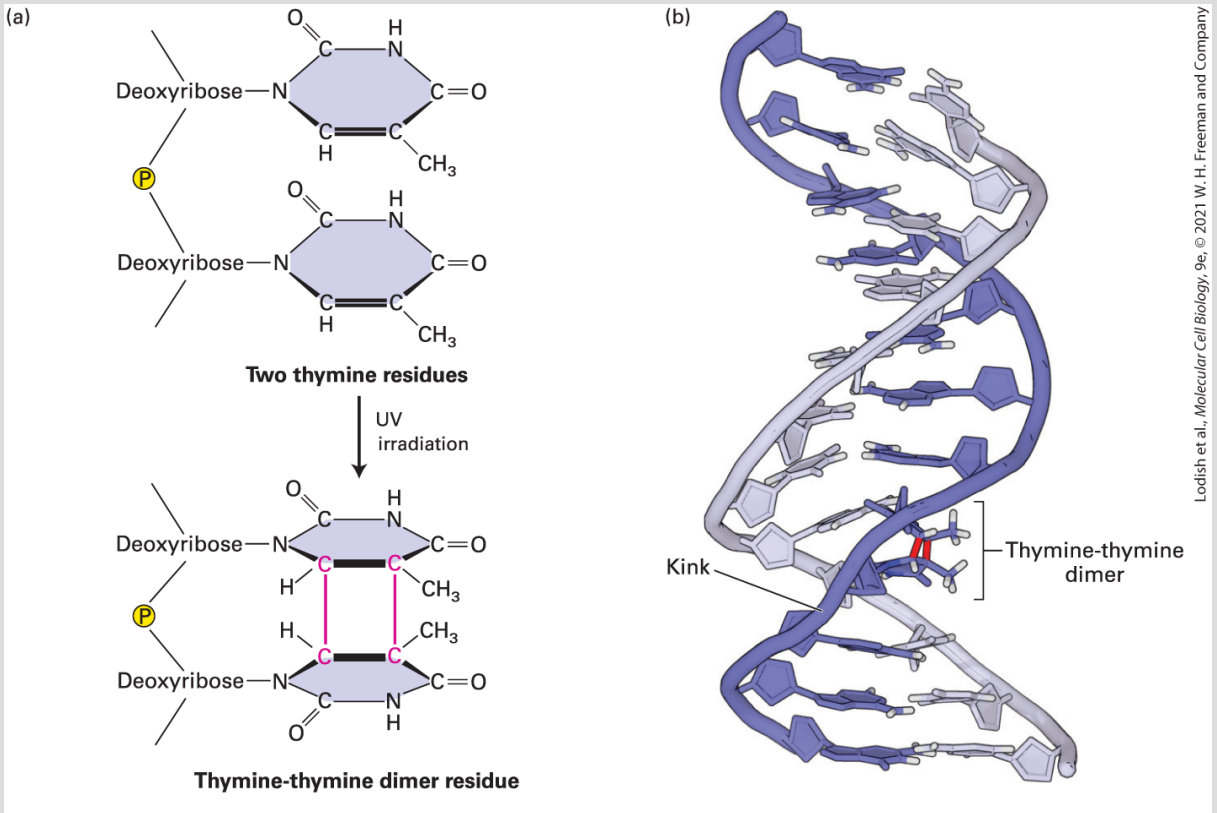
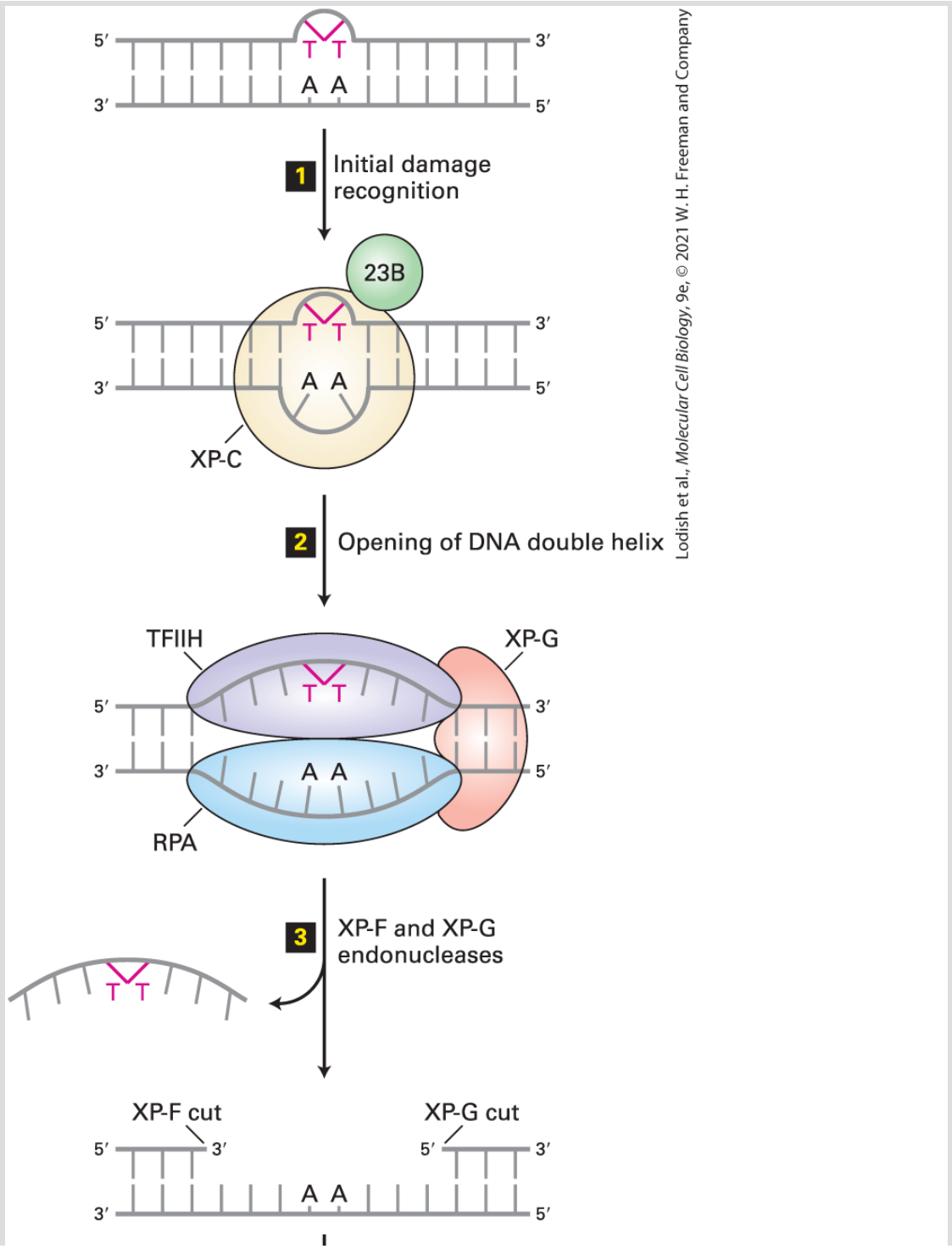


FIGURE 5-17 Formation of thymine-thymine dimers. (a) The most common type of DNA damage caused by UV irradiation is the formation of thymine-thymine dimers. (b) These lesions can be repaired by an excision-repair mechanism that recognizes the distortion they

create in the shape of the DNA double helix. The red lines in (b) represent the UV-induced C — C bonds shown in (a).

[Part (b) Data from K. McAteer et al., 1998, *J. Mol Biol*. **282**:1013, PDB ID 1ttd.]



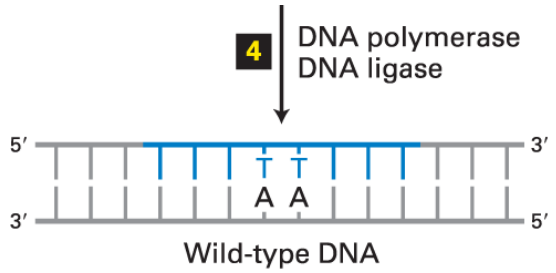


FIGURE 5-18 Nucleotide excision repair in human cells. A DNA lesion that causes distortion of the double helix, such as a thymine-thymine dimer, is initially recognized by a complex of the XP-C (*x*eroderma *p*igmentosum *C* protein) and 23B proteins (step **1**). This complex then recruits transcription factor TFIIH, whose helicase subunits, powered by ATP hydrolysis, partially unwind the double helix. XP-G and RPA proteins then bind to the complex and further unwind and stabilize the helix until a bubble of about 25 bases is formed (step **2**). Then XP-G (now acting as an endonuclease) and XP-F, a second endonuclease, cut the damaged strand at points 24–32 bases apart on each side of the lesion (step **3**). This releases the DNA fragment with the damaged bases, which is degraded to mononucleotides. Finally the gap is filled by DNA polymerase exactly as in DNA replication, and the remaining nick is sealed by DNA ligase (step **4**). See J. Hoeijmakers, 2001, *Nature* **411**:366, and O. Schärer, 2003, *Angewandte Chemie* **42**:2946.

Consider what happens to a thymine-thymine dimer not in the context of a double helix, but in a single strand of DNA at a replication fork. Such an unrepaired lesion of the template would not allow formation of a normal Watson-Crick base pair and the progress of a replicative polymerase would stall. In such events, progress of the replication fork is restored by replacement of the high-fidelity replicative polymerase with a low-fidelity *translesion polymerase*.

A replication fork stalled at an unrepaired thymine-thymine dimer on the template strand, produces a signal that leads to ubiquitylation of the sliding clamp PCNA (see Figure 5-13), which in turn triggers the replacement of a normal replicative polymerase such as Pol δ or Pol ε with the translesion polymerase, Pol η. Pol η has an active site that can accept distorted base pairs and lacks proofreading $3' \rightarrow 5'$ exonuclease activity. These features allow Pol η to form nonstandard base interactions, permitting replication past a thymine-thymine dimer on the template, but also render the polymerase prone to making errors during the replication of normal DNA. After replication has progressed past the lesion, the replication will eventually reset to the normal replicative polymerase. Ultimately, the process of replication past a lesion in the DNA will lead to a segment of DNA in the vicinity of the lesion that is relatively likely to contain mutations caused by errors in replication by translesion polymerase. This explains why mutagenesis with UV light causes a wide variety of base changes in the vicinity of T-T dinucleotides.

Some 30 proteins are involved in the nucleotide excision-repair process, the first of which were identified through a study of the defects in DNA repair in cultured cells from individuals with **xeroderma pigmentosum**, a hereditary disease associated with a predisposition to cancer. Individuals with this disease frequently develop skin cancers called *melanomas* and *squamous cell carcinomas* if their skin is exposed to the UV rays in sunlight. Mutations in any of at least seven different genes, called *XP-A* through *XP-G*, cause

sensitivity to UV by inactivation of components of the nucleotide excision-repair system depicted in Figure 5-18. A second type of xeroderma pigmentosum mutation, which also causes sensitivity to UV rays, called *XP-V,* is in the Pol η gene and prevents replication past UV-induced lesions. ■

Two Systems Use Recombination to Repair Double-Strand Breaks in DNA

Ionizing radiation (e.g., x- and γ-radiation) and some anticancer drugs cause double-strand breaks in DNA. The failure to repair such a lesion would lead to the loss of the portion of the chromosome that was distal to the site of the break and in most cases would be lethal. Two systems have evolved to repair double-strand breaks: breaks can be precisely repaired using information from the intact homologous chromosome by a process of *homologous recombination*, discussed in the next section. Failing precise repair, double-strand breaks can be repaired by the error-prone mechanism of *nonhomologous end joining* (*NHEJ*).

Error-Prone Repair by Nonhomolohous End Joining

In the **nonhomologous end joining** pathway, the free broken ends of DNA are prevented from further degradation by binding to a

complex of two proteins, Ku and DNA-PK. Two Ku and DNA-PK heterodimers bound to DNA can then come together and, with the assistance of additional proteins including DNA ligase, join the ends of two DNA molecules (Figure 5-19). Since diffusion of DNA within the viscous nucleoplasm is fairly slow, the ends of a broken chromosome can often be rejoined before they diffuse apart. Even when correct ends are rejoined, the inevitable action of exonucleases will result in the loss of several base pairs at the joining point. If a deletion resulting from inexact end joining occurs within the coding sequence, the function of the affected gene will likely be lost.

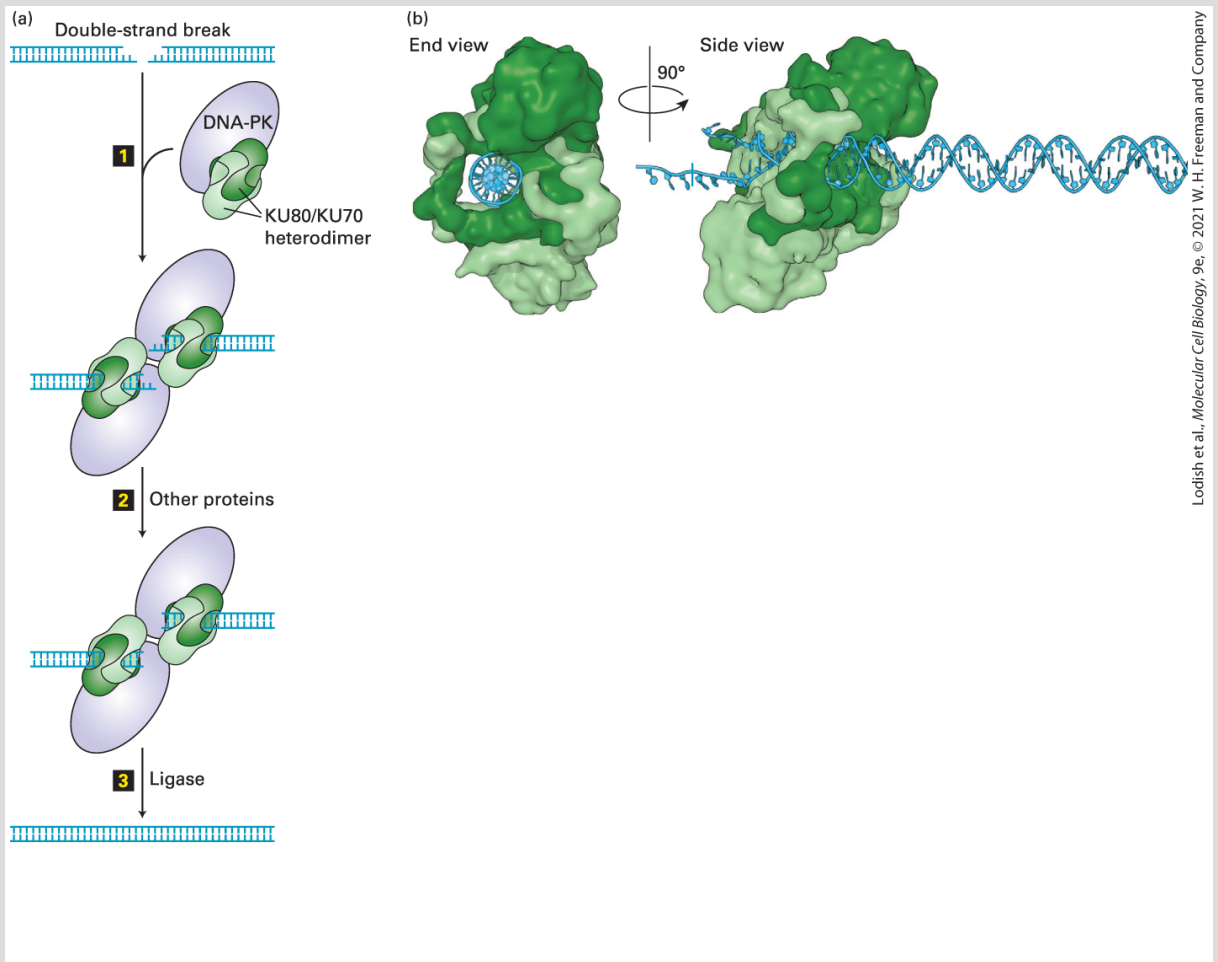


FIGURE 5-19 Nonhomologous end joining. (a) When sister chromatids are not available to help repair double-strand breaks, nucleotide sequences are butted together that were not apposed in the unbroken DNA. These DNA ends are usually from the same chromosome locus and, when linked together, several base pairs are lost. Occasionally, ends from different chromosomes are accidentally joined together. A complex of two proteins, Ku and DNA-dependent protein kinase (DNA-PK), binds to the ends of a double-strand break (step **1**). After formation of a synapse, the ends are further processed by nucleases, resulting in removal of a few bases (step **2**), and the two double-stranded molecules are ligated together (step **3**). As a result, the double-strand break is repaired, but several base pairs at the site of the break are removed. See G. Chu, 1997, *J. Biol. Chem.* **272**:24097; M. Lieber et al., 1997, *Curr. Opin. Genet. Devel.* **7**:99; and D. van Gant et al., 2001, *Nat. Rev. Genet*. **2**:196. (b) Structure of KU70/KU80 bound to a duplex DNA end. The complex is shown in a view down the DNA axis (*left*) and from the side (*right*). KU80 is light green, KU70 dark green, DNA blue.

[Part (b) Data from J. R. Walker, R. A. Corpina, and J. Goldberg, 2001, *Nature* **412**:607, PDB ID 1jey.]

Occasionally, however, broken ends from different chromosomes or far apart on the same chromosome are joined together, leading to gross chromosomal rearrangements that can affect the functioning of genes. For example, incorrect joining could create a hybrid gene that codes for the N-terminal portion of one amino acid sequence fused to the C-terminal portion of a completely different amino acid sequence; or a chromosomal rearrangement could bring the promoter of one gene close to the coding region of another gene, changing the level of expression or cell type in which that gene is expressed. Such translocations may generate chimeric genes that can have drastic effects on normal cell function, such as uncontrollable cell growth, which is the hallmark of cancer.

Homologous Recombination Can Repair DNA Damage and Generate Genetic Diversity

A variety of DNA lesions that are not repaired by the mechanisms discussed earlier can be repaired by mechanisms in which the damaged DNA sequence is copied from an undamaged copy of the same or a highly homologous sequence on the homologous chromosome in diploid organisms or the sister chromosome following DNA replication in haploid and diploid organisms. These mechanisms involve an exchange of strands between separate DNA molecules and hence are referred to as **recombination**.

In addition to providing a mechanism for DNA repair, similar recombination mechanisms generate genetic diversity among the individuals of a species by causing the exchange of large regions of chromosomes between the maternal and paternal pair of homologous chromosomes during *meiosis*, the special type of cell division that generates germ cells (sperm and eggs) (see Figure 6-3). In fact, the exchange of regions of homologous chromosomes, called *crossing over*, is required for proper segregation of chromosomes during the first meiotic cell division. Meiosis and the consequences of generating new combinations of maternal and paternal genes by recombination are discussed further in Chapter 6, and the mechanisms leading to proper segregation of chromosomes during meiosis are discussed in Chapter 19. Here we focus on the

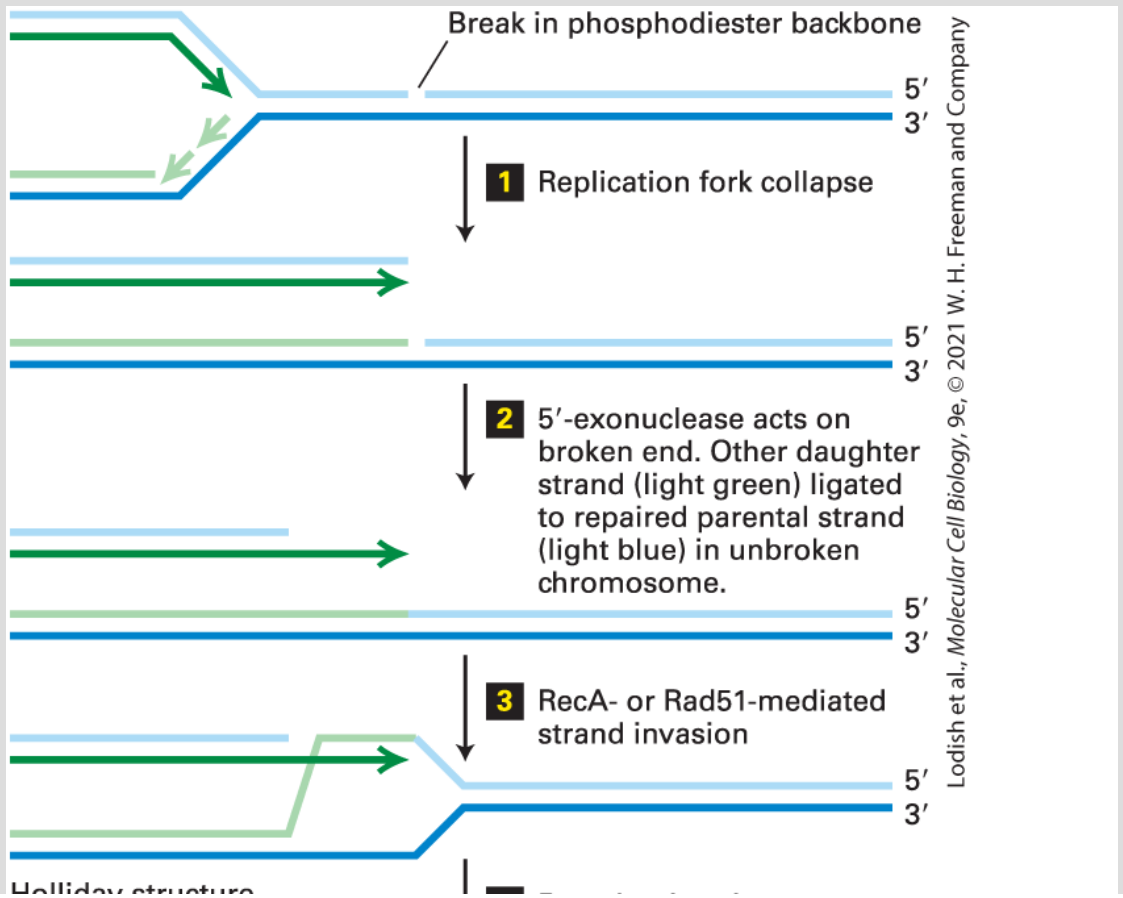
molecular mechanisms of DNA recombination, highlighting the exchange of DNA strands between two recombining DNA molecules.

At one time, homologous recombination was thought to be a minor repair process in human cells. This view changed when it was realized that several human cancers are potentiated by inherited mutations in genes that are essential for homologous recombination repair. For example, some women with an inherited susceptibility to breast cancer have a mutation in one allele of either the *BRCA1* or the *BRCA2* gene, both of which encode proteins participating in this repair process. Loss or inactivation of the second allele inhibits the homologous recombination repair pathway and thus tends to induce cancer in mammary or ovarian epithelial cells. Yeasts can use homologous recombination to repair double-strand breaks induced by γ-irradiation. Isolation and analysis of radiation-sensitive (*RAD*) mutants that are deficient in this repair system facilitated the study of this process. Most of the yeast Rad proteins have homologs in the human genome, and the human and yeast proteins function in an essentially identical fashion.

Repair of a Collapsed Replication Fork

An example of recombinational DNA repair is the repair of a replication fork that has come apart, known as a collapsed replication fork. If a break in the phosphodiester backbone (a nick) of one DNA strand is not repaired before a replication fork passes,

the replicated portions of the daughter chromosomes become separated when the replication helicase reaches the nick in the parent strand because there are no covalent bonds between the two fragments of the parent strand on either side of the nick. This process is called *replication fork collapse* (Figure 5-20, step **1**). If the break in the double-stranded daughter DNA molecule is not repaired, it is generally lethal to at least one daughter cell following cell division because of the loss of genetic information between the break and the end of the chromosome. The recombination process that repairs the resulting double-strand break and regenerates a replication fork involves multiple enzymes and other proteins, only some of which are mentioned here.



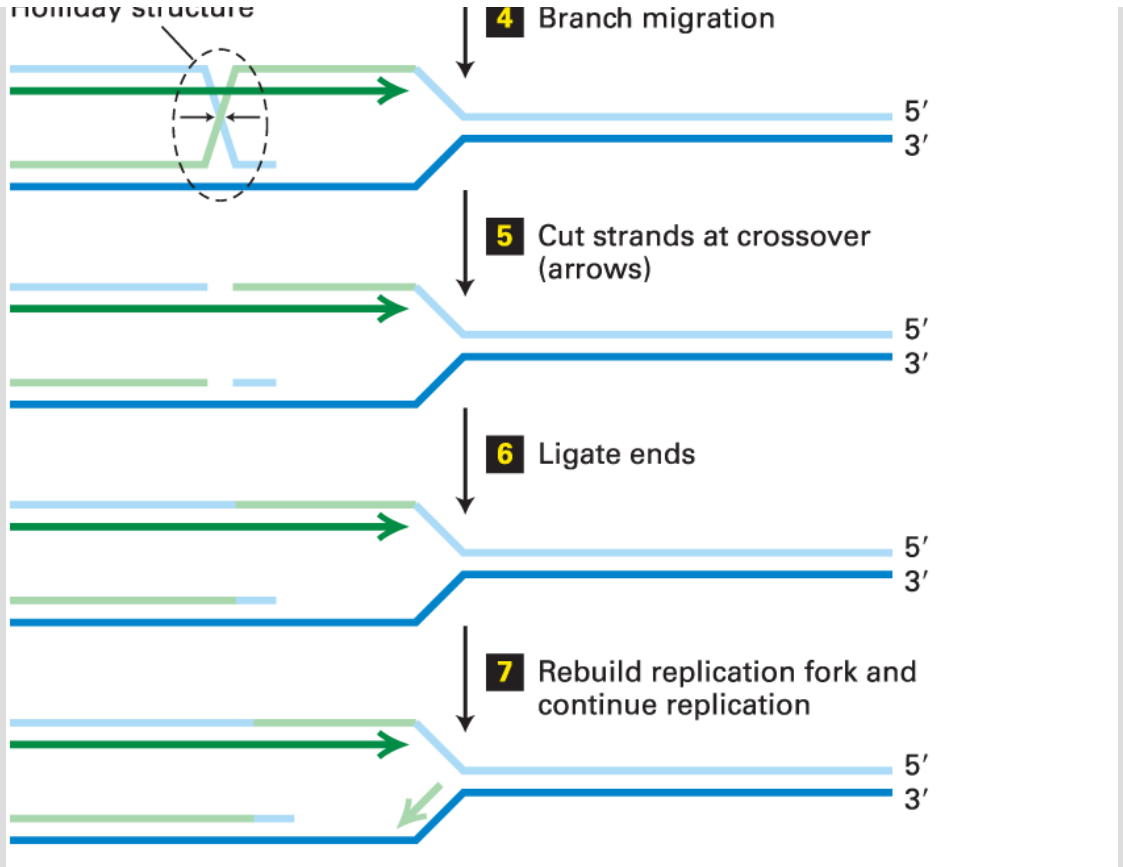


FIGURE 5-20 Recombinational repair of a collapsed replication fork. Parent strands are light and dark blue. The leading daughter strand is dark green, and the lagging daughter strand light green. Diagonal lines in step 3 and beyond represent a single phosphodiester bond from the DNA strand of the corresponding color. Small black arrows following step 4 represent cleavage of the phosphodiester bonds at the crossover of DNA strands in the Holliday structure. See the text for a discussion.

The first step in the repair of the double-strand break is exonucleolytic digestion of the strand whose 5′ end is at the broken end of DNA, leaving the strand with the 3′ end at the break single stranded (Figure 5-20, step 2). The lagging nascent strand (light green) base-paired to the unbroken parent strand (dark blue) is ligated to the unreplicated portion of the parent chromosome (light blue), as shown in Figure 5-20, step 2. A protein required for the

next step is RecA in bacteria, or the homologous Rad51 in *Saccharomyces cerevisiae* and other eukaryotes. Multiple RecA/Rad51 molecules bind to the single-stranded DNA and catalyze its hybridization to a perfectly or nearly perfectly complementary sequence in another, homologous, double-stranded DNA molecule. The complementary strand of this target double-stranded DNA (dark blue) is displaced as a single-stranded loop of DNA over the region of hybridization to the invading strand (Figure 5-20, step **3**). This RecA/Rad51–catalyzed *invasion* of a duplex DNA by a single-stranded complement of one of the strands is key to the recombination process. This process is called **strand invasion**, and because there is no change in the number of base pairs, it does not require an input of energy.

Next the hybrid region between target DNA and the invading strand is extended away from the break by a process called branch migration (Figure 5-20, step **4**). In this diagram, the diagonal lines represent only one phosphodiester bond. Molecular modeling and other studies show that the first base on either side of the branch is base-paired to a complementary nucleotide. As this branch *migrates* to the left, the number of base pairs remains constant; one new base pair formed with the invading strand (dark green) is matched by the loss of one base pair with the parent strand (blue).

When the hybrid region extends beyond the 5′ end of the broken strand that was digested by the 5′-exonuclease in step **2** (light

blue), the single-stranded parent DNA thus generated (light blue) can base-pair with the complementary region of the other parent strand (dark blue), which becomes single-stranded as the branch migrates to the left (see Figure 5-20, step **4**). The resulting structure is called a **Holliday structure**, after Robin Holliday, the geneticist who first proposed it as an intermediate in genetic recombination. Again, the diagonal lines in the diagram following step **4** represent single phosphodiester bonds, and all bases in the Holliday structure are base-paired to complementary bases in the parent strands. Cleavage of the phosphodiester bonds that *cross over* from one parent strand to the other (step **5**) and ligation of the 5′ and 3′ ends base-paired to the same parent strands (step **6**) result in the generation of a structure similar to a replication fork. Rebinding of replication fork proteins results in extension of the leading strand past the point of the original strand break and re-initiation of lagging-strand synthesis (step **7**), thus regenerating a replication fork. The overall process allows the ligated upper strand in the lower molecule following step **2** to serve as a template for extension of the leading strand in step **7**.

Repair of a Double-Strand Break by Homologous Recombination

A similar mechanism, called **homologous recombination**, can repair a double-strand break in a chromosome and can also exchange large segments of two double-stranded DNA molecules

(Figure 5-21). First, the broken ends of the DNA molecule are digested by 5′-exonucleases, leaving a single-stranded region of DNA with a 3′ end (step **1**). RecA in bacteria or Rad51 in eukaryotes then catalyzes invasion of one of these 3′ ends into the homologous region of the homologous chromosome, as described above for repair of a collapsed replication fork (step **2**). The 3′ end of the invading DNA strand is then extended by a DNA polymerase, displacing the parent strand as an enlarging single-stranded loop of DNA (dark blue) (step **3**). When the loop extends to a sequence that is complementary to the other broken and 5′-exonuclease–digested end of DNA (the fragment on the left following step **1**), the complementary sequences base-pair (diagram following step **3**). This 3′ end is then extended by a DNA polymerase using the displaced single-stranded loop of parent DNA (dark blue) as a template (step **4**).

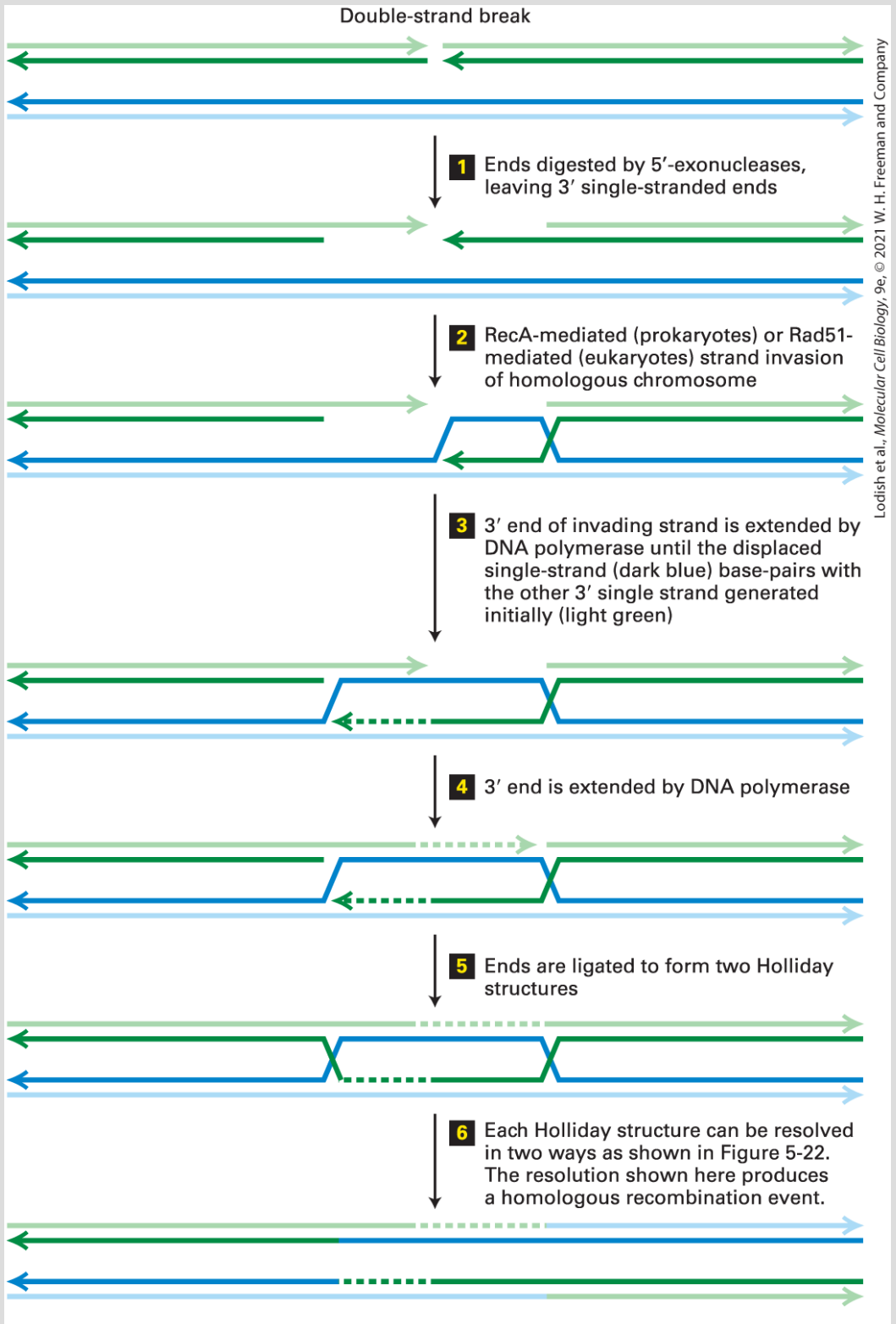


FIGURE 5-21 Repair of a double-strand break by homologous recombination. For simplicity, each DNA double helix is represented by two

parallel lines with the polarities of the strands indicated by arrowheads at their $3'$ ends. The upper molecule has a double-strand break. Note that in the diagram of the upper DNA molecule, the strand with its $3'$ end at the right is on the top, while in the diagram of the lower DNA molecule, this strand is drawn on the bottom. See the text for discussion. See T. L. Orr-Weaver and J. W. Szostak, 1985, *Microbiol. Rev.* **49**:33.

Next the new $3'$ ends are ligated (step **5**) to the exonuclease-digested $5'$ ends. This generates two Holliday structures in the paired molecules (step **5**). Branch migration of these Holliday structures can occur in either direction (not diagrammed). Finally, cleavage of the strands at the positions shown by the arrows, and ligation of the alternative $5'$ and $3'$ ends at each cleaved Holliday structure, generates two *recombinant* chromosomes that contain the DNA of one *parent* DNA molecule on one side of the initial break point (light and dark green strands), and the DNA of the other parent DNA molecule on the other side of the break point (light and dark blue) (step **6**). The region in the immediate vicinity of the initial break point forms a *heteroduplex*, in which one strand from one parent is base-paired to the complementary strand of the other parent (light or dark green strand base-paired to light or dark blue strand). Base-pair mismatches between the two parent strands are usually repaired by the repair mechanisms discussed above to generate a complementary base pair. In the process, sequence differences between the two parents are lost, a process referred to as **gene conversion**.

Figure 5-22 diagrams how cleavage of one or the other pair of strands at the four-way strand junction in the Holliday structure generates parent or recombinant molecules. This process, called *resolution* of the Holliday structure, separates DNA molecules initially joined by RecA/Rad51–catalyzed strand invasion. Each Holliday structure in the intermediate following Figure 5-21, step **5**, can be cleaved and religated in the two possible ways shown by the two sets of small black arrows in Figure 5-22 (step **1** or step **2**). Consequently, there are four possible products of the recombination process shown in Figure 5-21. After ligation of the cleaved ends, two of these products regenerate the parent chromosomes, with the exception of the heteroduplex region at the break point, which is repaired into the sequence of one parent or the other by a process known as gene conversion. The other two possible products generate recombinant chromosomes as shown in Figure 5-21.

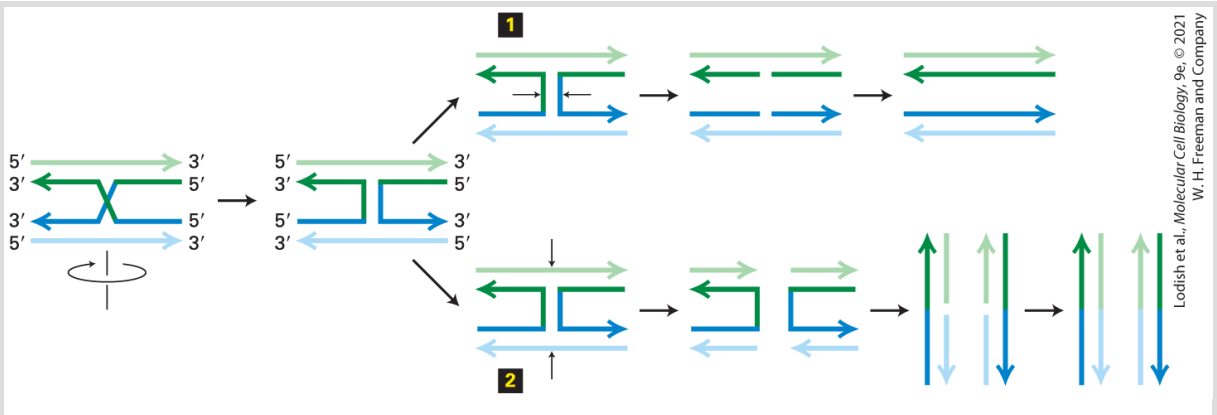


FIGURE 5-22 Alternative resolution of a Holliday structure. Diagonal and vertical lines represent a single phosphodiester bond. It is simplest to diagram the process by rotating the diagram of the bottom molecule 180° so that the top and bottom molecules have the same strand orientations. Cutting the bonds as shown in step **1** and ligating the ends as

indicated regenerates the original chromosomes. Cutting the strands as shown in step 2 and religating as shown at the bottom generates recombinant chromosomes.

KEY CONCEPTS OF SECTION 5.3

DNA Repair and Recombination

- Changes in the DNA sequence result from copying errors and the effects of various physical and chemical agents.
- Eukaryotic cells have three excision-repair systems for correcting mispaired bases and for removing chemical adducts from DNA. Base excision repair, mismatch repair, and nucleotide-excision repair operate with high accuracy and generally do not introduce errors.
- Damaged DNA bases in single-stranded template DNA at a replication fork will block the progress of a replicative DNA polymerase. These blocks can be overcome by a translesion polymerase that has relaxed requirements for proper base pairing during replication. Translesion polymerases are error prone and tend to generate mutations of their own in the vicinity of the original lesion.
- Repair of double-strand breaks by the nonhomologous end joining pathway can link segments of DNA from different chromosomes, possibly forming an oncogenic chromosomal translocation. This repair mechanism also produces a small deletion, even when segments from the same chromosome are joined.
- Error-free repair of double-strand breaks in DNA is accomplished by homologous recombination using the undamaged sister chromatid as a template.
- Inherited defects in the nucleotide excision-repair pathway, as in individuals with xeroderma pigmentosum, predispose them to skin cancer. Inherited colon cancer is frequently associated with mutant forms of proteins essential for the mismatch repair pathway. Defects in repair by homologous recombination are associated with inheritance of one mutant allele of the *BRCA1* or *BRCA2* gene and result in predisposition to breast and uterine cancer.

5.4 Transcription of Protein-Coding Genes and Formation of mRNA

The simplest definition of a *gene* is a unit of DNA that contains the information to specify synthesis of a single polypeptide chain or functional RNA (such as a tRNA). The DNA molecules of small viruses contain only a few genes, whereas the single DNA molecule in each of the chromosomes of higher animals and plants may contain several thousand genes. The vast majority of genes carry information used to build protein molecules, and it is the RNA copies of such *protein-coding genes* that constitute the mRNA molecules of cells.

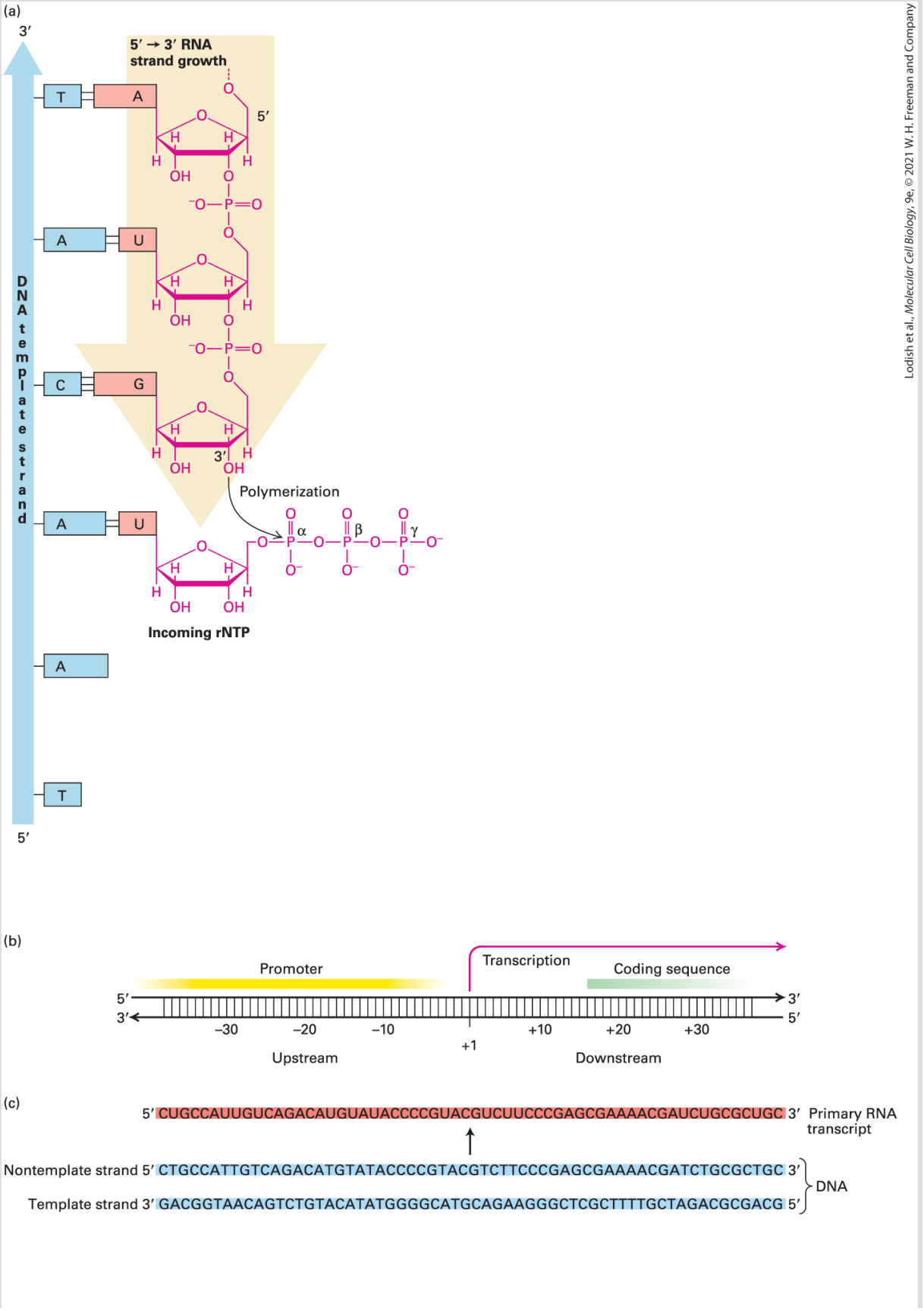
In its most elemental form, the process of transcription is the formation of an RNA copy of the information carried in DNA for one gene. RNA synthesis, catalyzed by an RNA polymerase and based on the same base-pairing mechanism as DNA replication copies, or *transcribes*, the four-base language of DNA containing A, G, C, and T into the four-base language of RNA, which is identical except that U replaces T. Because of the inherent instability of RNA within the cell, the amount of transcribed RNA can readily be regulated by alterations in the rate of transcription initiation. This allows distinct sets of genes to be expressed in the multiple different types of cells that make up a multicellular organism. Regulation of transcription is addressed in Chapter 8.

For protein encoding genes, the product of transcription is an mRNA, which is subsequently *translated* into the 20–amino acid language of proteins. In this section, we focus on the formation of functional mRNAs from protein-coding genes (see Figure 5-1, step **1**). A similar process yields precursors of rRNAs and tRNAs, encoded by rRNA and tRNA genes; these precursors are then further modified to yield functional rRNAs and tRNAs (see Chapters 8 and 9).

A Template DNA Strand Is Transcribed into a Complementary RNA Strand by RNA Polymerase

The process of transcription is closely related to that of DNA replication, in which one DNA strand acts as a *template* and, in the case of transcription, determines the order in which ribonucleoside triphosphate (rNTP) monomers are linked together to form a complementary RNA chain. Incoming rNTPs base-pair with complementary bases in the template DNA strand, and are then joined to a growing RNA strand in a polymerization reaction catalyzed by **RNA polymerase**. The polymerization reaction involves a nucleophilic attack by the 3′ oxygen in the growing RNA chain on the α phosphate of the next nucleotide precursor to be added, which results in the formation of a phosphodiester bond and the release of

pyrophosphate (PP_i). As with DNA replication, RNA molecules are always synthesized in the $5' \rightarrow 3'$ direction (Figure 5-23a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 5-23 RNA is synthesized $5' \rightarrow 3'$. (a) Polymerization of ribonucleotides by RNA polymerase during transcription. The ribonucleotide to be added at the $3'$ end of a growing RNA strand is specified by base pairing between the next base in the template DNA strand and the complementary incoming ribonucleoside triphosphate (rNTP). A phosphodiester bond is formed when RNA polymerase catalyzes a reaction between the $3'$ oxygen of the growing strand and the α phosphate of a correctly base-paired rNTP. RNA strands are always synthesized in the $5' \rightarrow 3'$ direction and are opposite in polarity to their template DNA strands. (b) The DNA nucleotide where RNA polymerase begins transcription is designated $+1$. The direction the polymerase travels on the DNA is termed *downstream*, and downstream bases are marked with positive numbers. The opposite direction is termed *upstream*, and upstream bases are marked with negative numbers. Some important gene features lie upstream of the transcription start site, including the promoter sequence that localizes RNA polymerase to the gene. (c) The DNA strand that is being transcribed is the template strand; its complement is the nontemplate strand. The RNA being synthesized is complementary to the template strand and is therefore identical with the nontemplate strand sequence, except with uracil in place of thymine.

By convention, the site on the DNA template at which RNA polymerase begins transcription is numbered $+1$ (Figure 5-23b). *Downstream* denotes the direction in which a template DNA strand is transcribed; *upstream* denotes the opposite direction. Nucleotide positions in the DNA sequence downstream from a start site are indicated by a positive (+) sign; those upstream, by a negative (−) sign. Because RNA is synthesized $5' \rightarrow 3'$, RNA polymerase moves down the template DNA strand in a $3' \rightarrow 5'$ direction. The newly synthesized RNA is complementary to the template DNA strand; therefore, it is identical to the nontemplate DNA strand, with uracil in place of thymine (Figure 5-23c).

Stages in Transcription

To carry out transcription, RNA polymerase performs several distinct functions, as depicted in Figure 5-24. During transcription *initiation*, RNA polymerase, with the help of initiation factors (called s-factors in bacteria and general transcription factors in eukaryotes), recognizes and binds to a specific sequence of double-stranded DNA called a **promoter** (step **1**). After binding, RNA polymerase and the initiation factors separate the DNA strands to make the bases in the template strand available for base pairing with the bases of the rNTPs that the RNA polymerase will polymerize (step **2**). Approximately 12–14 base pairs of DNA around the transcription start site are separated, which allows the template strand to enter the active site of the enzyme. The *active site* is where catalysis of phosphodiester bond formation between rNTPs that are complementary to the template strand takes place. The 12–14 base-pair region of melted DNA in the active site of the polymerase is known as the *transcription bubble*. Transcription initiation is considered complete when the first two ribonucleotides of an RNA chain are linked by a phosphodiester bond (step **3**).

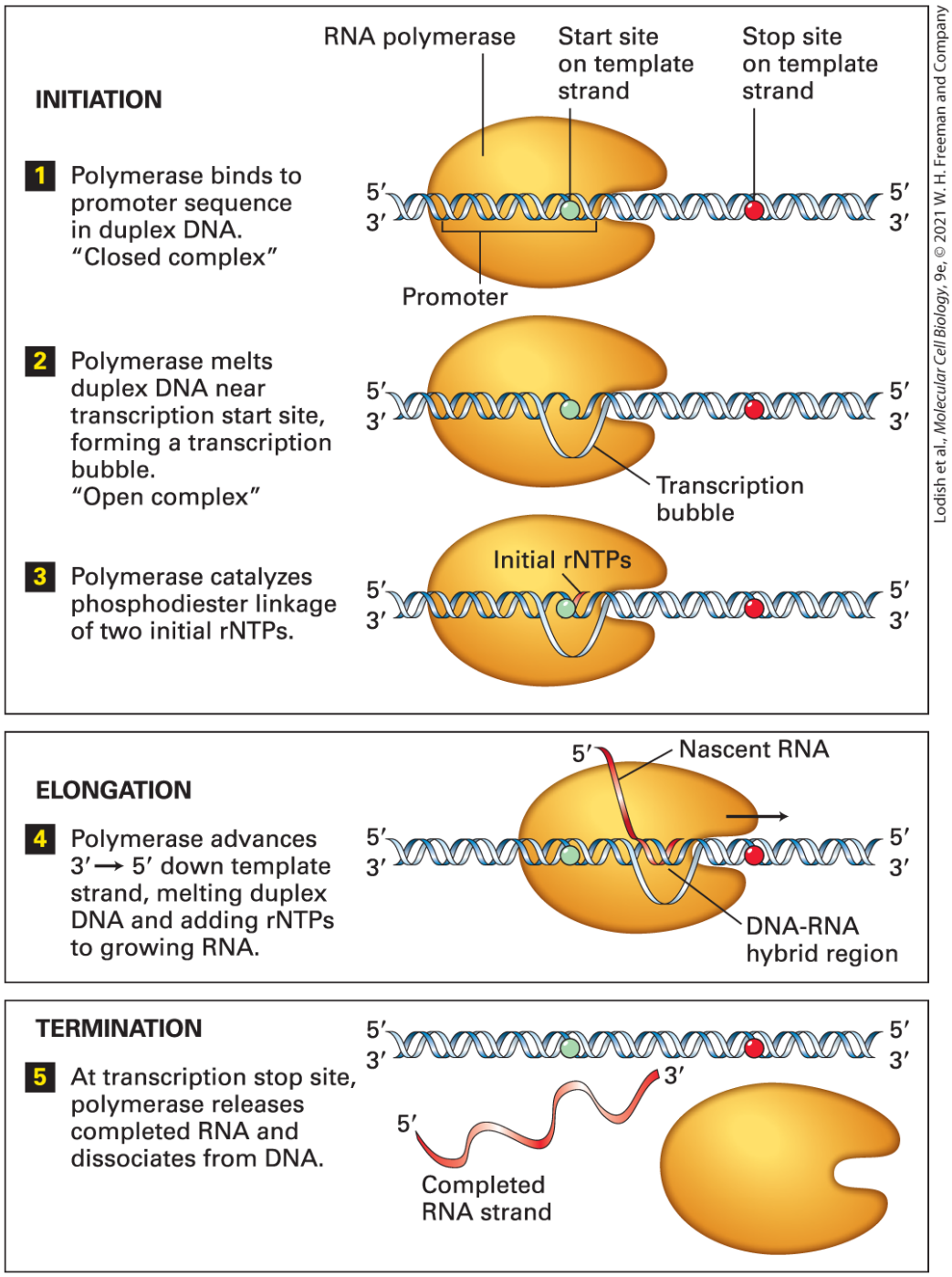


FIGURE 5-24 Three stages in transcription. During initiation of transcription, RNA polymerase forms a transcription bubble and begins polymerization of ribonucleotides (rNTPs) at the start site, which is located within the promoter region. Once a DNA region has been transcribed, the separated strands reassociate into a double helix. The nascent RNA is displaced from its template

strand except at its 3′ end. The 5′ end of the RNA strand exits the RNA polymerase through a channel in the enzyme. Termination occurs when the polymerase encounters a specific termination sequence (stop site). See the text for details. For simplicity, the diagram depicts transcription of four turns of the DNA helix encoding some 40 nucleotides of RNA. Most RNAs are considerably longer, requiring transcription of a longer region of DNA.

After several ribonucleotides have been polymerized, RNA polymerase dissociates from both the promoter DNA and the initiation factors (called σ-*factors* in bacteria, and *general transcription factors* in archaea and eukaryotes). During the *strand elongation* stage, RNA polymerase moves along the template DNA, opening the double-stranded DNA in front of its direction of movement and guiding the strands back together so that they reassociate at the upstream end of the transcription bubble (step **4**). One ribonucleotide at a time is added by the polymerase to the 3′ end of the growing (*nascent*) RNA chain. During strand elongation, the enzyme maintains a melted region of approximately 14–20 base pairs in the transcription bubble. Approximately eight nucleotides at the 3′ end of the growing RNA strand remain base-paired to the template DNA strand in the transcription bubble. The *elongation complex*, comprising RNA polymerase, template DNA, and the nascent RNA strand, can be extraordinarily stable. For example, RNA polymerase transcribes the longest known mammalian gene, containing about 2 million base pairs, without dissociating from the DNA template or releasing the nascent RNA. RNA synthesis occurs at a rate of about 1000–2000 nucleotides per minute, so the

elongation complex must remain intact for more than 24 hours to ensure continuous synthesis of pre-mRNA from this very long gene. Note that although biochemically similar to DNA replication, transcription of RNA occurs about 30 times more slowly in terms of the rate of nucleotides added.

During transcription *termination*, the final stage in RNA synthesis, the completed RNA molecule is released from the RNA polymerase and the polymerase dissociates from the template DNA (step **5**). Once it is released, an RNA polymerase is free to transcribe the same gene again or another gene.

Structure of RNA Polymerases

The RNA polymerases of bacteria, archaea, and eukaryotic cells are fundamentally similar in structure and function. Bacterial RNA polymerases are composed of two related large subunits (β' and β), two copies of a smaller subunit (α), and one copy of a fifth subunit (ω) that is not essential for transcription or cell viability, but that stabilizes the enzyme and assists in the assembly of its subunits. Archaeal and eukaryotic RNA polymerases have several additional small subunits associated with this core complex, which we describe in Chapter 8. Schematic diagrams of the transcription process generally show RNA polymerase bound to an unbent DNA molecule, as in Figure 5-24. However, x-ray crystallography and other studies

of an elongating bacterial RNA polymerase indicate that the DNA bends at the transcription bubble (Figure 5-25).

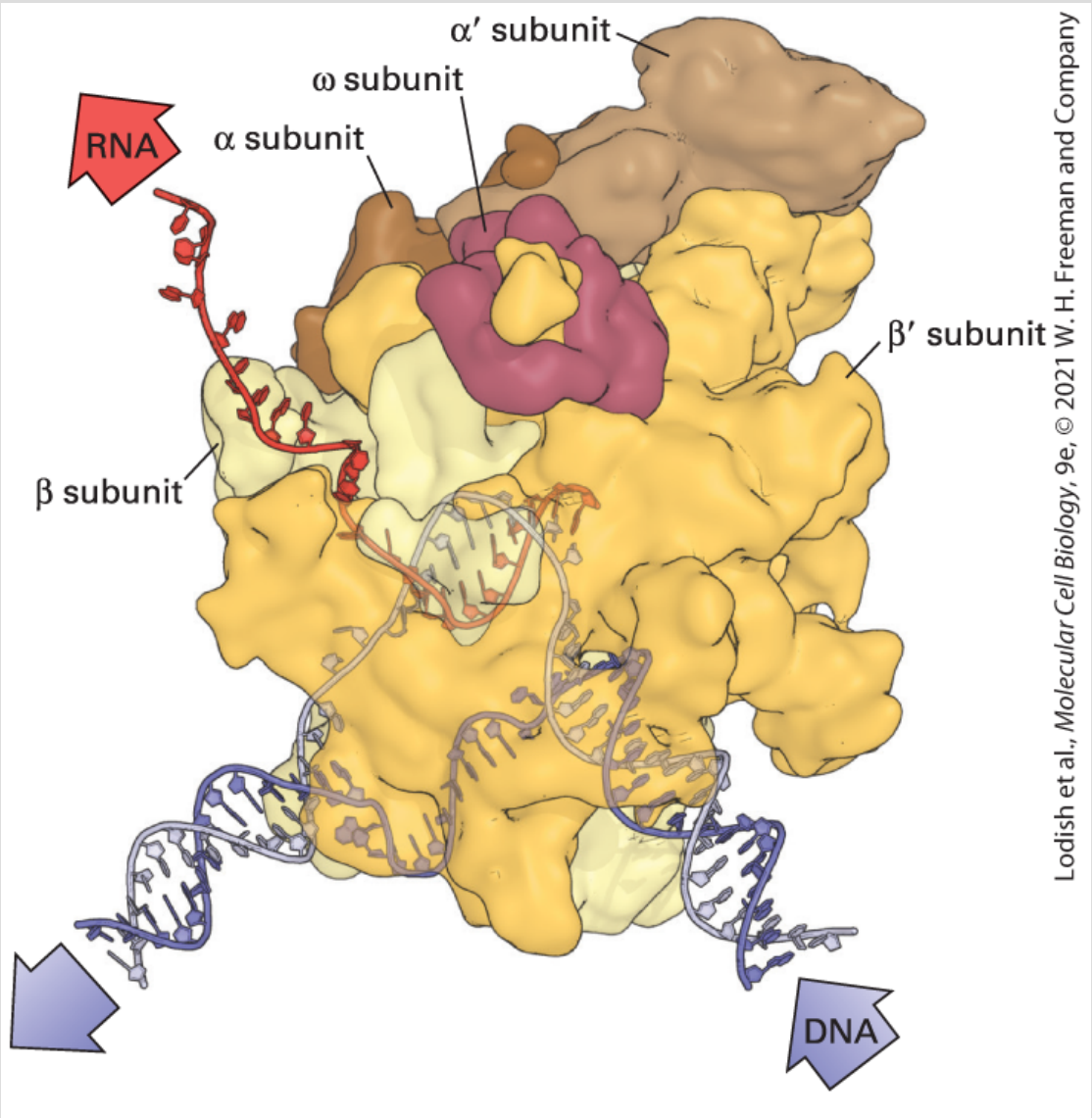


FIGURE 5-25 Bacterial RNA polymerase. This structure corresponds to the polymerase molecule in the elongation stage of transcription. In this diagram, transcription is proceeding in the rightward direction. Arrows indicate where downstream DNA enters the polymerase and upstream DNA exits at an angle from the downstream DNA. The template strand is light violet; the nontemplate strand, dark violet; the nascent RNA, red. The RNA polymerase β′ subunit is gold; the β subunit, light yellow; and the α subunits visible from this angle, brown. Nucleotides

complementary to the template DNA are added to the $3'$ end of the nascent RNA strand on the right side of the transcription bubble. The newly synthesized nascent RNA exits the polymerase at the upstream side through a channel formed by the β subunit. The ω subunit is also visible from this angle.

[Data courtesy of Seth Darst; see N. Korzheva et al., 2000, *Science* **289**:619–625, and N. Opalka et al., 2003, *Cell* **114**:335–345.]

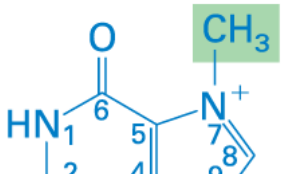
Eukaryotic Precursor mRNAs Are Processed to Form Functional mRNAs

Early research on the structure of eukaryotic genes involved studies of viruses that infect animals. When researchers analyzed the regions of a viral DNA molecule that encode viral mRNAs, they were surprised to observe that the sequence of a single viral mRNA was encoded in several regions of the viral DNA separated by DNA sequences that are not present in the mRNA. Later, the development of gene cloning and DNA sequencing (see Chapter 6) allowed researchers to compare the genomic DNA sequences of multicellular organisms with the sequences of their mRNAs. This research revealed that most cellular mRNAs are also encoded in several separate regions of genomic DNA, called *exons*, separated by sequences of DNA called *introns*. Further studies showed that a gene is first transcribed into a long primary transcript that includes both the exon sequences and the intron sequences that separate them. Subsequently, the introns are removed and the exons are spliced together (see Chapter 9). Although introns are common in

multicellular eukaryotes, they are extremely rare in bacteria and archaea and uncommon in many unicellular eukaryotes, such as baker's yeast.

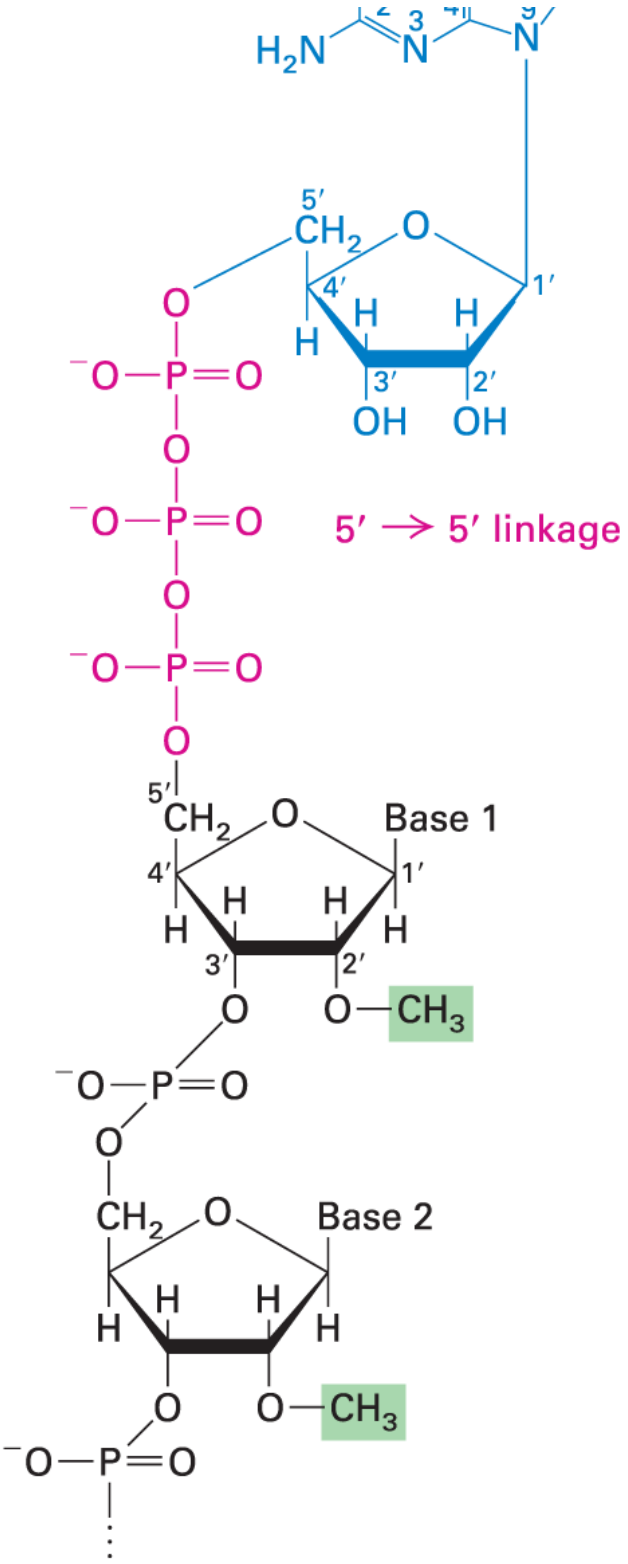
In eukaryotic cells, the site of RNA synthesis — the nucleus — is separated from the site of translation — the cytoplasm. In the nucleus, the primary transcripts of protein-coding genes — **precursor mRNAs (pre-mRNAs)** — must undergo several modifications, collectively termed *RNA processing*, to yield a functional mRNA (see Figure 5-1, step 2). This mRNA must then be exported to the cytoplasm before it can be translated into protein. Thus transcription and translation cannot occur concurrently in eukaryotic cells.

All eukaryotic pre-mRNAs are initially modified at the two ends, and these modifications are retained in mRNAs. As the 5′ end of a nascent RNA chain emerges from the surface of RNA polymerase, it is immediately acted on by several enzymes that together synthesize the *5′ cap*, a 7-methylguanylate that is connected to the terminal nucleotide of the RNA by an unusual 5′, 5′ triphosphate linkage (Figure 5-26). The cap protects an mRNA from enzymatic degradation and assists in its export to the cytoplasm, where it is bound by a protein factor required to begin translation.



7-Methylguanylate

n and Company



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freema

FIGURE 5-26 Structure of the 5′ methylated cap. The distinguishing chemical features of the 5′ methylated cap on eukaryotic mRNA are (1) the 5′ → 5′ linkage of

7-methylguanylate to the initial nucleotide of the mRNA molecule and (2) the methyl group on the 2′ hydroxyl of the ribose of the first nucleotide (base 1). Both of these features occur in all animal cells and in cells of higher plants; yeasts lack the methyl group on nucleotide 1. The ribose of the second nucleotide (base 2) is also methylated in vertebrates. See A. J. Shatkin, 1976, *Cell* **9**:645.

Processing at the 3′ end of a pre-mRNA involves cleavage by an endonuclease to yield a free 3′-hydroxyl group, to which a string of adenylic acid residues is added one at a time by an enzyme called *poly(A) polymerase.* The resulting *poly(A) tail* contains 100–250 bases, and is shorter in yeasts and invertebrates than in vertebrates. Poly(A) polymerase is part of a complex of proteins that can locate and cleave a transcript at a specific site and then add the correct number of A bases, in a process that does not require a template. As discussed further in Section 5.6 and in Chapter 10, the poly(A) tail has important functions both in translation of mRNA and in stabilizing pre-mRNAs in the nucleus and fully processed mRNAs in the nucleus and cytoplasm.

Another step in the processing of many different eukaryotic mRNA molecules is **RNA splicing**: the internal cleavage of a transcript to excise the introns and stitch together the coding exons. Figure 5-27 summarizes the basic steps in eukaryotic mRNA processing, using the β-globin gene as an example. We examine the cellular machinery for carrying out processing of mRNA, as well as tRNA and rRNA, in Chapter 9.

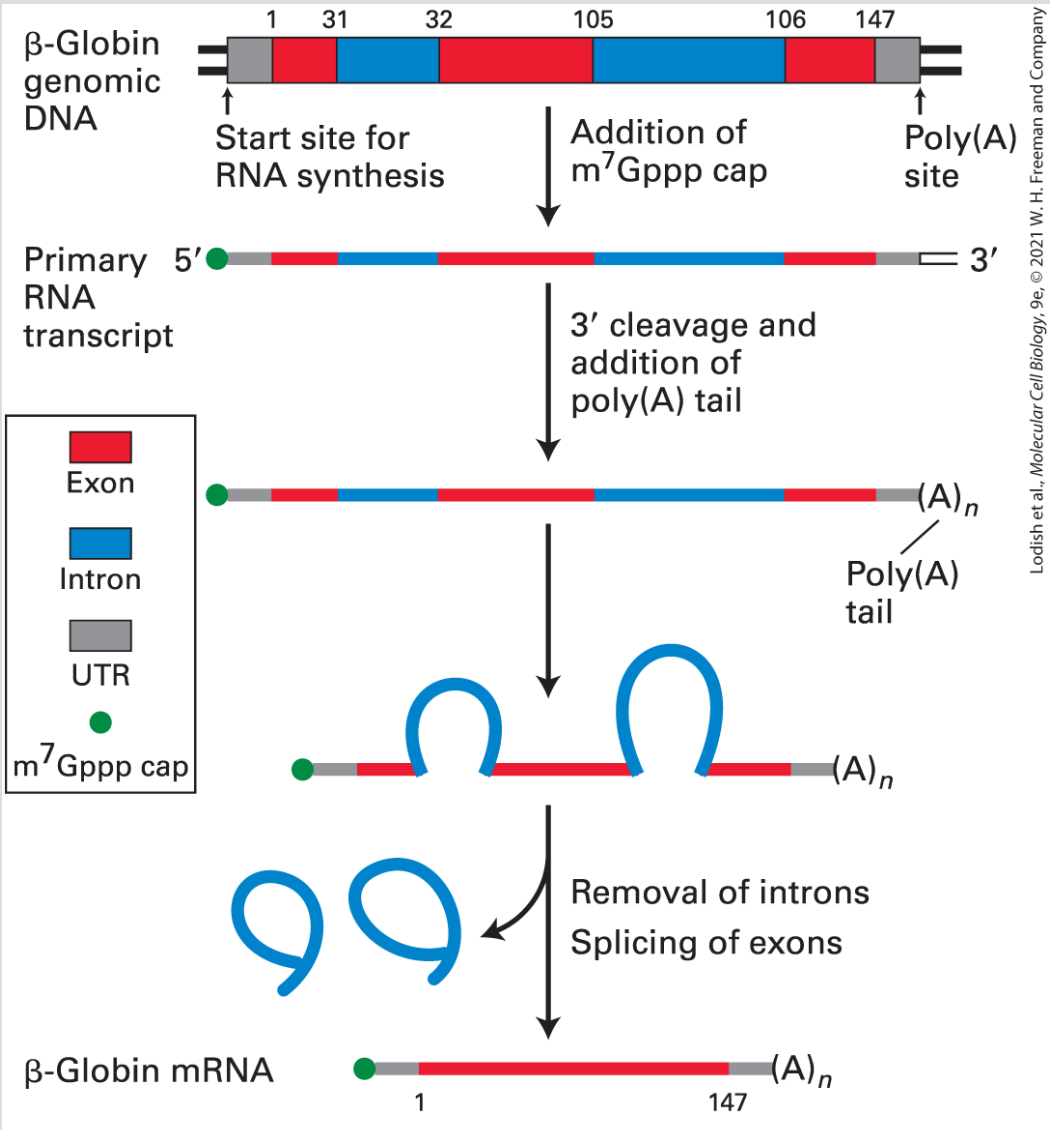


FIGURE 5-27 Overview of RNA processing. RNA processing produces functional mRNA in eukaryotes. The β-globin gene contains three protein-coding exons (constituting the coding region) and two intervening noncoding introns. The introns interrupt the protein-coding sequence between the codons for amino acids 31 and 32 and 105 and 106. Transcription of eukaryotic protein-coding genes starts before the sequence that encodes the first amino acid and extends beyond the sequence that encodes the last amino acid, resulting in noncoding regions at the ends of the primary transcript. These untranslated regions (UTRs) are retained during processing. The $5'$ cap $(\mathrm{m^7Gppp})$ is added during formation of the primary RNA transcript, which extends beyond the poly(A) site. After cleavage at the poly(A) site and addition of multiple A residues to the $3'$

end, splicing removes the introns and joins the exons. The small numbers refer to positions in the 147–amino acid sequence of β-globin.

The functional eukaryotic mRNAs produced by RNA processing retain noncoding regions, referred to as *untranslated regions (UTRs)*, at each end. In mammalian mRNAs, the 5′ UTR may be 100 or more nucleotides long, and the 3′ UTR may be several kilobases in length. Bacterial mRNAs also usually have 5′ and 3′ UTRs, but these regions are much shorter than those in eukaryotic mRNAs, generally containing fewer than 10 nucleotides. As discussed in Chapter 9, the 5′ UTR and 3′ UTR sequences participate in regulation of mRNA translation and stability, and 3′ UTRs also function in the localization of many mRNAs to specific regions of the cytoplasm.

Alternative RNA Splicing Increases the Number of Proteins That Can Be Expressed from a Single Eukaryotic Gene

The vast majority of genes in multicellular eukaryotes are made up of multiple exons separated by introns. As noted in Chapter 3, many proteins from higher eukaryotes have a multidomain tertiary structure (see Figure 3-10) and it is often the case that the boundaries between protein domains correspond to junctions

between exons of the underlying gene. This tendency for exons to correspond to discrete folded domains in the encoded protein suggests that many eukaryotic proteins have evolved by genetic recombination events that bring exons together in new combinations, a process known as exon shuffling.

We will see many examples throughout this book of mammalian proteins that contain multiple repeated copies of a protein domain with identical or nearly identical amino acid sequences. Evidently these sequences arose by recombination events that led to a string of multiple copies of the same exon separated by introns. The presence of multiple intron genes permits expression of multiple variations of the protein from a single gene by means of **alternative splicing**. In higher eukaryotes, alternative splicing is an important mechanism for production of different forms of a protein, called **isoforms**, by different types of cells.

Fibronectin, a multidomain protein found in mammals, provides a good example of alternative splicing (Figure 5-28). Fibronectin is a long, adhesive protein secreted into the extracellular space that can bind other proteins together. What and where it binds depends on which domains are spliced together. The fibronectin gene contains numerous exons, grouped into several regions corresponding to specific domains of the protein. Fibroblasts produce fibronectin mRNAs that contain exons EIIIA and EIIIB; these exons encode a protein domain that binds tightly to proteins in the fibroblast plasma membrane. Consequently, this fibronectin isoform adheres

fibroblasts to the extracellular matrix. Alternative splicing of the fibronectin primary transcript in hepatocytes, the major type of cell in the liver, yields mRNAs that lack the EIIIA and EIIIB exons. As a result, the fibronectin secreted by hepatocytes into the blood does not adhere tightly to fibroblasts or to most other cell types, which allows it to circulate. During formation of blood clots, however, other fibrin-binding domains of hepatocyte fibronectin bind to fibrin, one of the principal constituents of blood clots. Yet another domain of the bound fibronectin then interacts with integrins on the membranes of passing platelets, thereby expanding the clot by addition of platelets.

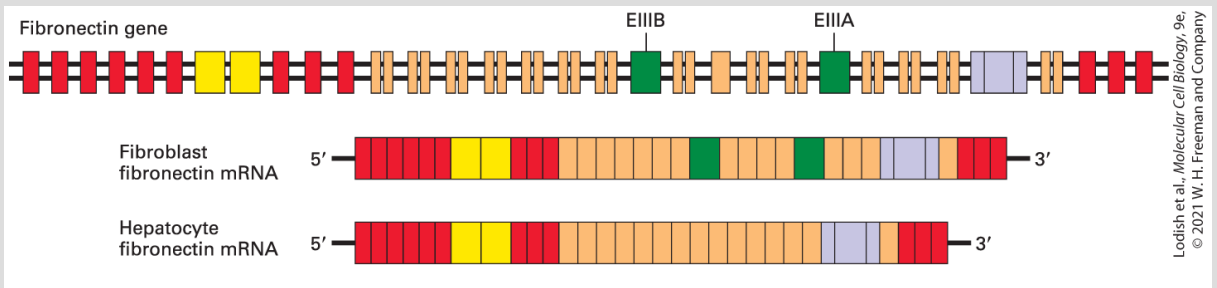


FIGURE 5-28 Alternative splicing. The ~75-kb fibronectin gene (*top*) contains multiple exons; splicing of the fibronectin transcript varies by cell type. The EIIIB and EIIIA exons (green) encode binding domains for specific proteins on the surface of fibroblasts. The fibronectin mRNA produced in fibroblasts includes the EIIIA and EIIIB exons, whereas these exons are spliced out of fibronectin mRNA in hepatocytes. In this diagram, introns (black lines in the top diagram of the fibronectin gene) are not drawn to scale; most of them are much longer than any of the exons.

More than 20 different isoforms of fibronectin have been identified, each encoded by a different, alternatively spliced mRNA composed of a unique combination of fibronectin gene exons. Sequencing of

large numbers of mRNAs isolated from various tissues and comparison of their sequences with genomic DNA has revealed that nearly 90 percent of all human genes are expressed as alternatively spliced mRNAs. Clearly alternative RNA splicing greatly expands the number of protein variants encoded by the genomes of higher, multicellular organisms.

KEY CONCEPTS OF SECTION 5.4

Transcription of Protein-Coding Genes and Formation of mRNA

- Transcription of DNA is carried out by RNA polymerase, which adds one ribonucleotide at a time to the $3'$ end of a growing RNA chain (see Figure 5-23). The sequence of the template DNA strand determines the order in which ribonucleotides are polymerized to form an RNA chain.
- During transcription initiation, RNA polymerase binds to a specific site in DNA (the promoter), locally melts the double-stranded DNA to reveal the unpaired template strand, and polymerizes the first two nucleotides complementary to the template strand. The melted region of 12–14 base pairs is known as the transcription bubble.
- During strand elongation, RNA polymerase moves down the DNA, melting the DNA ahead of the polymerase so that the template strand can enter the active site of the enzyme, and allowing the complementary strands of the region just transcribed to reanneal behind it. The transcription bubble moves with the polymerase as the enzyme adds ribonucleotides complementary to the template strand to the $3'$ end of the growing RNA chain.
- When RNA polymerase reaches a termination sequence in the DNA, the enzyme stops transcription, leading to release of the completed RNA and dissociation of the enzyme from the template DNA.
- In eukaryotic DNA, each protein-coding gene is transcribed from its own promoter. The initial primary transcript very often contains noncoding regions (introns) interspersed with coding regions (exons).
- Eukaryotic primary transcripts must undergo RNA processing to yield functional RNAs. During processing, the ends of nearly all primary transcripts from protein-coding

genes are modified by addition of a $5'$ cap and $3'$ poly(A) tail. Transcripts from genes containing introns undergo splicing, the removal of the introns and joining of the exons (see Figure 5-27).

- The individual domains of multidomain proteins found in higher eukaryotes are often encoded by individual exons or a small number of exons. Distinct isoforms of such proteins are often expressed in specific cell types as the result of alternative splicing of exons.

5.5 The Decoding of mRNA by tRNAs

As we saw in Chapter 3, the linear order of amino acids in each protein determines its three-dimensional structure and activity. Translation is the whole process by which the nucleotide sequence of an mRNA is used as a template to join the amino acids of a polypeptide chain in the correct order (see Figure 5-1, step **3**). In eukaryotic cells, protein synthesis occurs in the cytoplasm, where three types of RNA molecules come together to perform different but cooperative functions (Figure 5-29):

1. **Messenger RNA (mRNA)** carries the genetic information transcribed from DNA in a linear form. The mRNA is read in sets of three-nucleotide sequences, called **codons**, each of which specifies a particular amino acid.
2. **Transfer RNA (tRNA)** is the key to deciphering the codons in mRNA and serves as an adaptor to link three bases of RNA sequence to a particular amino acid. Each type of amino acid has its own subset of tRNAs, which are covalently bound to that amino acid and carry it to the growing end of a polypeptide chain when the next codon in the mRNA calls for it. The correct tRNA with its attached amino acid is selected at each step because each specific tRNA molecule contains a three-nucleotide sequence, an **anticodon**, that can base-pair with its complementary codon in the mRNA.

3. **Ribosomal RNA (rRNA)** associates with a set of proteins to form ribosomes. These complex structures form the scaffold for stepwise alignment of the codons in mRNA with the anticodons in tRNA. The ribosome also catalyzes the sequential formation of peptide bonds between amino acids carried by tRNAs for the assembly of a polypeptide chain. Ribosomes are composed of a large and a small subunit, each of which contains its own rRNA molecule or molecules.

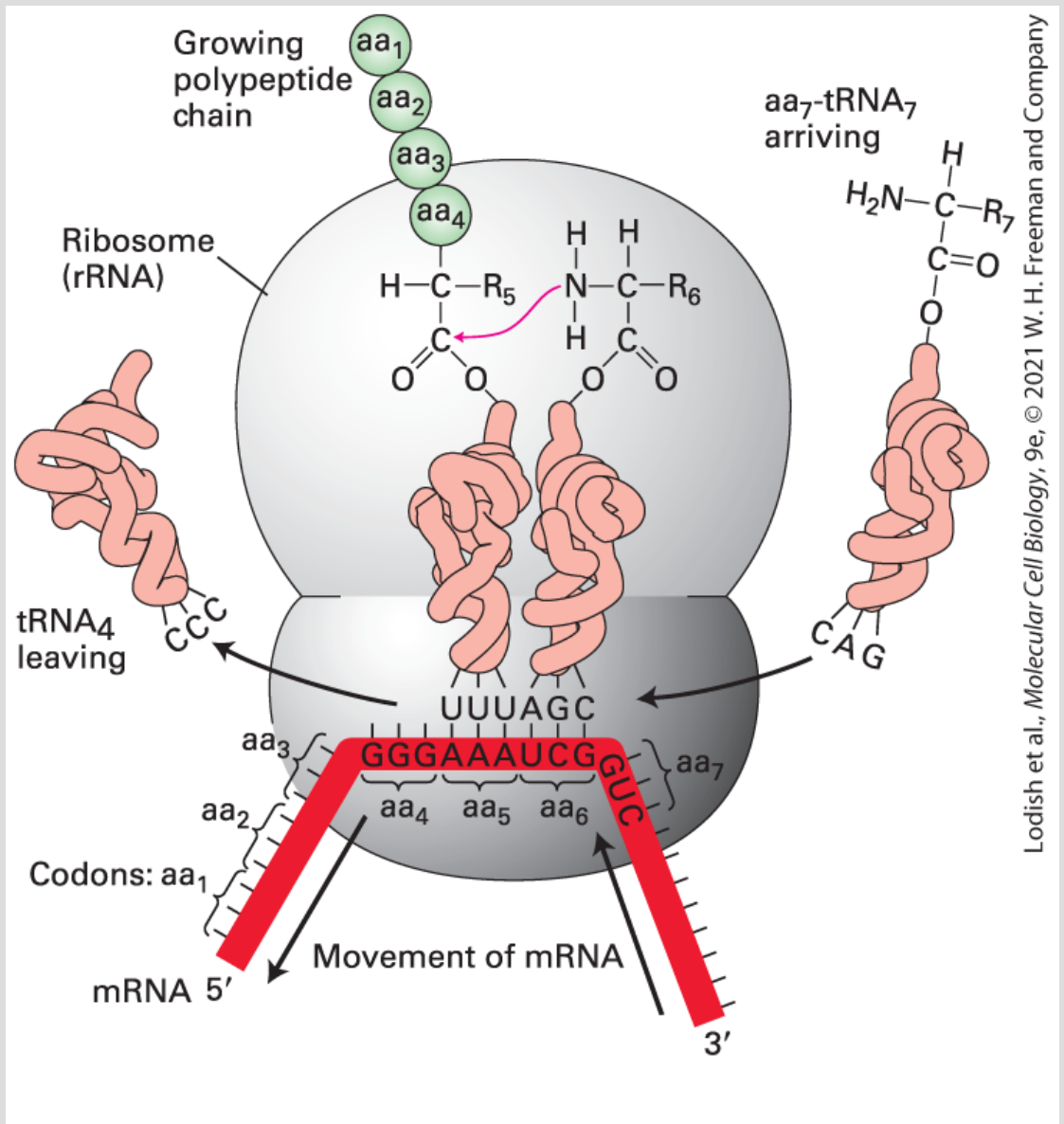


FIGURE 5-29 Three roles of RNA in protein synthesis. Messenger RNA (mRNA) is translated into protein by the joint action of transfer RNA (tRNA) and the ribosome, which is composed of numerous proteins and three (bacterial) or four (eukaryotic) ribosomal RNA (rRNA) molecules (not shown). Note the base pairing between tRNA anticodons and complementary codons in the mRNA. Formation of a peptide bond between the amino-group N on the incoming aa-tRNA and the carboxy-terminal C on the growing protein chain (green) is catalyzed by one of the rRNAs.
$\mathrm{aa} = \mathrm{amino\ acid}; \mathrm{R} = \mathrm{side\ group}$. Note that these are simplified diagrams of tRNAs and the ribosomal subunits. Their actual structures are shown in Figure 5-32b and Figure 5-35.

These three types of RNA participate in the synthesis of proteins in all organisms. In this section, we focus on the decoding of mRNA by tRNAs and how the structure of each of these RNAs relates to its specific task. How they work together with ribosomes and protein factors to synthesize proteins is detailed in the following section. Because translation is essential for protein synthesis, the two processes are commonly referred to interchangeably. However, the polypeptide chains resulting from translation must undergo additional steps after translation, including folding and often other changes (e.g., chemical modifications, association with other chains) that are required for the production of mature, functional proteins (see Chapter 3).

Messenger RNA Carries Information from DNA in a Three-Letter Genetic

Code

As noted previously, the genetic code used by cells is a *triplet* code, in which every three-nucleotide sequence, or codon, is read from a specified starting point in the mRNA. Of the 64 possible codons in the genetic code (one of four nucleotides at each of the three positions of a codon yields $4 \times 4 \times 4 = 64$ possible codons), 61 specify individual amino acids, and 3 are stop codons. Table 5-1 shows that most amino acids are encoded by more than one codon. Only two — methionine and tryptophan — have a single codon; at the other extreme, leucine, serine, and arginine are each specified by six different codons. The different codons that code for a given single amino acid are said to be *synonymous*. The code itself is termed *degenerate*, meaning that a particular amino acid can be specified by several codons.

TABLE 5-1 • The Genetic Code (Codons to Amino Acids)*

First position (5′ end)	Second position: U	C	A	G	Third position (3′ end)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu (Met)*	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met (Start)	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val (Met)*	Ala	Glu	Gly	G

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

*AUG is the most common initiation codon; GUG usually codes for valine; and CUG for leucine, but rarely these codons can also code for methionine to initiate a protein chain.

Synthesis of all polypeptide chains in prokaryotic and eukaryotic cells begins with the amino acid methionine. In bacteria, a specialized form of methionine with a formyl group linked to its amino group is used. In most mRNAs, the *start (initiation) codon* specifying this amino-terminal methionine is AUG. In a few bacterial mRNAs, GUG is used as the initiation codon, and CUG is

occasionally used as an initiation codon for methionine in eukaryotes. The three codons UAA, UGA, and UAG do not specify amino acids, but rather constitute *stop (termination) codons* that mark the carboxyl terminus of polypeptide chains in almost all cells. The sequence of codons that runs from a specific start codon to a stop codon is called a **reading frame**. This precise linear array of ribonucleotides in groups of three in mRNA specifies the precise linear sequence of amino acids in a polypeptide chain and also signals where synthesis of the chain starts and stops.

The great utility of the genetic code table is that it allows the amino acid sequence of the encoded protein to be deduced from the sequence of an mRNA. Since the genetic code is read as a continuous sequence of triplet codons without divisions between them, a particular mRNA theoretically could be translated in three different reading frames (Figure 5-30). Once the sequence of an mRNA is known (usually determined from the DNA sequence of the transcribed gene), it is necessary to determine which of the three possible reading frames is translated by the ribosome into protein sequence. In principle, this could be done by identifying the AUG codon that is the start of translation. In practice, however, finding the correct initiation codon is error prone and the most often used method is to search for a long **open reading frame**. The idea behind this method is that in sequences that do not code for proteins, including the two incorrect reading frames of an mRNA, stop codons should appear randomly as 3 out of 64 codons, and thus

noncoding reading frames should only have relatively short stretches (on average of about 20 amino acid codons) between stop codons. Thus a given mRNA will have two reading frames that are broken up by multiple stop codons and usually only have one long open reading frame not interrupted by stop codons representing the sequence of the encoded protein.

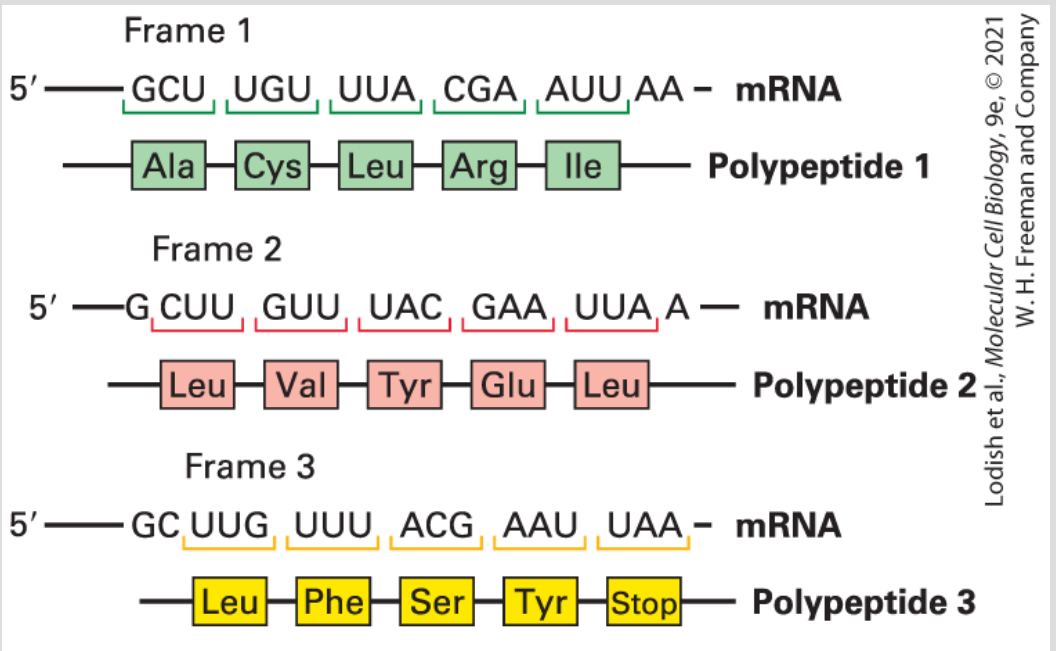


FIGURE 5-30 Three possible reading frames in an mRNA sequence. If translation of the mRNA sequence shown begins at three different upstream start sites (not shown), then three overlapping reading frames are possible. In this example, the codons are shifted one base to the right in the middle frame and two bases to the right in the third frame, which ends in a stop codon. As a result, the same mRNA nucleotide sequence can specify different amino acids. The correct reading frame for a given mRNA sequence can usually be readily identified as the longest open reading frame that is not interrupted by stop codons.

A related use of the genetic code is to deduce the effect that a gene mutation will have on the encoded protein. To see the range of possibilities, imagine the effect of mutations in a serine codon UCG that occurs in the middle of a protein-coding sequence. A mutation changing G $\rightarrow$ A would give a UCA codon, which also encodes serine and is thus a *synonymous mutation*, and would be predicted to have no effect on the function of the encoded protein. By contrast, a mutation changing C $\rightarrow$ U would give a UUG codon, which codes for leucine: this is a *missense mutation* that may have an effect on the function of the protein depending on the importance of serine at that position for protein function. On the other hand, a mutation changing C $\rightarrow$ A would give UAG, which is a stop codon and thus would cause a *nonsense mutation* that would terminate the protein sequence early and would likely cause a complete loss of function. Finally, an addition or subtraction of one or two bases will shift the reading frame. Known as a *frameshift mutation*, such a change in sequence would lead to a scrambled amino sequence with an early termination of the protein. Unless the frameshift mutation occurred near the end of the coding sequence, the likely consequence would be a complete loss of function.

The genetic code shown in Table 5-1 is known as the *universal code* since meaning of each codon is the same in most known organisms — strong evidence that life on Earth evolved only once. The 20 amino acids are not distributed among the 61 sense codons in haphazard fashion, and the underlying structure of the universal

code provides clues to how an optimal code evolved. The most prominent pattern to emerge when the code is considered as a whole is that changing one pyrimidine for another (C $\rightarrow$ U or U $\rightarrow$ C) or changing one purine for another (G $\rightarrow$ A or A $\rightarrow$ G) in either the first or the third position typically either leads to a synonymous codon or changes to a codon for an amino acid with similar chemical properties. These types of sequence changes, as we saw in Section 5.3, are the most frequent kinds of mistakes to be made during DNA replication. Evidently the code evolved to minimize the deleterious effects on the amino acid sequence of proteins of the most common types of mutations in DNA.

The Folded Structure of tRNA Promotes Its Decoding Functions

The two information-carrying polymers, DNA and mRNA, perform their functions with the flexibility to accommodate a vast number of different sequences. On the other hand, tRNAs have specific fixed sequences that allow the tRNA molecules to fold into a three-dimensional structure that is crucial for their function. To participate in protein synthesis, a tRNA molecule must become chemically linked to a particular amino acid via a high-energy bond, forming an **aminoacyl-tRNA** (Figure 5-31). Enzymes called *aminoacyl-tRNA synthetases* are responsible for adding the correct amino acid to a given tRNA and, as will be described, they do this

through the recognition of particular features of the tRNA structure. During the process of translation, an aminoacyl-tRNA interacts with the ribosome and other protein factors so that the anticodon sequence is properly base-paired with a codon in mRNA. All of these interactions also depend on the folded structure of the tRNA (see Figure 5-29).

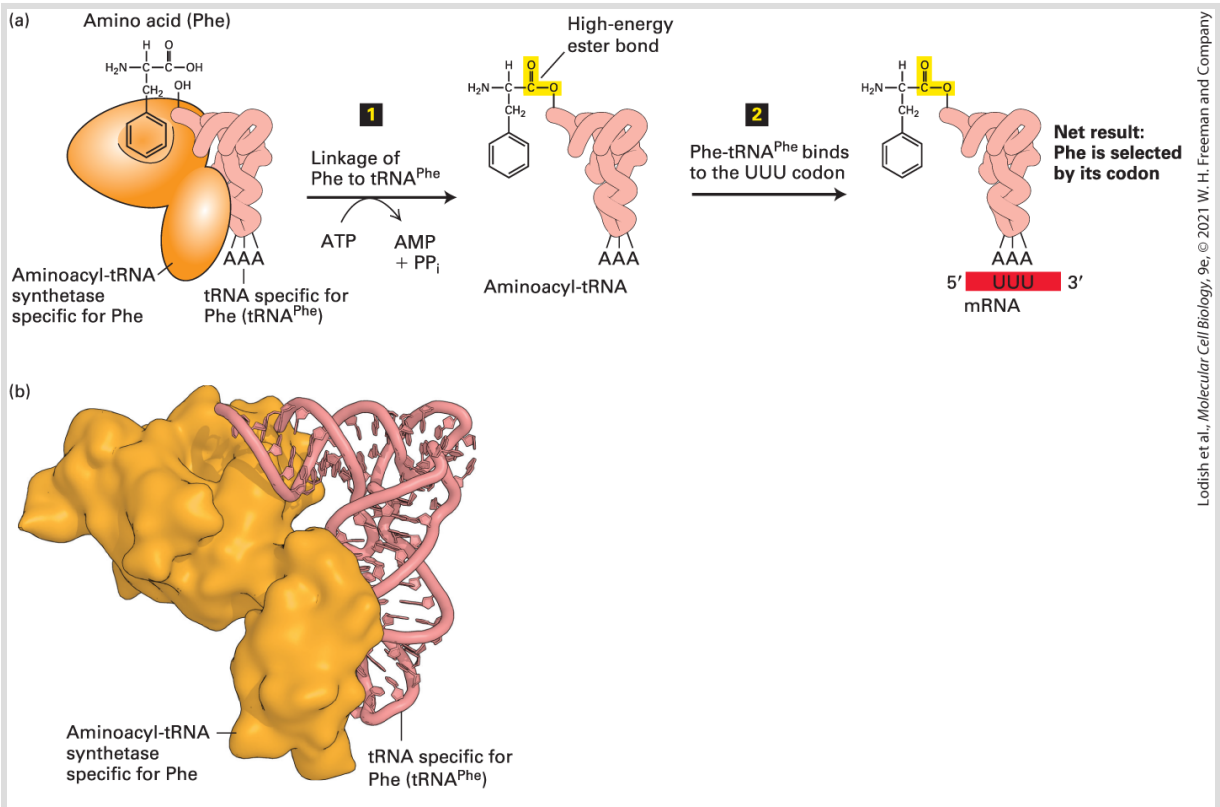


FIGURE 5-31 Translating nucleic acid sequence into amino acid sequence. (a) The process for translating nucleic acid sequences in mRNA into amino acid sequences in proteins involves two steps. Step **1**: An aminoacyl-tRNA synthetase first couples a specific amino acid, via a high-energy ester bond (yellow), to either the $2'$ or $3'$ hydroxyl of the terminal adenosine in the corresponding tRNA. Step **2**: A three-base sequence in the tRNA (the anticodon) then base-pairs with a codon in the mRNA specifying the attached amino acid. If an error occurs in either step, the wrong amino acid may be incorporated into a polypeptide chain. $\mathrm{Phe} = \mathrm{phenylalanine}$. (Note that this is a simplified diagram of $\mathrm{tRNA}^{\mathrm{Phe}}$; its actual

structure is shown in Figure 5-32b.) (b) Molecular model of the human mitochondrial aminoacyl-tRNA synthetase for Phe in complex with $tRNA^{Phe}$.

[Data from L. Klipcan et al., 2012, *J. Mol. Biol*. **415**:527, PDB ID 3tup.]

Some 30–40 different tRNAs have been identified in bacterial cells and as many as 50–100 in animal and plant cells. Thus the number of tRNAs in most cells is more than the number of amino acids used in protein synthesis (20), but less than the number of amino acid codons in the genetic code (61). Consequently, many amino acids have more than one tRNA to which they can attach (explaining how there can be more tRNAs than amino acids); in addition, many tRNAs can pair with more than one codon (explaining how there can be more codons than tRNAs).

tRNA molecules, which are 70–80 nucleotides long, are initially formed by transcription from a tRNA gene template. The maturation of the primary transcript to a functional tRNA molecule with precise three-dimensional structure takes place in three stages. The first stage is the chemical modification of certain bases, creating nonstandard nucleotides, such as inosine, dihydrouridine, and pseudouridine. As we will see shortly, some of these modified bases are known to play an important role in protein synthesis. The second stage involves complementary base pairing between antiparallel arrangements of segments of the same tRNA molecule to produce a form that resembles a cloverleaf when drawn in two dimensions (Figure 5-32a). The four short, double-helical segments

that form are stabilized mostly by Watson-Crick base pairing, although other nonstandard base pairs, such as G·U, are accommodated by the helical structures that are much less uniform than that of DNA. Three of the four helical stems have loops containing seven or eight bases at their ends, while the remaining, unlooped stem contains the free 3′ and 5′ ends of the chain. The three nucleotides composing the anticodon are located at the center of the middle loop, in an accessible position that facilitates codon-anticodon base pairing. In all tRNAs, the 3′ end of the unlooped *acceptor stem*, to which a specific amino acid is attached, has the sequence CCA, which in most cases is added after synthesis and processing of the tRNA are complete. Finally, the first and third loops join together by base pairing to form a compact, L-shaped structure when viewed in three dimensions. The anticodon loop and acceptor stem form the ends of the folded tRNA molecule (Figure 5-32b).

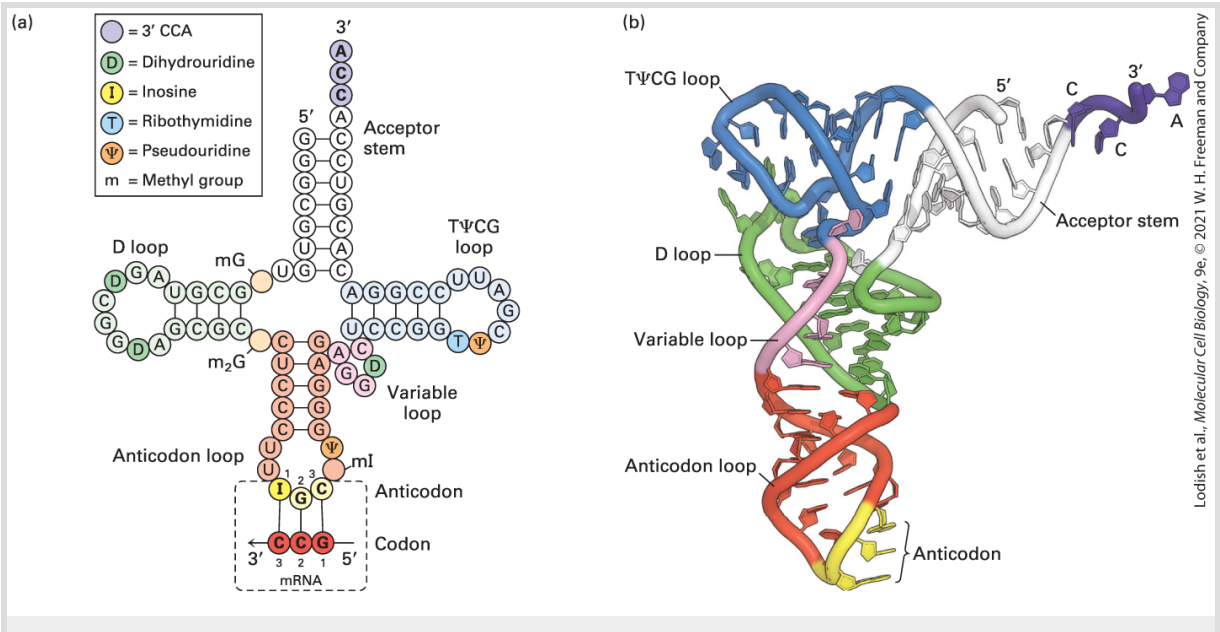


FIGURE 5-32 Structure of tRNAs. (a) Although the exact nucleotide sequence varies among tRNAs, they all fold into four base-paired stems and three loops. The CCA sequence at the $3'$ end is also found in all tRNAs. Attachment of an amino acid to the $3'$ A yields an aminoacyl-tRNA. Some of the A, C, G, and U residues are modified post-transcriptionally in most tRNAs (see key). Dihydrouridine (D) is nearly always present in the D loop; likewise, ribothymidine (T) and pseudouridine (ψ) are almost always present in the TψCG loop. Yeast alanine tRNA, represented here, also contains other modified bases. The triplet at the tip of the anticodon loop base-pairs with the corresponding codon in mRNA. See R. W. Holly et al., 1965, *Science* **147**:1462. (b) Three-dimensional model of the generalized backbone of all tRNAs. Note the L shape of the molecule.

[Part (b) Data from J. G. Arnez and D. Moras, 1997, *Trends Biochem. Sci.* **22**:211, PDB ID 1vtq.]

Nonstandard Base Pairing Often Occurs Between Codons and Anticodons

If perfect Watson-Crick base pairing between codons and anticodons were required, cells would have to contain at least 61 different types of tRNAs, one for each codon that specifies an amino acid. As noted previously, however, many cells contain fewer than 61 tRNAs. The explanation for the smaller number lies in the capability of a single tRNA anticodon to recognize more than one, but not necessarily every, codon corresponding to a given amino acid. This broader recognition can occur because of nonstandard pairing between bases in the so-called *wobble* position: that is, the third ($3'$) base in an mRNA codon and the corresponding first ($5'$) base in its tRNA anticodon.

The first and second bases of a codon almost always form standard Watson-Crick base pairs with the third and second bases, respectively, of the corresponding anticodon, but four nonstandard interactions can occur between bases in the wobble position. Particularly important is the G·U base pair, which fits into the short, 3-bp RNA-RNA double-stranded region formed between the codon and the anticodon almost as well as the standard G·C pair. Thus a tRNA anticodon with G in the first (wobble) position can base-pair with the two corresponding codons that have either pyrimidine (C or U) in the third position (see Figure 5-26). For example, the phenylalanine codons UUU and UUC ($5' \rightarrow 3'$) are both recognized by the tRNA that has GAA ($5' \rightarrow 3'$) as its anticodon. In fact, any two codons of the type NNPyr ($\mathrm{N} = \text{any base}; \mathrm{Pyr} = \text{pyrimidine}$) encode a single amino acid and are decoded by a single tRNA with G in the first (wobble) position of the anticodon.

Although adenine is rarely found in the anticodon wobble position, many tRNAs in plants and animals contain inosine (I), a deaminated product of adenine, at this position. Inosine can form nonstandard base pairs with A, C, and U. A tRNA with inosine in the wobble position thus can recognize the corresponding mRNA codons with A, C, or U in the third (wobble) position (see Figure 5-33). For this reason, inosine-containing tRNAs are heavily employed in translation of the synonymous codons that specify a single amino acid. For example, four of the six codons for leucine (CUA, CUC, CUU, and UUA) are all recognized by the same tRNA with the

anticodon 3′-GAI-5′; the inosine in the wobble position forms nonstandard base pairs with the third base in each of these four codons.

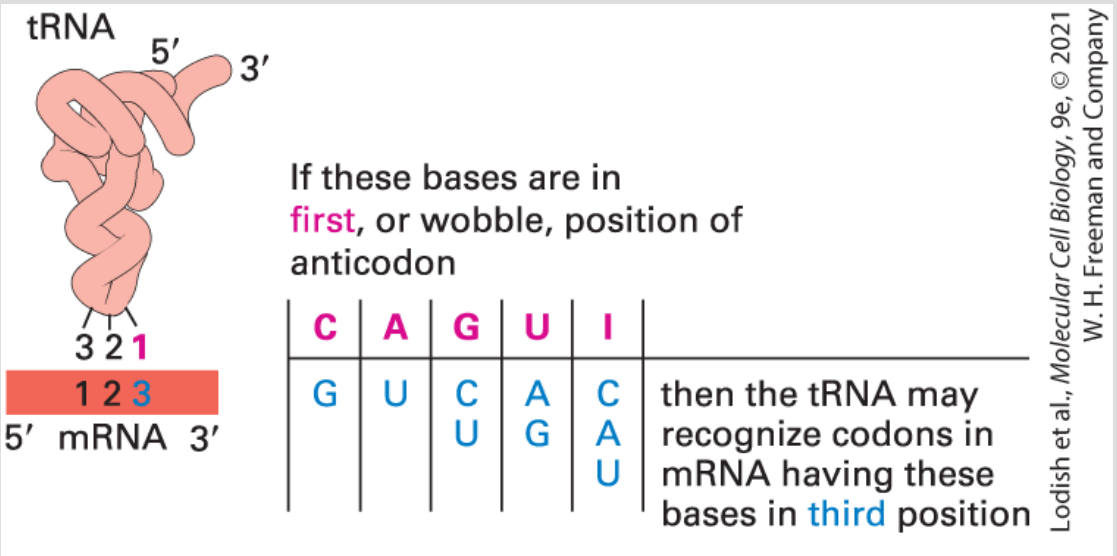


FIGURE 5-33 Nonstandard base pairing at the wobble position. The base in the third (or wobble) position of an mRNA codon often forms a nonstandard base pair with the base in the first (or wobble) position of a tRNA anticodon. Wobble pairing allows a tRNA to recognize more than one mRNA codon, although each of those tRNAs will bear the same amino acid. Note that a tRNA with I (inosine) in the wobble position can read (become paired with) three different codons, and a tRNA with G or U in the wobble position can read two codons. Although A is theoretically possible in the wobble position of the anticodon, it is almost never found in nature. (Note that this is a simplified diagram of a tRNA. The actual structure of a tRNA is shown in Figure 5-32b.)

Amino Acids Are Linked to Their Cognate tRNAs with Great Accuracy

Recognition of the codon or codons specifying a given amino acid by a particular tRNA is actually the second step in decoding the genetic message. The first step, attachment of the appropriate amino acid to a tRNA, is catalyzed by a specific aminoacyl-tRNA synthetase. Each of the 20 different synthetases recognizes *one* amino acid and *all* its compatible, or *cognate*, tRNAs. These coupling enzymes link an amino acid to the free 2′ or 3′ hydroxyl of the adenosine at the 3′ terminus of the tRNA molecule by an ATP-requiring reaction. In this reaction, the amino acid is linked to the tRNA by a high-energy bond and is thus said to be *activated.* The energy of this bond subsequently drives the formation of the peptide bonds linking adjacent amino acids in a growing polypeptide chain. The equilibrium of the aminoacylation reaction is driven further toward activation of the amino acid by hydrolysis of the high-energy phosphoanhydride bond in the released pyrophosphate (see Figure 5-31a).

Each aminoacyl-tRNA synthetase must specifically recognize two substrates; the amino acid and the corresponding set of cognate tRNAs. Any errors made in tRNA charging could lead to a deleterious error in translation and the aminoacyl-tRNA synthetases have evolved to have high fidelity, mischarging a tRNA only about one in 10^4 times. As one would expect, most aminoacyl-tRNA synthetases recognize their cognate tRNAs by interacting with the anticodon loop. However, for amino acids that have multiple codons, other features common to all the cognate tRNAs contribute to specific

recognition. In an extreme case, in *E. coli*, the four tRNAs for alanine are not recognized by their anticodon loop at all. Instead, all four tRNAs contain a unique G·U base pair in the acceptor stem, which is used for specific recognition by the corresponding aminoacyl tRNA synthetase. In other instances, specific bases in incorrect tRNAs that are structurally similar to a cognate tRNA will inhibit charging of the incorrect tRNA. Thus recognition of the correct tRNA depends on both positive interactions and the absence of negative interactions.

A second source of errors in charging of tRNAs by aminoacyl-tRNA synthetases is selection of the wrong amino acid. In cases in which mischarging by a structurally very similar amino acid is possible, the aminoacyl-tRNA synthetase has evolved a second active site with a *proofreading* activity. For example, a single active site in the aminoacyl-tRNA synthetase for isoleucine would not be able to effectively exclude valine, which is similar to isoleucine but lacks one methyl group. For this enzyme, there is a second proofreading site that can remove an incorrect amino acid that has been added to $tRNA^{Ile}$. The hydrophobic pocket for this site is too small to admit isoleucine, but it can admit valine that has been erroneously linked to $tRNA^{Ile}$ and will remove it. This mechanism of proofreading by aminoacyl-tRNA synthetases is similar in outline to proofreading of incorrect nucleotide incorporation by the $3' \rightarrow 5'$ exonuclease activity of DNA polymerase.

KEY CONCEPTS OF SECTION 5.5

The Decoding of mRNA by tRNAs

- Genetic information from DNA is carried by mRNA in the form of a non-overlapping, degenerate triplet code.
- Each amino acid is encoded by one or more three-nucleotide sequences (codons) in mRNA. Each codon specifies one amino acid, but most amino acids are encoded by multiple codons (see Table 5-1).
- The AUG codon for methionine is the most common start codon, specifying the amino acid at the amino-terminus of a protein chain. Three codons (UAA, UAG, UGA) function as stop codons and specify no amino acids.
- There are three possible reading frames for the translation of a given mRNA sequence. The protein coding sequence can be identified as a long open reading frame that is not interrupted by stop codons.
- The genetic code can be used to determine the linear sequence of amino acids in a protein sequence and to deduce the effect of gene mutations on the encoded protein.
- All tRNAs have a similar three-dimensional structure that includes an acceptor stem for attachment of a specific amino acid and a stem-loop with a three-base anticodon sequence at its end (see Figure 5-32). The anticodon can base-pair with its corresponding codon in mRNA.
- Because of nonstandard interactions, a tRNA may base-pair with more than one mRNA codon; conversely, a particular codon may base-pair with multiple tRNAs. In each case, however, only the proper amino acid is inserted into a growing polypeptide chain.
- Each of the 20 aminoacyl-tRNA synthetases recognizes a single amino acid and covalently links it to a cognate tRNA, forming an aminoacyl-tRNA (see Figure 5-31). This reaction activates the amino acid so that it can participate in peptide bond formation.

5.6 Stepwise Synthesis of Proteins on Ribosomes

The previous sections have introduced two of the major participants in protein synthesis: mRNAs and aminoacyl-tRNAs. Here we first describe the third key player in protein synthesis — the rRNA-containing ribosome — before taking a detailed look at how all three components are brought together to carry out the biochemical events leading to the formation of polypeptide chains by ribosomes. The complex process of translation can be divided into three stages — initiation, elongation, and termination — which we consider in order. We focus our description on translation in eukaryotic cells, but the mechanism of translation is fundamentally the same in all cells.

Ribosomes Are Protein-Synthesizing Machines

In simple terms, the ribosome serves both as a scaffold to facilitate the binding aminoacyl-tRNAs to sequential codons in the mRNA and as an enzyme to catalyze the peptidyl transfer reaction for the stepwise formation of peptide bonds in the growing polypeptide chain. The ribosome is responsible for all protein synthesis and is the most abundant RNA-protein complex in the cell. Elongation of polypeptides occurs at a rate of 3 to 5 amino acids added per second,

so that although small proteins may be made in a minute or less, the formation of very large proteins such as titin, a muscle protein of about 30,000 amino acid residues, takes 2–3 hours. The ribosome must therefore be precise, and it must be persistent.

With the aid of the electron microscope, ribosomes were first discovered as small, discrete, RNA-rich particles found to be most abundant in those cells most active in protein synthesis. However, their direct role in protein synthesis was not recognized until reasonably pure ribosome preparations were obtained. In vitro radiolabeling experiments with such preparations showed that radioactive amino acids were first incorporated into growing polypeptide chains associated with ribosomes before appearing in finished proteins.

Although there are differences between the ribosomes of bacteria, archaea, and eukaryotes, the great structural and functional similarities between ribosomes from all species reflect the common evolutionary origin of the most basic constituents of living cells. A ribosome is composed of three (in bacteria and archaea) or four (in eukaryotes) distinct rRNA molecules and as many as 80 proteins, organized into a large subunit and a small subunit (Figure 5-34 and Table 5-2). The ribosomal subunits and the rRNA molecules are commonly designated in Svedberg units (S), a measure of the sedimentation rate of macromolecules centrifuged under standard conditions — essentially, a logarithmic measure of size. The small ribosomal subunit contains a single rRNA molecule, referred to as

small rRNA. The large subunit contains a molecule of *large rRNA* and one molecule of 5S rRNA, plus an additional molecule of 5.8S rRNA in eukaryotes. The lengths of the rRNA molecules, the numbers of proteins in each subunit, and consequently, the sizes of the subunits differ between bacterial and eukaryotic cells (see Table 5-2). The assembled ribosome is 70S in bacteria and 80S in vertebrates.

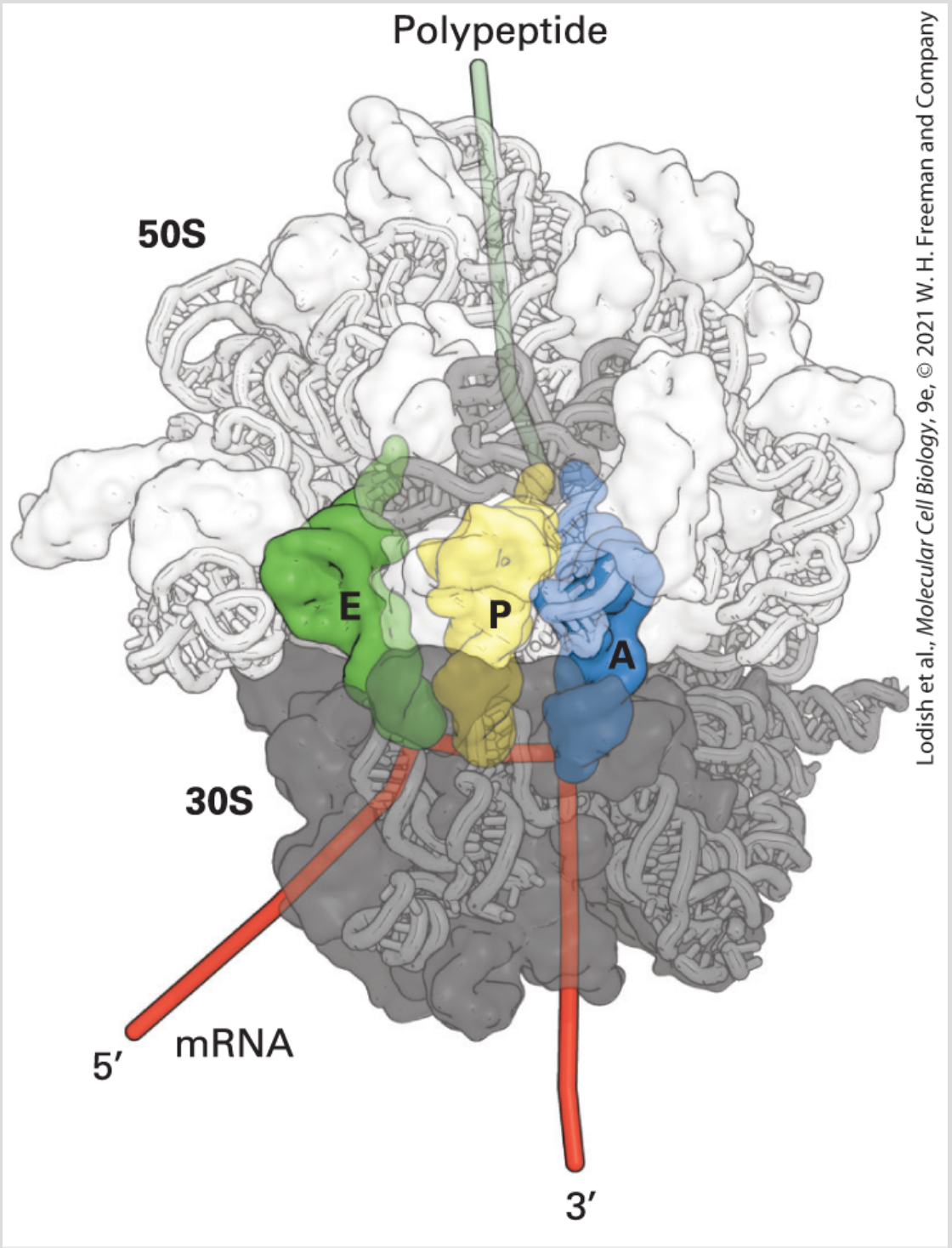


FIGURE 5-34 Structure of the bacterial ribosome. Model of the *Thermus thermophilus* ribosome viewed along the interface between the large (50S) and small (30S) subunits. The 16S rRNA and proteins in the small subunit are dark gray. RNA is depicted as a tube model and protein surfaces are shown. The 23S rRNA and proteins in the large subunit are light gray, and the 5S rRNA is an intermediate shade

of gray. The surface of the ribosome is made partially transparent to display the positions of bound tRNAs. Note that the ribosomal proteins are located primarily on the surface of the ribosome.

[Data from A. Korostelev et al., 2006, *Cell* **126**:1065–1077, PDB ID 4v4i.]

TABLE 5-2 • Ribosome Components

	Common core	*E. coli*	*S. cerevisiae*	Human
Ribosome	2.0 MDa	2.3 MDa	3.3 MDa	4.3 MDa
	34 proteins	54 proteins	79 proteins	80 proteins
	3 rRNAs	3 rRNAs	4 rRNAs	4 rRNAs
Large subunit		50S	60S	60S
	19 proteins	33 proteins	46 proteins	47 proteins
	23S rRNA: 2843 bases	23S rRNA: 2904 bases	25S rRNA: 3396 bases*	28S rRNA: 5034 bases*
			5.8S rRNA: 158 bases*	5.8S rRNA: 156 bases*
	5s rRNA: 121 bases	5S rRNA: 121 bases	5S rRNA: 121 bases	5S rRNA: 121 bases
Small subunit		30S	40S	40S
	15 proteins	21 proteins	33 proteins	33 proteins
	16S rRNA: 1458 bases	16S rRNA: 1542 bases	18S rRNA: 1800 bases	18S rRNA: 1870 bases

*5.8S rRNA in eukaryotes is base-paired to 25S or 28S rRNA.
Data from G. Yusupov and M. Yusupov, *Ann. Rev. Biochem.*, 2014, **83**:467.

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

The sequences of the small and large rRNAs from thousands of organisms are now known. Although the primary nucleotide sequences of these rRNAs vary considerably, the same parts of each type of rRNA contain complementary sequences that can form internally base-paired segments, indicating that they all share the same core structure. Since every organism has ribosomes, differences between their rRNA sequences have proven to be the most reliable means to determine the evolutionary distances between organisms. Even today, when whole genome sequences are

often available, phylogenetic comparisons usually begin with the sequence of the respective rRNA genes.

The three-dimensional structures of bacterial and yeast ribosomes (see Figure 5-35) and of the large subunit of an archaeal ribosome have been determined by x-ray crystallography. The structures of human (see Figure 5-35d) and plant ribosomes have also been determined by cryoelectron microscopy. The structure of the rRNAs in the common core, where mRNAs and tRNAs are bound and where peptide bond formation is catalyzed, is similar in all three domains of life. However, archaeal rRNAs and proteins are more similar to those of eukaryotic ribosomes than to those of bacterial ribosomes, reflecting their later divergence from a common ancestor (see Figure 1-1). For the most part, the multiple ribosomal proteins are much smaller than the rRNAs and associate with the surface of the ribosomes. Although the number of protein molecules in ribosomes greatly exceeds the number of RNA molecules, RNA constitutes about 60 percent of the mass of a bacterial ribosome and about 50 percent of the mass of a human ribosome. Eukaryotic ribosomes are generally similar to bacterial ribosomes but are larger because of eukaryote-specific insertions of RNA segments into regions of the common core rRNAs as well as the presence of a larger number of proteins (see Figure 5-35 and Table 5-2). Basic aspects of protein synthesis are thought to be similar among all three domains of life, although initiation of translation in eukaryotes, discussed later, is more complex and subject to additional mechanisms of regulation.

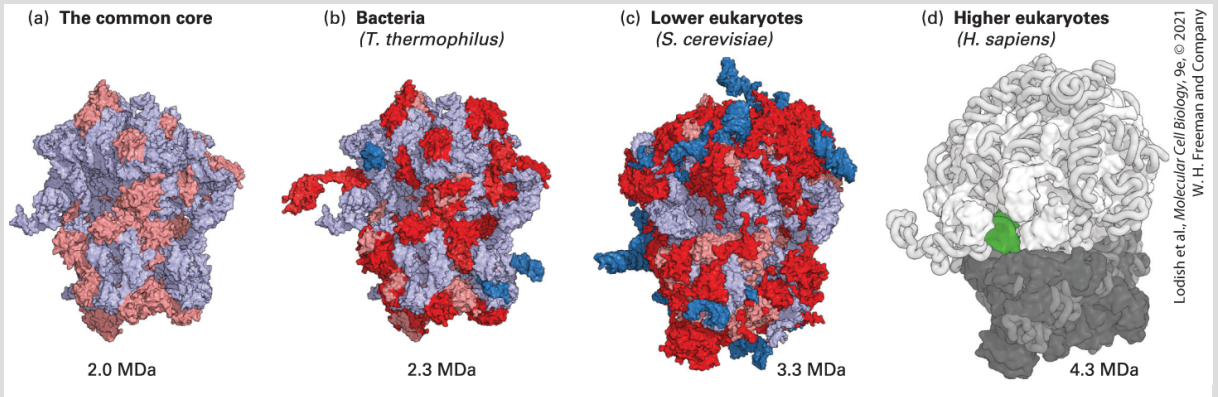


FIGURE 5-35 Comparison of the common core structure at the center of ribosomes from all domains of life and bacterial, yeast, and human ribosomes. (a) RNA in the common core structure is shown in light blue and protein domains common to all ribosomes are shown in pink. Additions to the common core structure are shown in dark blue for RNA and red for proteins in ribosomes from *T. thermophilus* (b) and *S. cerevisiae* (c). (d) Human ribosome structure from cryoelectron microscopy. A tRNA visible in the E site is shown in green.

[Parts (a, b, c) Data from G. Yusupova and M. Yusupov, 2014, *Annu. Rev. Biochem.* **83**:467; data for (d) from H. Khatter et al., 2015, *Nature* **520**:640; PDB ID 4ug0.]

The most therapeutically effective antibiotics target processes in bacterial cells that are essential for life and are fundamentally different between bacterial and mammalian cells. Because ribosomes satisfy both criteria, the bacterial ribosome has been a major target for the development of new antibiotics. The high-resolution structures of ribosomes are providing new insights into the mechanism by which many antibiotics inhibit bacterial protein synthesis without affecting the function of mammalian ribosomes. These insights are providing important clues for the design and synthesis of new antibiotics. Such research is desperately needed as the occurrence of bacteria resistant to currently available antibiotics

becomes increasingly more common, especially in hospitals, where antibiotic-resistant bacteria are under positive selection. ■

Methionyl-$tRNA_i^{Met}$ Recognizes the AUG Start Codon

As noted earlier, the AUG codon for methionine functions as the start codon in the vast majority of mRNAs. A critical aspect of translation initiation is to begin protein synthesis at the start codon, thereby establishing the correct reading frame for the entire mRNA. Both bacteria and eukaryotes contain two different methionine tRNAs: $tRNA_i^{Met}$ can initiate protein synthesis, and $tRNA^{Met}$ can incorporate methionine only into a growing protein chain. The same aminoacyl-tRNA synthetase (MetRS) charges both tRNAs with methionine, but *only* Met-$tRNA_i^{Met}$ (i.e., activated methionine attached to $tRNA_i^{Met}$) can bind at the appropriate site on the small ribosomal subunit, the *P site*, to begin synthesis of a polypeptide chain. The regular Met-$tRNA^{Met}$, and all other charged tRNAs, bind only to the *A site*, as described later. tRNAs enter the exit or *E site* after transferring their covalently bound amino acids to the growing polypeptide chain.

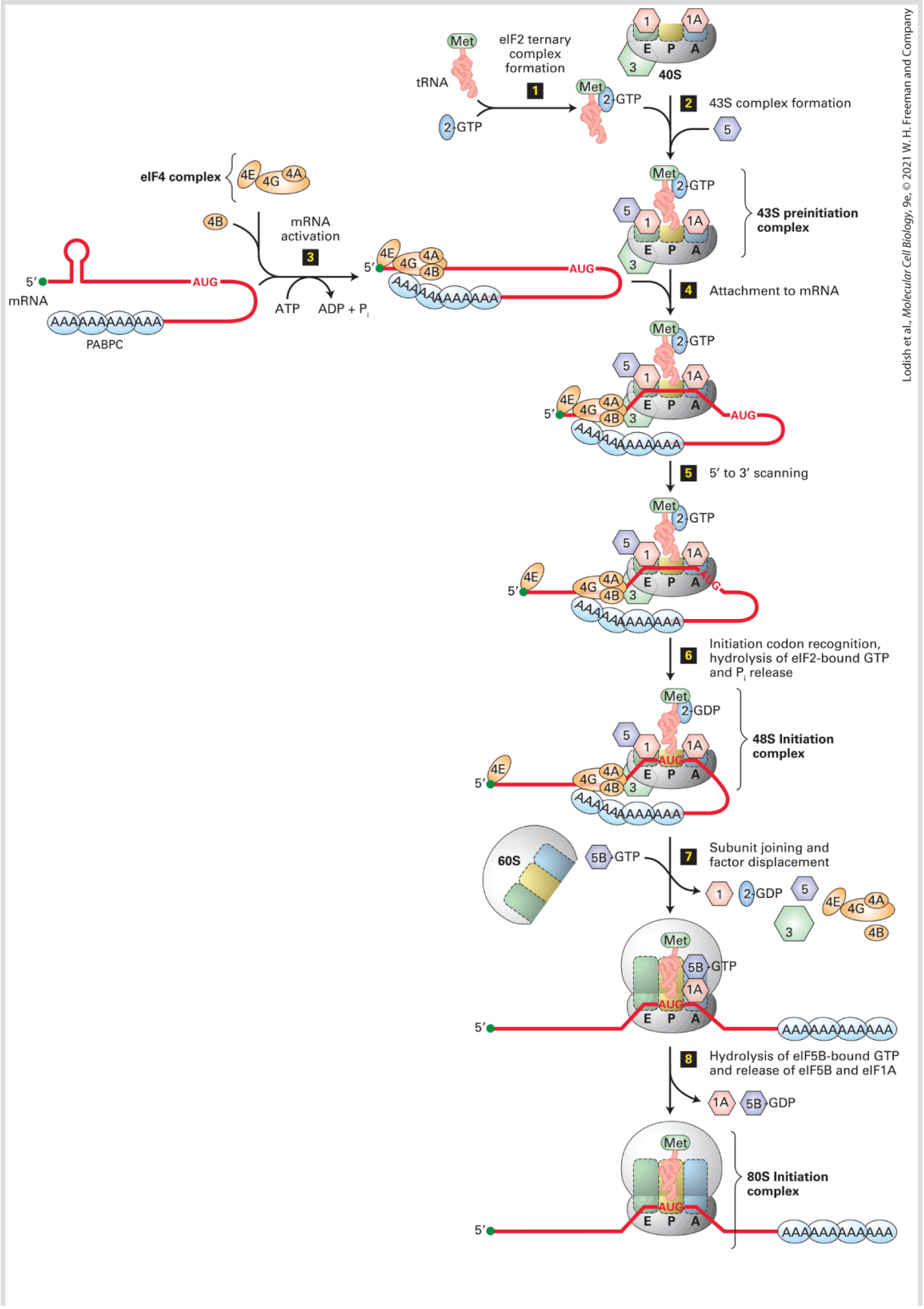
Eukaryotic Translation Initiation Usually Occurs at the First AUG

Downstream from the 5′ End of an mRNA

During the first stage of translation, the small and large ribosomal subunits assemble around an mRNA that has a Met-$tRNA_i^{Met}$ correctly positioned at the start codon in the ribosomal P site. In eukaryotes, the assembly of this complex is mediated by a special set of proteins known as **eukaryotic translation initiation factors (eIFs)**. As each individual component joins the complex, it is guided by interactions with specific eIFs. Several of the initiation factors bind GTP, and the hydrolysis of GTP to GDP functions as a switch to ensure that the assembly pathway proceeds in an orderly and irreversible fashion. By coupling of assembly to the essentially irreversible step of GTP hydrolysis ensures that, once formed, the appropriate complexes are stable.

The current model for initiation of translation in vertebrates is depicted in Figure 5-36. Large and small ribosomal subunits released from a previous round of translation are kept apart by the binding of eIFs 1, 1A, and 3 to the small 40S subunit (Figure 5-36, *top*). The first step of translation initiation is formation of a *43S preinitiation complex*. This preinitiation complex is formed when the 40S subunit with eIFs 1, 1A, and 3 associates with eIF5 and a ternary (three-part) complex consisting of the Met-$tRNA_i^{Met}$ and eIF2 bound to GTP (Figure 5-36, steps **1** and **2**). The initiation factor eIF2 alternates between association with GTP and GDP; it can bind

Met-$tRNA_i^{Met}$ only when it is associated with GTP. Cells can inhibit protein synthesis by phosphorylating a serine residue on the eIF2 bound to GDP; the phosphorylated complex is unable to exchange the bound GDP for GTP and cannot bind Met-$tRNA_i^{Met}$, so protein synthesis cannot occur.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 5-36 Initiation of translation in eukaryotes. The current model of eukaryotic initiation involves eight steps. Step **1**: An eIF2 ternary complex forms when eIF2 · GTP binds a $\mathrm{tRNA}_{\mathrm{i}}^{\mathrm{Met}}$. Step **2**: When a ribosome dissociates at the termination of translation, the 40S subunit is bound by eIF1, eIF1A, and eIF3. A 43S preinitiation complex forms when this subunit associates with an eIF2 ternary complex and eIF5. Step **3**: An mRNA is activated when a multisubunit eIF4 complex binds: subunit eIF4E binds to the $5'$ cap, and subunit eIF4G binds multiple copies of the cytoplasmic poly(A)-binding protein (PABPC) bound to the mRNA poly(A) tail. For simplicity, binding of only one PABPC to eIF4G is shown. Then eIF4B, which stimulates eIF4A helicase activity, also joins this circular complex in which both the mRNA $5'$ cap and poly(A) tail are associated with the eIF4 complex. Step **4**: The 43S preinitiation complex binds an eIF4-mRNA complex. Step **5**: The RNA helicase activity of subunit eIF4A unwinds any RNA secondary structure at the $5'$ end of the mRNA as the 40S subunit scans in the $5' \rightarrow 3'$ direction until it recognizes the initiation codon. For simplicity, eIF4E is diagrammed as releasing from the remainder of the eIF4 complex, but in reality, it remains associated, forming a loop in the mRNA between the $5'$ cap and the scanning eIF4 complex. Step **6**: Recognition of the initiation codon causes eIF5 to stimulate hydrolysis of eIF2-bound GTP. This switches the conformation of the scanning complex to a 48S initiation complex with the anticodon of $\mathrm{tRNA}_{\mathrm{i}}^{\mathrm{Met}}$ base-paired to the initiator AUG in the 40S-subunit P site. Step **7**: The 60S subunit joins the 40S subunit, leading to the release of most of the earlier-acting eIFs as eIF5B-GTP binds to eIF1A in the ribosomal A site. The released eIF4 complex and eIF4B associate with the $5'$ cap and PABPC as shown in step **3** to prepare for interaction with another 43S preinitiation complex. For simplicity, this association is not shown. Step **8**: Correct association of the 40S and 60S subunits results in hydrolysis of eIF5B-bound GTP, release of eIF5B-GDP and eIF1A, and formation of the 80S initiation complex with $\mathrm{tRNA}_{\mathrm{i}}^{\mathrm{Met}}$ base-paired to the initiation codon in the ribosomal P site. See R. J. Jackson et al., 2010, *Nat. Rev. Mol. Cell Biol.* **11**:113.

The mRNA to be translated is bound by the multisubunit eIF4 complex, which interacts with both the $5'$ cap and the cytoplasmic poly(A)-binding protein (PABPC) bound in multiple copies to the mRNA poly(A) tail. Both interactions are required for translation of most mRNAs. This binding results in the formation of a circular

complex (Figure 5-36, step **3**). The eIF4 cap-binding complex consists of several subunits with different functions. The eIF4E subunit binds the 5′ cap on mRNAs (see Figure 5-13). The large eIF4G subunit binds cooperatively to several PABPC proteins bound to the mRNA poly(A) tail and also forms a scaffold to which the other eIF4 subunits bind. The mRNA-eIF4 complex then associates with the preinitiation complex through an interaction between eIF4G and eIF3 (step **4**).

The initiation complex slides along, or *scans*, the associated mRNA, with eIF4A, stimulated by eIF4B, unwinding the RNA secondary structure by using energy from ATP hydrolysis (step **5**). Scanning stops when the $tRNA_i^{Met}$ anticodon recognizes the start codon, which is the first AUG downstream from the 5′ end in most eukaryotic mRNAs. Recognition of the start codon leads to hydrolysis of the GTP associated with eIF2, an irreversible step that prevents further scanning, resulting in formation of the *48S initiation complex* (step **6**). This commitment to the correct initiation codon is facilitated by eIF5, an eIF2 GTPase–activating protein (GAP, see Figure 3-35). Selection of the initiating AUG is facilitated by specific surrounding nucleotides called the *Kozak sequence*, for Marilyn Kozak, who defined it: (5′)ACC**AUGG**(3′). The A preceding the AUG (in bold) and the G immediately following it are the most important nucleotides affecting translation initiation efficiency.

Association of the large (60S) subunit with the small subunit, which is mediated by eIF5B bound to GTP, results in displacement of many of the initiation factors (step **7**). Correct association between the ribosomal subunits results in hydrolysis of the eIF5B-bound GTP to GDP and the release of eIF5B-GDP and eIF1A (step **8**), completing the formation of an *80S initiation complex*. Coupling of the ribosome subunit–joining reaction to GTP hydrolysis by eIF5B allows the initiation process to continue only when the subunit interaction has occurred correctly. It also makes this an irreversible step, so that the ribosomal subunits do not dissociate until the entire mRNA is translated and protein synthesis is terminated.

The eukaryotic protein-synthesizing machinery begins translation of most cellular mRNAs within about 100 nucleotides of the 5′-capped end, as just described. However, some cellular mRNAs contain an internal ribosome entry site (IRES) located far downstream from the 5′ end. It is thought that cellular IRESs form RNA structures that interact with a complex of eIF4A and eIF4G, which then associates with eIF3 bound to a 40S subunit with eIF1 and eIF1A. This assembly then binds an eIF2 ternary complex to assemble an initiation complex directly on a neighboring AUG codon. In addition, translation of some viral RNAs, which lack a 5′ cap, is initiated at viral IRES sequences. These RNAs fall into different classes depending on how many of the standard eIFs are required for initiation. In the case of cricket paralysis virus, the ~200-nt-long IRES folds into a complex structure that interacts

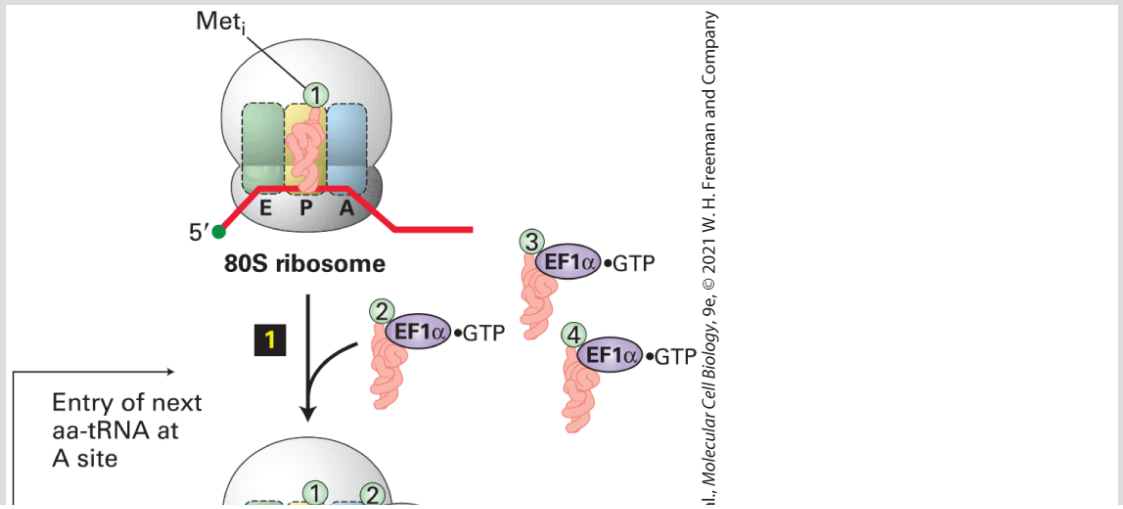
directly with the 40S ribosomal subunit and leads to initiation without any of the eIFs or even the initiator Met-$tRNA_i^{Met}$.

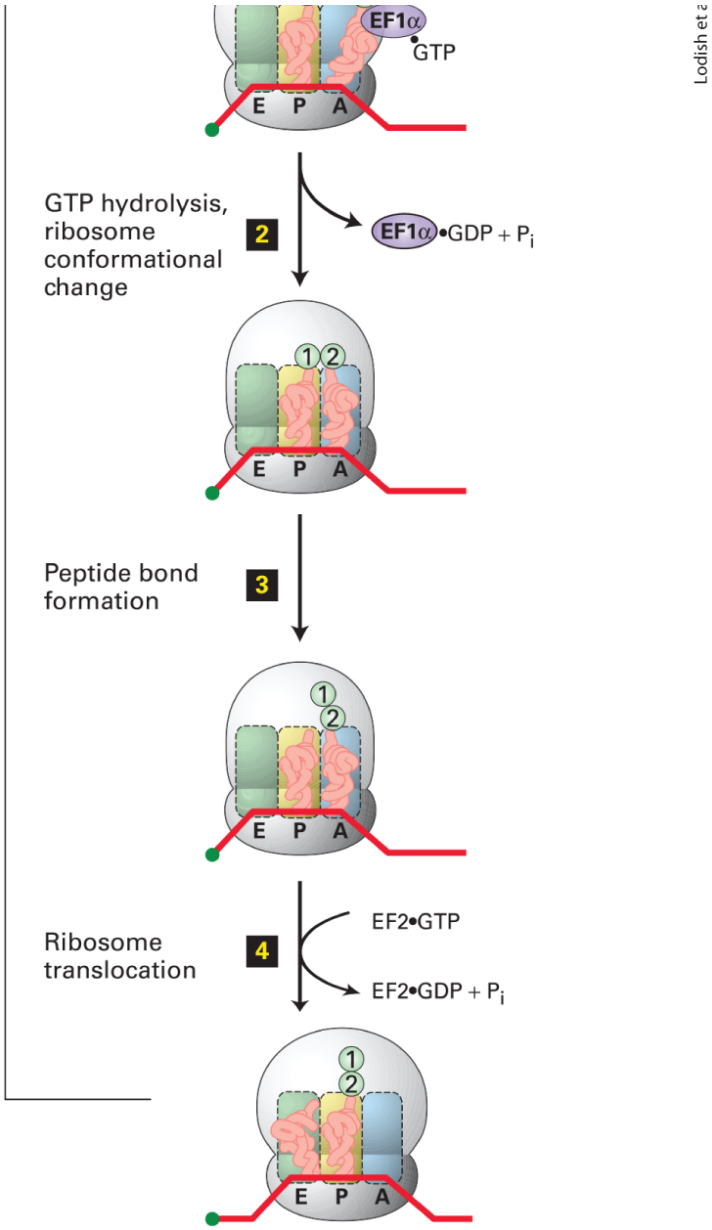
During Chain Elongation Each Incoming Aminoacyl-tRNA Moves Through Three Ribosomal Sites

The correctly positioned ribosome–Met-$tRNA_i^{Met}$ complex is now ready to begin the task of stepwise addition of amino acids by in-frame translation of the mRNA. As is the case with initiation, a set of specialized proteins, termed translation **elongation factors (EFs)**, is required to carry out this process of chain elongation. The key steps in elongation are the entry of each succeeding aminoacyl-tRNA with an anticodon complementary to the next codon, the formation of a peptide bond, and the movement, or *translocation*, of the ribosome one codon at a time along the mRNA.

At the completion of translation initiation, as noted already, Met-$tRNA_i^{Met}$ is bound to the P site on the assembled 80S ribosome (Figure 5-37, *top*). This region of the ribosome is called the *P* site because the tRNA chemically linked to the growing poly*p*eptide chain is located here. The second aminoacyl-tRNA is brought into the ribosome as a ternary complex in association with EF1α·GTP and becomes bound to the *A* site, so named because it is where *a*minoacyl-tRNAs bind (step **1**). EF1α·GTP bound to various

aminoacyl-tRNAs diffuse into the A site, but the next step in translation proceeds only when the tRNA anticodon base-pairs with the second codon in the coding region. When that occurs properly, the GTP in the associated EF1α·GTP is hydrolyzed. The hydrolysis of GTP promotes a conformational change in EF1α that leads to release of the resulting EF1α·GDP complex and tight binding of the aminoacyl-tRNA in the A site (step **2**). This conformational change also positions the aminoacylated 3′ end of the tRNA in the A site close to the 3′ end of the Met-$tRNA_i^{Met}$ in the P site. GTP hydrolysis, and hence tight binding, does not occur if the anticodon of the incoming aminoacyl-tRNA cannot base-pair with the codon at the A site. In this case, the ternary complex diffuses away, leaving an empty A site that can associate with other aminoacyl-tRNA–EF1α·GTP complexes until a correctly base-paired tRNA is bound. Thus GTP hydrolysis by EF1α is another proofreading step that allows protein synthesis to proceed only when the correct aminoacyl-tRNA is bound to the A site. This phenomenon contributes to the fidelity of protein synthesis.





Lodish et a

FIGURE 5-37 Chain elongation in eukaryotes. Once the 80S ribosome with Met-$\mathrm{tRNA_i^{Met}}$ in the ribosome P site is assembled (*top*), a ternary complex bearing the second amino acid ($\mathrm{aa_2}$) coded by the mRNA binds to the A site (step **1**). Following a conformational change in the ribosome induced by hydrolysis of GTP in EF1α·GTP (step **2**), the large rRNA catalyzes peptide bond formation between $\mathrm{Met_i}$ and $\mathrm{aa_2}$ (step **3**). Hydrolysis of GTP in EF2·GTP causes another conformational change in the ribosome that results in its translocation one codon along the mRNA and shifts the unacylated $\mathrm{tRNA_i^{Met}}$ to the E site and the tRNA with the bound peptide to the P site (step **4**). The cycle can begin again with binding of a ternary complex bearing

aa_3 to the now open A site. In the second and subsequent elongation cycles, the tRNA at the E site is ejected during step **2** as a result of the conformational change induced by hydrolysis of GTP in EF1α·GTP.

With the initiating Met-$tRNA_i^{Met}$ at the P site and the second aminoacyl-tRNA tightly bound at the A site, the α-amino group of the second amino acid reacts with the *activated* (ester-linked) methionine on the initiator tRNA, forming a peptide bond (Figure 5-37, step **3**; see Figure 5-29). Since the vast majority of cellular catalysts are protein enzymes, it was long assumed that the catalyst for peptide transfer would be one of the ribosomal proteins. It therefore came as quite a surprise when the high-resolution crystal structure of the bacterial large ribosomal subunit revealed that no proteins lie near the site of peptide bond synthesis and that the peptidyltransferase reaction was likely catalyzed by the large rRNA. The catalytic ability of the large rRNA in bacteria has been demonstrated by carefully removing the vast majority of the protein from large ribosomal subunits. The nearly pure bacterial 23S rRNA can catalyze a peptidyltransferase reaction between analogs of aminoacyl-tRNA and peptidyl-tRNA. The 2′-hydroxyl of the terminal A of the peptidyl-tRNA in the P site also participates in catalysis.

Following peptide bond synthesis, the ribosome translocates a distance equal to one codon along the mRNA. This translocation step is monitored by hydrolysis of the GTP in eukaryotic EF2·GTP. Once translocation has occurred correctly, the bound GTP is hydrolyzed, another irreversible process that prevents the ribosome

from moving along the RNA in the wrong direction or from translocating an incorrect number of nucleotides. As a result of conformational changes in the ribosome that accompany proper translocation and the resulting GTP hydrolysis by EF2, $tRNA_i^{Met}$, now without its activated methionine, is moved to the *E (e*xit) site on the ribosome; concurrently, the second tRNA, now covalently bound to a dipeptide (a peptidyl-tRNA), is moved to the P site (Figure 5-37, step **4**). Translocation thus returns the ribosome conformation to a state in which the A site is open and able to accept another aminoacyl-tRNA complexed with EF1α·GTP, beginning another cycle of chain elongation.

Repetition of the elongation cycle depicted in Figure 5-37 adds amino acids one at a time to the carboxyl terminus of the growing polypeptide, as directed by the mRNA sequence, until a stop codon is encountered. In subsequent cycles, the conformational change that occurs in step **2** ejects the unacylated tRNA from the E site. As the nascent polypeptide chain becomes longer, it threads through a channel in the large ribosomal subunit, exiting at a position opposite the side that interacts with the small subunit (Figures 5-34).

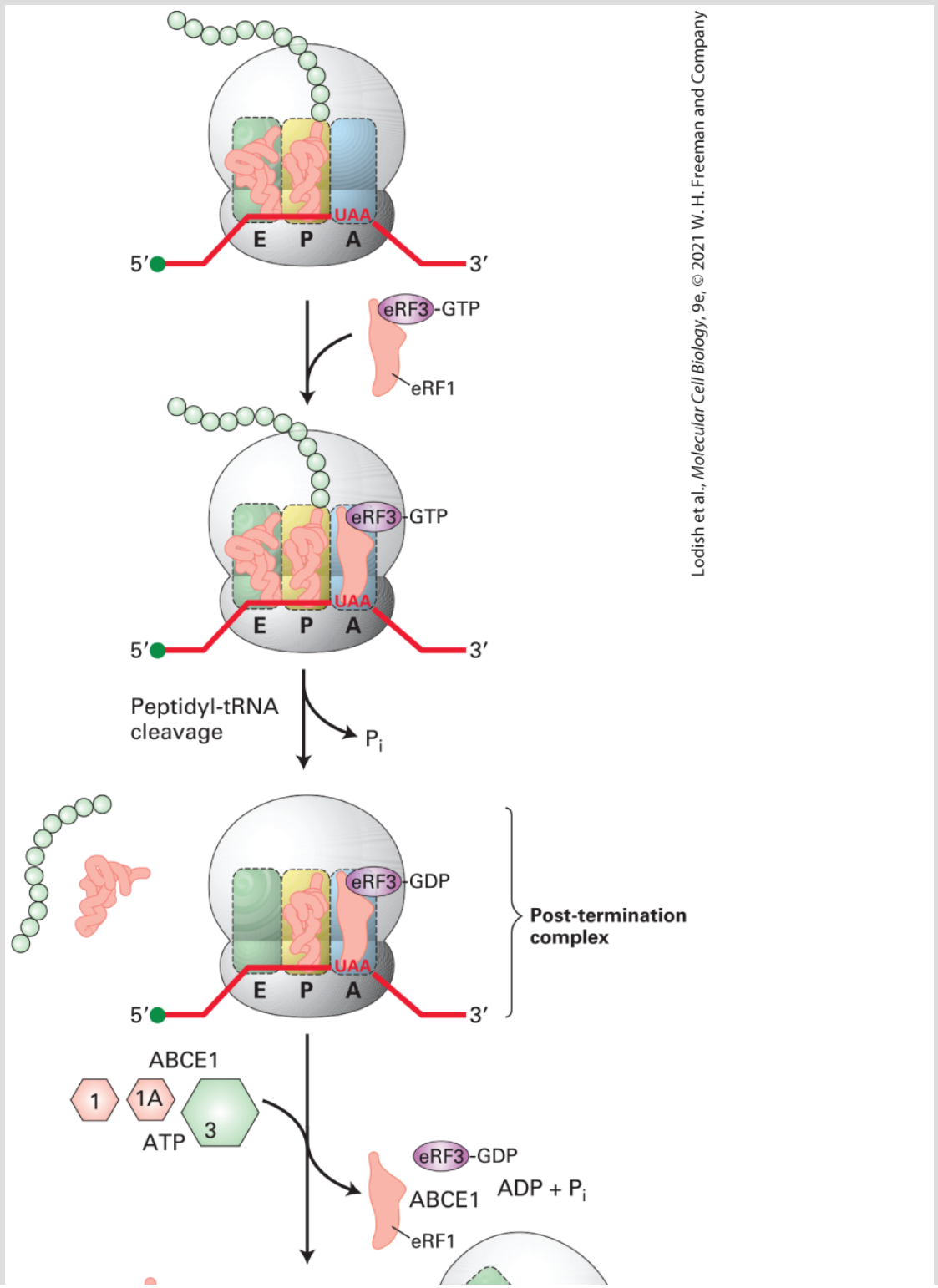
In the absence of the ribosome, the three-base-pair RNA-RNA hybrid between the tRNA anticodons and the mRNA codons in the A and P sites would not be stable; RNA-RNA duplexes between separate RNA molecules must be considerably longer to be stable under physiological conditions. However, multiple interactions between

the large and small rRNAs and the general domains of tRNAs (e.g., the D and TψCG loops, see Figure 5-32) stabilize the tRNAs in the A and P sites, while other RNA-RNA interactions sense correct codon-anticodon base pairing, ensuring that the genetic code is read properly. Then, interactions between rRNAs and the general domains of all tRNAs result in the movement of the tRNAs between the A, P, and E sites as the ribosome translocates along the mRNA one three-nucleotide codon at a time.

Translation Is Terminated by Release Factors When a Stop Codon Is Reached

The final stages of translation, like initiation and elongation, require highly specific molecular signals that decide the fate of the mRNA–ribosome–peptidyl-tRNA complex. Two types of eukaryotic protein **release factors (RFs)** have been discovered. The first, eRF1, whose shape is similar to that of tRNAs, acts by binding to the ribosomal A site and recognizing stop codons directly. Like some of the initiation and elongation factors discussed previously, the second release factor, eRF3, is a GTP-binding protein. The eRF3·GTP complex acts in concert with eRF1 to promote cleavage of the peptidyl-tRNA bond, thus releasing the completed protein chain and terminating translation (Figure 5-38). The peptidyl-tRNA bond of the tRNA in the P site is not cleaved until one of the three stop codons is correctly

recognized by eRF1, another example of a proofreading step in protein synthesis.



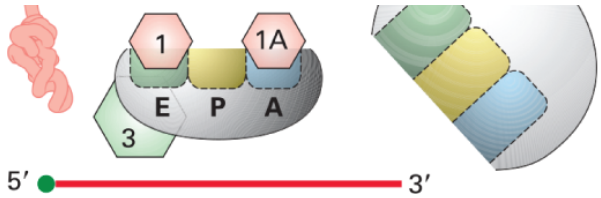


FIGURE 5-38 Termination of translation in eukaryotes. When a ribosome bearing a nascent protein chain reaches a stop codon (UAA, UGA, UAG), release factor eRF1 enters the A site together with eRF3·GTP. Hydrolysis of the bound GTP is accompanied by cleavage of the peptide chain from the tRNA in the P site and ejection of the tRNA in the E site, forming a post-termination complex. The ribosomal subunits are separated by the action of the ABCE1 ATPase together with eIF1, eIF1A, and eIF3. The 40S subunit is released bound to these eIFs, ready to initiate another cycle of translation (see Figure 5-36).

Release of the completed protein leaves a free tRNA in the P site and the mRNA still associated with the 80S ribosome, to which eRF1 and eRF3·GDP are still bound in the A site. In eukaryotes, ribosome recycling occurs when this post-termination complex is bound by a protein called *ABCE1*, which uses energy from ATP hydrolysis to separate the subunits and release the mRNA and tRNA in the P site. Initiation factors eIF1, eIF1A, and eIF3, which are also required for separation of the subunits, load onto the 40S subunit, making it ready for another round of initiation (see Figure 5-36, *top*). In reality, a free mRNA is never released, as diagrammed in Figure 5-25 for simplicity. Rather, the mRNA has other ribosomes associated with it in various stages of elongation, PABPC bound to the poly(A) tail, and the eIF4 complex associated with the 5′ cap, ready to associate with another 43S preinitiation complex (see Figure 5-36).

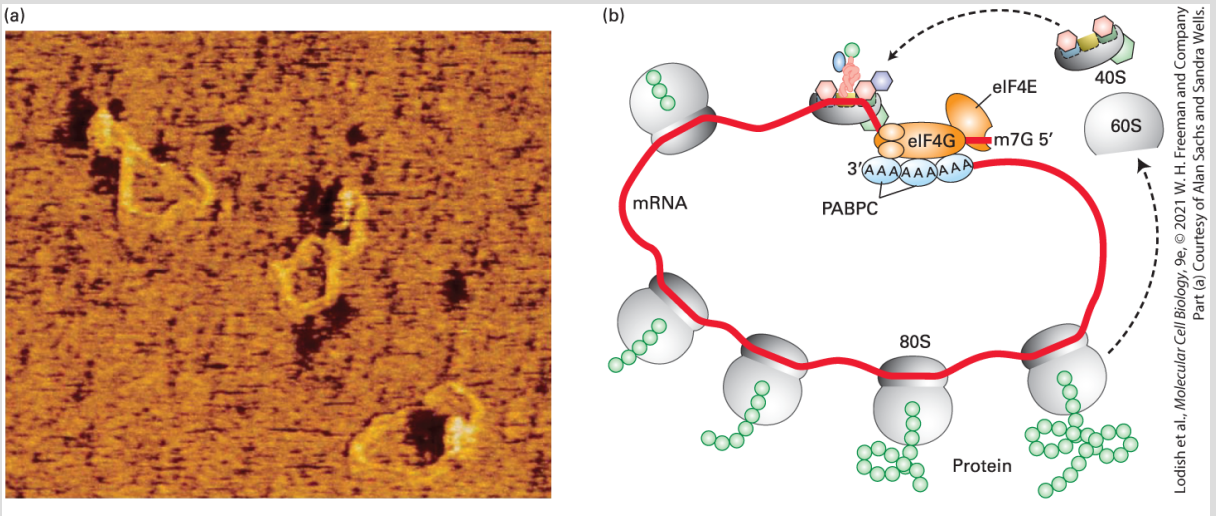
In addition to these functions in protein synthesis, ribosomes also associate transiently with protein chaperones that assist in folding the polypeptide chain as it emerges from the ribosome surface (see Figure 3-19). As we will see in Chapter 13, ribosomes that synthesize proteins destined to be inserted into the endoplasmic reticulum (ER) associate with a ribonucleoprotein complex called *SRP* (signal recognition particle) that arrests protein synthesis until the nascent polypeptide encounters specialized channels for insertion into the ER. SRP also assists with the insertion and threading of these proteins through these ER channels when protein synthesis is permitted to resume.

Polysomes and Rapid Ribosome Recycling Increase the Efficiency of Translation

Translation of a single eukaryotic mRNA molecule to yield a typically sized protein takes 1 to 2 minutes. Two phenomena significantly increase the overall rate at which cells can synthesize a protein: the simultaneous translation of a single mRNA molecule by multiple ribosomes, and rapid recycling of ribosomal subunits after they disengage from a stop codon. Simultaneous translation of an mRNA by multiple ribosomes is readily observable in electron micrographs and by sedimentation analysis, revealing mRNA molecules attached to multiple ribosomes bearing nascent growing

polypeptide chains. These structures, referred to as **polyribosomes** or *polysomes*, were seen to be circular in electron micrographs of some tissues.

Subsequent studies with purified initiation factors explained the circular shape of polyribosomes and suggested the mechanism by which ribosomes recycle efficiently. These studies revealed that multiple copies of the cytoplasmic poly(A)-binding protein (PABPC) interact with both an mRNA poly(A) tail and the eIF4G subunit of eIF4. Since the eIF4E subunit of eIF4 binds to the cap structure on the 5′ end of an mRNA, the two ends of an mRNA molecule are bridged by the intervening proteins, forming circular mRNA (Figure 5-39a). Because the two ends of a polysome are relatively close together, ribosomal subunits that disengage from a stop codon are positioned near the 5′ end, facilitating reinitiation by the interaction of the 40S subunit and its associated initiation factors with eIF4 bound to the 5′ cap. The circular pathway depicted in Figure 5-39b is thought to enhance ribosome recycling and thus increase the efficiency of protein synthesis.



EXPERIMENTAL FIGURE 5-39 Circular structure of mRNA increases translation efficiency. Eukaryotic mRNA forms a circular structure owing to interactions of three proteins. (a) In the presence of purified yeast poly(A)-binding protein [PABP; there is only one PABP in *S. cerevisiae*, rather than a nuclear (PABPN) and cytoplasmic (PABPC) protein as in higher eukaryotes], eIF4E, and eIF4G, eukaryotic mRNAs form circular structures, visible in this force-field electron micrograph. In these structures, protein-protein and protein-mRNA interactions form a bridge between the $5'$ and $3'$ ends of the mRNA. (b) Model of protein synthesis on circular polysomes and recycling of ribosomal subunits. Multiple individual ribosomes can simultaneously translate a eukaryotic mRNA, shown here in a circular form stabilized by interactions between proteins bound at the $3'$ and $5'$ ends. When a ribosome completes translation and dissociates from the $3'$ end, the separated subunits can rapidly find the nearby $5'$ cap $(\mathrm{m}^7\mathrm{G})$ and PABPC-bound poly(A) tail and initiate another round of synthesis.

GTPase-Superfamily Proteins Function in Several Quality-Control Steps of Translation

We have now seen that one or more GTP-binding proteins participate in each stage of translation. These proteins belong to the **GTPase superfamily** of switch proteins that cycle between a GTP-bound form and a GDP-bound form (see Figure 3-35). Usually correct execution of some preceding assembly step is required to trigger hydrolysis of the bound GTP. The hydrolysis of bound GTP to bound GDP is essentially irreversible and causes a conformational change in the GTPase itself, thus ensuring that the next interdependent assembly step associated with translation proceeds rapidly, unidirectionally, and with high fidelity. As the details of translation have been worked out, two distinct but not mutually exclusive functions of GTP hydrolysis have been identified. The first is to make an assembly event unidirectional by locking associated proteins into place once components have come together in the correct manner. A good example of this is successful association of the large and small ribosomal subunits coupled to hydrolysis of GTP by eIF5B (Figure 5-36, step **8**). Recall that if the correct complex does not form, eIF5B·GTP does not hydrolyze the GTP and the complex is unstable, free to diffuse apart and try again. When the precise alignment of the subunits required for elongation occurs, eIF5B hydrolyzes the GTP to GDP, locking the correct complex in place. Energy released from the high energy β-γ bond in GTP drives the reaction in one direction. A second function of GTP hydrolysis is to create an additional step by which the correct fit for a molecular association can be checked to increase the fidelity of translation. The best example of such a proofreading activity occurs during the crucial process of matching the correct aminoacyl-tRNA with the

next available codon in an mRNA. An initial match is made when the anticodon of an aminoacyl-tRNA in complex with EF1α·GTP docks with the codon in the A site. Then as a distinct step, hydrolysis of EF1α·GTP to EF1α·GDP occurs only after correct base pairing between anticodon and codon is confirmed. GTP hydrolysis causes a conformational change in EF1α that results in the release of its bound tRNA, allowing the aminoacylated 3′ end of the charged tRNA to move into the position required for peptide bond formation (see Figure 5-37, step **2**). In effect, the base pairing between anticodon and codon is checked twice; first in the initial selection of an amino acyl-tRNA and again in a step coupled to hydrolysis of EF1α·GTP. Molecular simulations estimate that this second proofreading step increases the fidelity of translation about 15-fold.

Nonsense Mutations Can Be Bypassed by Suppressing tRNA Mutations

As we have seen, a mutation that converts a codon normally encoding an amino acid into a stop codon will lead to premature termination of translation. Nonsense mutations, as they are called, result in a truncated protein that is usually nonfunctional. In genetic studies with the bacterium *E. coli*, it was discovered that the effect of a nonsense mutation can be suppressed by a second mutation in a tRNA gene. This occurs when the sequence in a tRNA gene that encodes the anticodon is changed to a triplet that is complementary

to a stop codon. If, for example, a nonsense mutation in a gene converted a UCG (serine) codon into a UAG (stop) codon, the stop may be read often enough to produce some full-length protein if the anticodon of a tRNA were changed to CUA, which can base-pair with the UAG stop codon. One way to do this would be to mutate the gene for $tRNA^{Tyr}$ in a way that changes its anticodon from GUA to CUA. Importantly, the mutant $tRNA^{Tyr}$ can still be recognized by the tyrosine aminoacyl-tRNA synthetase, since the anticodon sequence is not a necessary feature for recognition by the corresponding aminoacyl-tRNA synthetase. Cells that have both the original nonsense mutation and the second mutation in the anticodon of the $tRNA^{Tyr}$ gene consequently can insert a tyrosine at the position of the mutant stop codon, allowing protein synthesis to continue past the original nonsense mutation. This mechanism is not highly efficient, so translation of normal mRNAs with a UAG stop codon terminates at the normal position in most instances. If enough of the protein encoded by the original gene with the nonsense mutation is produced to provide its essential functions, the effect of the first mutation is said to be *suppressed* by the second mutation in the anticodon of the tRNA gene.

This mechanism of *nonsense suppression* is a powerful tool in genetic studies of bacteria. For example, it allows us to isolate mutant bacterial viruses that cannot grow in normal cells but can grow in cells expressing a nonsense-suppressing tRNA because the mutant virus has a nonsense mutation in an essential gene. Such mutant

viruses grown on nonsense-suppressing cells can then be used in experiments to analyze the function of the mutant gene by infecting normal cells that do not suppress the mutation and analyzing what step in the viral life cycle is defective in the absence of the mutant protein.

KEY CONCEPTS OF SECTION 5.6

Stepwise Synthesis of Proteins on Ribosomes

- Bacterial, archaeal, and eukaryotic ribosomes — the large ribonucleoprotein complexes on which translation occurs — consist of a small and a large subunit. Each subunit contains numerous different proteins and one major rRNA molecule (small or large). The large subunit also contains one accessory 5S rRNA in bacteria and archaea and two accessory rRNAs in eukaryotes (5S and 5.8S).
- Analogous rRNAs from many different species fold into quite similar three-dimensional structures containing numerous stems and loops, as well as binding sites for proteins, mRNAs, and tRNAs. Much smaller ribosomal proteins are associated with the periphery of the rRNAs.
- Of the two methionine tRNAs found in all cells, only one $(\mathrm{tRNA_i^{Met}})$ functions in initiation of translation.
- Each stage of translation — initiation, chain elongation, and termination — requires specific protein factors, including GTP-binding proteins that hydrolyze their bound GTP to GDP when a step has been completed successfully.
- During initiation, the ribosomal subunits assemble near the translation start site in an mRNA molecule on which a tRNA carrying the amino-terminal methionine $(\mathrm{Met\text{-}tRNA_i^{Met}})$ has base-paired with the start codon (see Figure 5-36).
- Chain elongation entails a repetitive four-step cycle: loose binding of an incoming aminoacyl-tRNA to the A site on the ribosome; tight binding of the correct aminoacyl-tRNA to the A site accompanied by release of the previously used tRNA from the E site; transfer of the growing polypeptide chain to the incoming amino acid catalyzed by large rRNA; and translocation of the ribosome to the next codon, thereby moving the peptidyl-tRNA in the A site to the P site and the now unacylated tRNA in the P site to the E site (see Figure 5-38).

- In each cycle of chain elongation, the ribosome undergoes two conformational changes monitored by two GTP-binding proteins. The first of these proteins (EF1α) permits tight binding of the incoming aminoacyl-tRNA to the A site and ejection of a tRNA from the E site, and the second (EF2) monitors translocation.
- Termination of translation is carried out by two types of termination factors: those that recognize stop codons and those that promote hydrolysis of peptidyl-tRNA (see Figure 5-39). Once again, correct recognition of a stop codon is monitored by a GTPase (eRF3).
- The efficiency of protein synthesis is increased by the simultaneous translation of a single mRNA by multiple ribosomes, forming a polyribosome or polysome. In eukaryotic cells, protein-mediated interactions bring the two ends of a polyribosome close together, thereby promoting the rapid recycling of ribosomal subunits, which further increases the efficiency of protein synthesis (see Figure 5-39b).

5.7 Viruses: Parasites of the Cellular Genetic System

We wrap up this chapter with a discussion of **viruses**, which are radically different from the organisms we've so far discussed, but that nevertheless depend on fundamental genetic processes. Viruses carry instructions encoded in DNA or RNA that are copied by polymerization using standard Watson-Crick base pairing, and their genes are expressed according to the universal genetic code. The viral genome can be considered to be a self-replicating genetic program. In addition, viruses can mutate and they can evolve through the process of natural selection. Thus viruses have an autonomous existence in terms of the biological informational that they carry, but just as the instructions of a computer program cannot fully come into being without a computer to run on, the genetic instructions of a virus need a cellular host to synthesize viral proteins and to replicate the viral genome. Viruses are categorized according to the type of nucleic acid that makes up their genome. RNA viruses, which usually replicate in the host-cell cytoplasm, have an RNA genome, and DNA viruses, which commonly replicate in the host-cell nucleus, have a DNA genome (see Figure 5-1). Viral genomes may be single or double stranded, depending on the specific type of virus. The entire infectious virus particle, called a **virion**, consists of nucleic acid encased in a shell of protein, known as the **capsid**, which both protects the viral nucleic acid and functions in the process of host-cell infection. The genomes of the

simplest known viruses only encode four proteins; the most complex can encode some two hundred proteins. In addition to their obvious importance as causes of disease, viruses are extremely useful as research tools in the study of basic biological processes, such as those discussed in this chapter.

Most Viral Host Ranges Are Narrow

The process of infection usually begins when a protein on the surface of the virion binds to a specific protein on the surface of the host cell that the virus hijacks for use as a receptor. This interaction is usually highly specific and thus limits the *host range* — the types of cells that a virus can infect — to a very narrow range of cell types in particular species.

A virus that infects only bacteria is called a **bacteriophage**, or simply a **phage**. Viruses that infect animal or plant cells are referred to generally as animal viruses or plant viruses. Most animal viruses have host ranges restricted to certain phyla, and some infect only closely related species, such as primates. Some animal viruses, such as vesicular stomatitis virus, can infect insects as well as many different types of mammals. For viruses that can grow not only in plants or animals but also in the insects that feed on them, the highly mobile insects can serve as vectors for transferring the virus between susceptible hosts. The host-cell range of some animal viruses is further restricted to a limited number of cell types

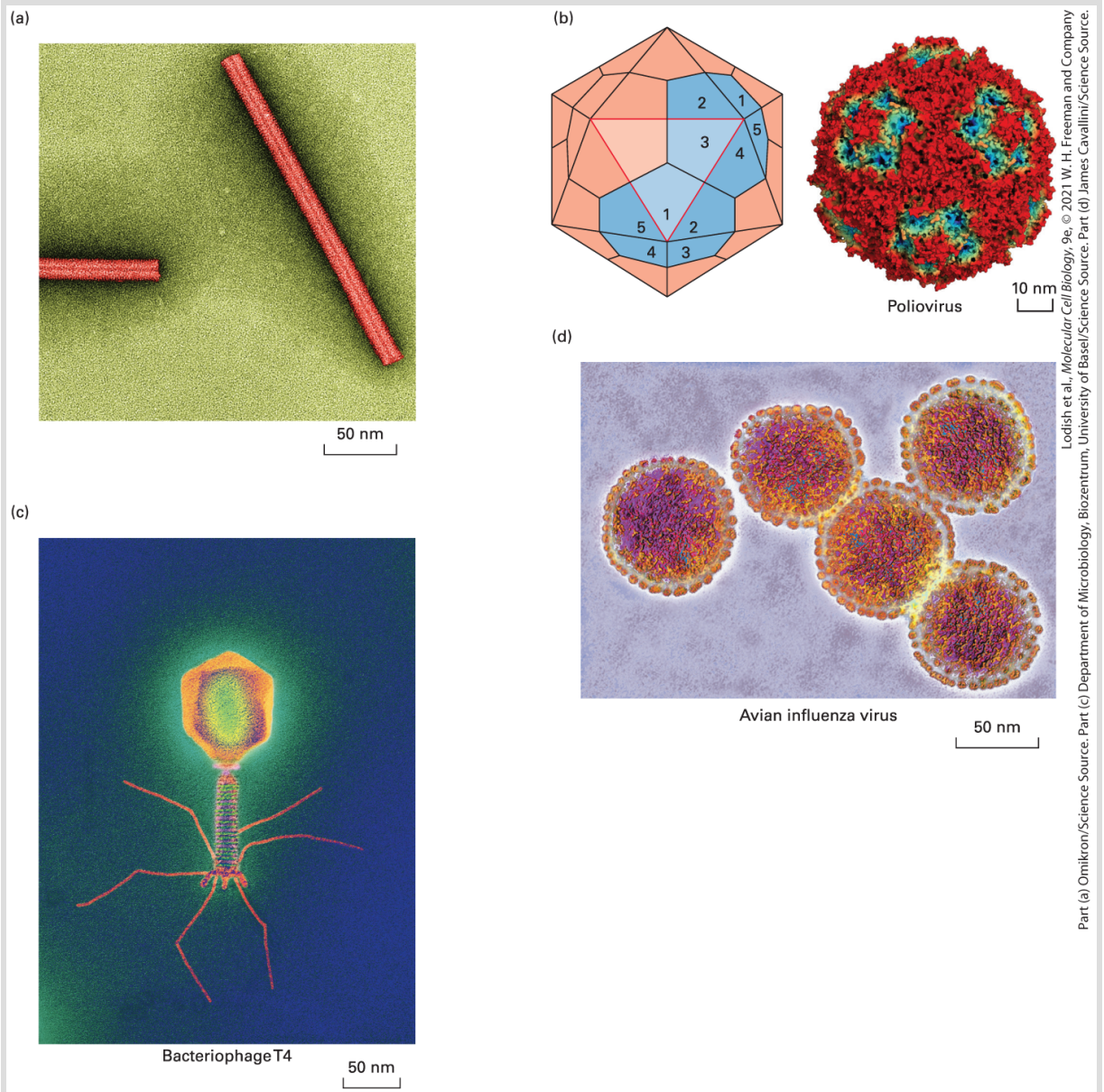
because only those cells have surface receptors to which the virions can attach. One example is poliovirus, which infects only cells in the intestine and, unfortunately for its host, motor neurons in the spinal cord, causing paralysis. Another is HIV-1, discussed further in this section, which infects cells called $CD4^+$ *T lymphocytes* that are essential for the immune response (see Chapter 24) as well as certain neurons and other cells of the central nervous system called *glial cells*.

Viral Capsids Are Regular Arrays of One or a Few Types of Protein

The viral capsid is usually composed of multiple copies of only one protein or a few different proteins, each of which is encoded by a single viral gene. Because of this structure, a virus is able to encode all the information for making a relatively large capsid in a small number of genes. This efficient use of genetic information is important because only a limited amount of DNA or RNA, and therefore a limited number of genes, can fit into a virion capsid.

Nature has found two basic ways of packaging a viral genome within a regular structure composed of multiple copies of capsid proteins. In some viruses, multiple copies of a single capsid protein form a *helical* structure that encloses and protects the viral RNA or DNA, which runs in a helical groove within the protein tube. Viruses with

such a helical structure, such as tobacco mosaic virus, have a rodlike shape (Figure 5-40a). The other major structural type is based on the *icosahedron*, a solid, approximately spherical object built of 20 identical faces (Figure 5-40b). Since each face is an equilateral triangle that must be composed of at least three protein subunits, an icosahedral capsid must be composed of at least 60 subunits. During infection, some icosahedral viruses interact with host-cell receptors via clefts between the capsid subunits. Others interact via long fibrous proteins extending from the vertices of the icosahedron.



Part (a) Omikron/Science Source. Part (c) Department of Microbiology, Biozentrum, University of Basel/Science Source. Part (d) James Cavallini/Science Source.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 5-40 Virion structures. (a) Helical tobacco mosaic virus. (b, *left*) Diagram of the structure of poliovirus, a small icosahedral virus, made of 20 equilateral triangular faces, one of which is outlined in red. Each face is composed of three outlined structural elements called *capsomeres*. The numbers show how five capsomeres associate at the 12 vertices of the icosahedron. (b, *right*) Space-filling model of poliovirus based on x-ray crystallography. The model is color-coded according to distance from the center of the virion, red farthest, blue closest. The virion binds to host cell receptors (not shown), which are long, narrow cell surface proteins that enter the blue “canyons” around each vertex. (c) Bacteriophage T4. (d) Influenza virus, an example of an enveloped virus.

Part (b) data from D. J. Filman et al., 1989. *EMBO J.* **8**:1567, PDB ID 2plv.

In many DNA bacteriophages, the viral DNA is located within an icosahedral *head* that is attached to a rodlike *tail*. During infection, viral proteins at the tip of the tail bind to host-cell receptors, and the viral DNA then passes down the tail into the cytoplasm of the host cell (Figure 5-40c).

In some viruses, a symmetrically arranged **nucleocapsid** composed of the viral genome associated with multiple copies of one or a few proteins is covered by an external membrane, or **viral envelope**, which consists mainly of a phospholipid bilayer but also contains one or more types of virus-encoded glycoproteins (Figure 5-40d). The phospholipids in the viral envelope are similar to those in the plasma membrane of an infected host cell. The viral envelope is, in fact, derived by budding from that membrane, but contains mainly viral glycoproteins, as we will discuss shortly.

Lytic Viral Growth Cycles Lead to Death of Host Cells

Although details vary among different types of viruses, those that exhibit a *lytic cycle* of growth proceed through the following stages:

1. *Adsorption* — Virion interacts with a host cell by binding of multiple copies of capsid protein to specific receptors on the

cell surface.

2. *Entry* — The virus enters the cytoplasm of the host cell. Membrane enveloped viruses may enter through the endocytic pathway described in Chapter 14 while others fuse directly with the host plasma membrane. Viruses with a protein capsid must enter cells through the endocytic pathway so as not to disrupt the integrity of the plasma membrane.
3. *Early gene expression* — The first viral genes to be expressed are usually involved in the replication of the viral genome and the efficient expression of viral structural genes. For positive strand RNA viruses, early genes are expressed by direct translation of the genomic RNA in the cytoplasm. For DNA viruses, viral DNA enters the cell nucleus where viral mRNAs are produced with the aid of the host-cell transcription machinery. Viral mRNAs are translated by the host-cell translation machinery.
4. *Genome replication* — Production of multiple copies of the viral genome is carried out either by viral proteins alone or with the help of host-cell proteins. RNA viruses usually replicate in the cytoplasm whereas DNA viruses must replicate within the nucleus.
5. *Structural gene expression* — The virus uses host cell machinery for transcription and translation of viral structural proteins. These include proteins of the nucleocapsid and abundant membrane glycoprotein for membrane enveloped viruses.
6. *Assembly* — Viral proteins and replicated genomes associate to form progeny virions in the cytoplasm.
7. *Release* — The host cell may rupture suddenly (**lysis**), releasing all the newly formed virions at once, or disintegrate gradually,

releasing the virions slowly. For membrane enveloped viruses the last stage of assembly involves entry into membrane bounded organelles the secretory pathy and then release from the cell by exocytosis or by budding directly from the plasma membrane.

We will use coronavirus, which is the cause of the COVID-19 pandemic, as an example to show how a lytic virus navigates the compartments of a eukaryotic cell. Note that the features of the coronavirus life cycle depicted in Figure 5-41 are governed by the basic properties of the coronavirus as a membrane enveloped virus with an RNA genome. Coronavirus begins entry into a target cell by binding of the spike protein of the viral envelope to a receptor protein on the surface of the host cell. Binding to the receptor allows the virus to enter endocytic vesicles which bud into the interior of the cell as described in Chapter 14. Within an endocytic compartment, the relatively low pH triggers a proteolytic cleavage of the viral spike protein, activating the spike protein for fusion of the viral membrane with the endosome membrane and releasing the nucleocapsid into the cytoplasm.

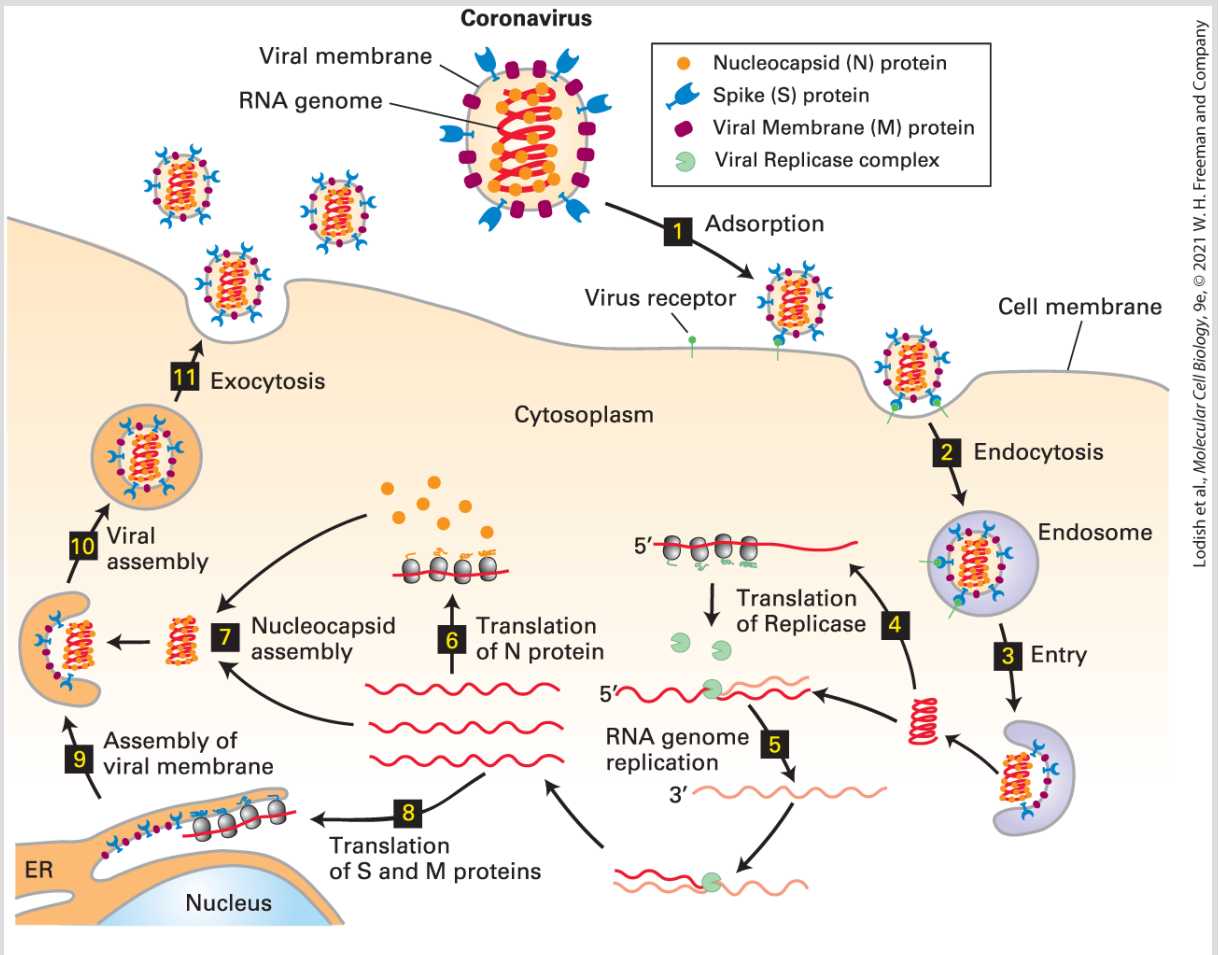


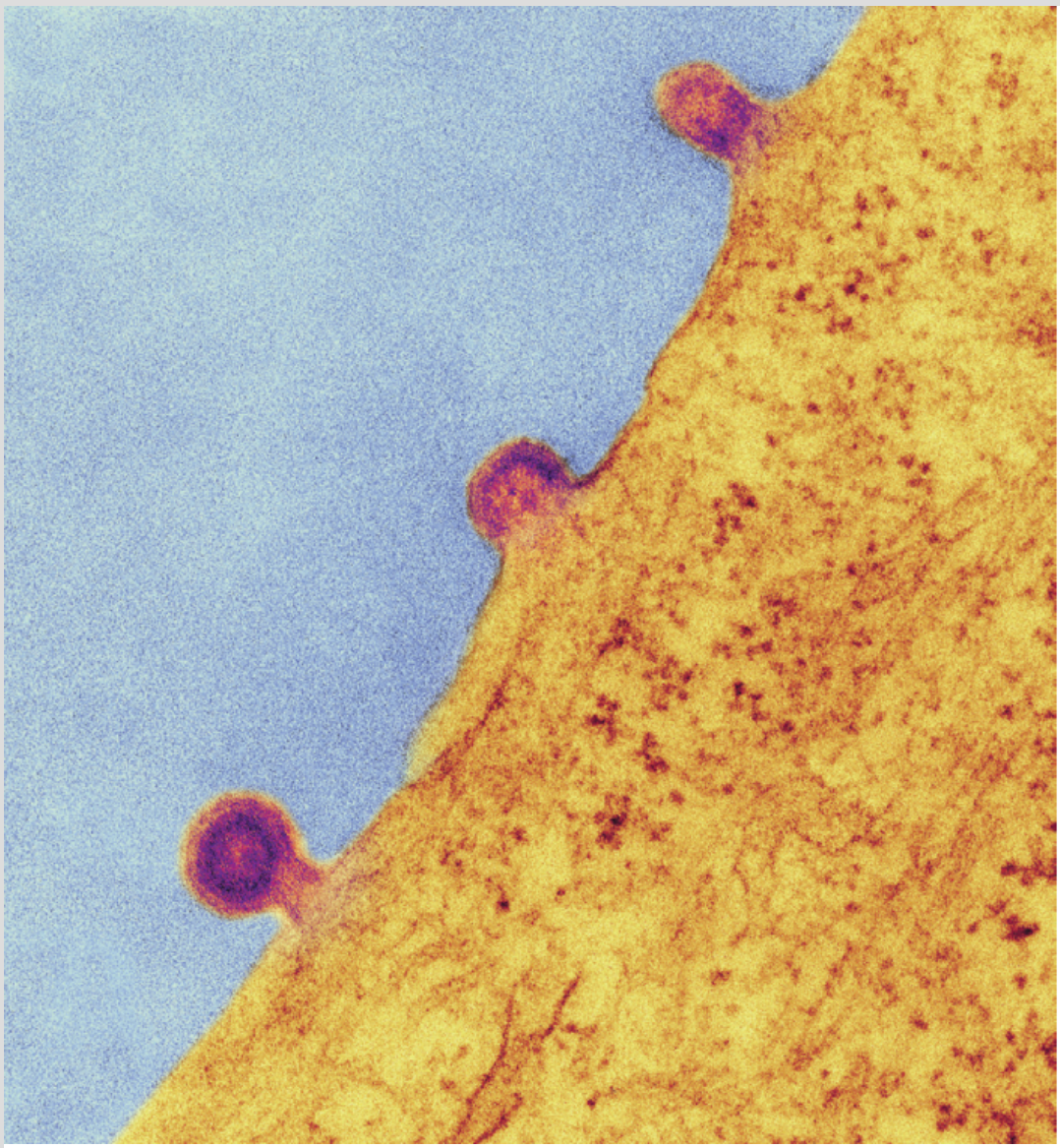
FIGURE 5-41 Lytic replication cycle of a membrane enveloped RNA coronavirus. Coronavirus has a genome of ~30 kb of single stranded RNA, encoding about 30 viral proteins. Within the virion, the RNA genome is in a complex with the viral nucleocapsid (N) protein forming a helical nucleocapsid. The nucleocapsid is packaged within the viral membrane whose major constituents are the membrane (M) and spike (S) proteins. Infection begins with adsorbtion of the virus to a target cell by binding of the S protein to a specific cell surface receptor (step 1), leading to engulfment into an endosome (step 2). The S protein can catalyze fusion of membranes after it has been activated by a proteolytic cleavage that takes place within the acidic interior of the endosome. S protein catalyzed fusion of the viral membrane with the membrane of the endosome enables the nucleocapsid to enter the cytoplasm (step 3). Host cell ribosomes in the cytoplasm recognize the viral genomic RNA as an mRNA and translate early sequences encoding the viral Replicase (step 4). The Replicase produces complementary copies of the genome that are acted upon by the Replicase to produce viral mRNAs for expression of the viral structural proteins as well as many copies of the full length viral RNA genome (step 5).

Translated viral N protein (step 6) assembles with newly replicated viral genomic RNA to form nucleocapsid complexes (step 7). At the same time, the viral membrane proteins M and S are translated and assembled in the membrane of the ER (step 8). In membranes derived from the ER, these viral membrane proteins assemble into a precursor of the viral membrane (step 9), which then envelopes the nucleocapsid as the final stage of viral assembly (step 10). Mature virions formed by budding into the interior of the ER are secreted from the cell by exocytosis (step 11), ready to carry out another round of cellular infection and growth.

Coronavirus is a plus stranded RNA virus, meaning the single stranded RNA genome can immediately engage host-cell ribosomes in the cytoplasm for translation of early viral gene products. The early genes of coronavirus encode a protein complex known as Replicase, which is responsible for replication of the viral RNA genome. Replicase, which contains an RNA template dependent RNA polymerase, first uses the RNA genome as a template to produce complementary RNA strands. These complementary strands, in turn, are used as templates for the production of more copies of the RNA genome as well as of viral mRNAs suitable for translation of the viral structural proteins. The translated viral nucleocapsid protein assembles with viral RNA to form mature helical nucleocapsid structures.

Meanwhile, coronavirus membrane structural proteins are synthesized on the ER membrane (see Chapter 13) and cluster in regions of ER membrane that are precursors of the viral membrane. Complete virions are assembled by envelopment of nucleocapsid by viral membrane as the virus buds into the interior of the ER. Finally,

the mature virions are released from the cell by exocytosis (see Chapter 14). Some membrane enveloped viruses form by a similar envelopment process but are released from the cell by budding directly from the plasma membrane, as shown in an electron micrographs of a cell infected with measles virus in Figure 5-42.



Thomas Deerinck, NCMIR/Science Source.

EXPERIMENTAL FIGURE 5-42 Release of progeny virions by budding. Membrane enveloped viruses are released from infected cells by budding from the plasma

membrane. In this transmission electron micrograph of a cell infected with measles virus, virion buds are clearly visible protruding from the cell surface. Measles virus is an enveloped RNA virus with a helical nucleocapsid.

We have seen that for RNA viruses, replication of the RNA genome and the production and translation of viral mRNAs can take place entirely within the cytoplasm. For DNA viruses, the DNA genome must enter the nucleus in order to be transcribed into viral mRNA and to be replicated by nuclear DNA polymerases. Soon after entry into a host cell, the viral DNA genome is transported into the cell nucleus through nuclear pores, as described in Chapter 13. Once inside the nucleus, the viral DNA is transcribed into RNA by the host's transcription machinery. Processing of the viral RNA primary transcript by host-cell enzymes yields viral mRNA, which is transported to the cytoplasm and translated into viral proteins. The early viral proteins are then transported back into the nucleus, where they collaborate with the cellular replication machinery to replicate viral DNA. Viral proteins that comprise the capsid are usually expressed later in the infection process. Association of the capsid proteins with the newly replicated viral DNA occurs in the nucleus, yielding mature progeny virions.

After the synthesis of hundreds to many thousands of new virions has been completed, depending on the type of virus and host cell, the host cells lyse, releasing all the virions at once, or the host cell releases the virions gradually as the cell progressively disintegrates.

Viral DNA Is Integrated into the Host-Cell Genome in Some Nonlytic Viral Growth Cycles

The genomes of a number of animal viruses can integrate into the host-cell genome. Among the most important of these are the **retroviruses**, which are enveloped viruses with a genome consisting of two identical strands of RNA. These viruses are so named because their RNA genome acts as a template for the formation of a DNA molecule — a flow of genetic information that is opposite to the more common transcription of DNA into RNA. In the retroviral life cycle (Figure 5-43), a viral enzyme called **reverse transcriptase** initially copies the viral RNA genome into single-stranded DNA that is complementary to the viral RNA; the same enzyme then catalyzes the synthesis of a complementary DNA strand. (This complex reaction is detailed in Chapter 7, where we consider closely related intracellular parasites called *retrotransposons*.) The resulting double-stranded DNA is integrated into the chromosomal DNA of the infected cell by an integrase enzyme in the virion. Finally, the integrated DNA, called a *provirus*, is transcribed by the host cell's own RNA polymerase into RNA, which is either translated into viral proteins or packaged within virion capsid proteins to form progeny virions that are released by budding from the host-cell membrane. Because most retroviruses do not kill their host cells, infected cells can replicate, producing daughter cells with integrated proviral

DNA. These daughter cells continue to transcribe the proviral DNA and bud progeny virions.

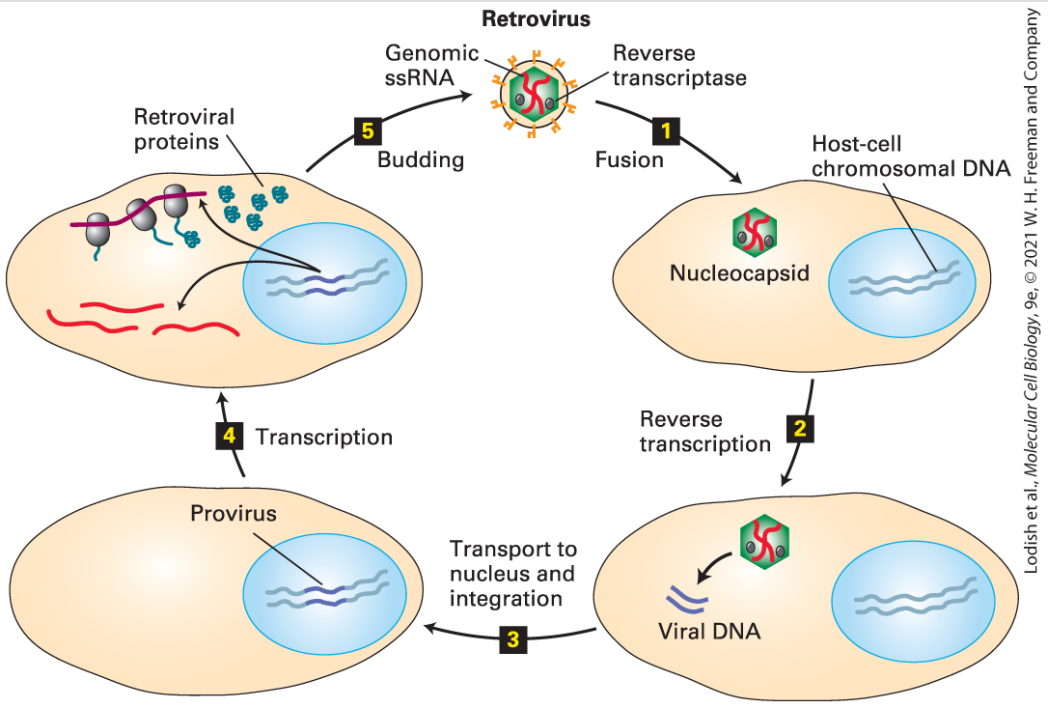


FIGURE 5-43 Retroviral life cycle. Retroviruses have a genome of two identical copies of single-stranded RNA and an outer envelope. Step 1: After viral glycoproteins in the retroviral envelope interact with a specific host-cell membrane protein, the envelope fuses directly with the plasma membrane, allowing entry of the nucleocapsid into the cytoplasm of the cell. Step 2: Viral reverse transcriptase and other proteins copy the viral ssRNA genome into a double-stranded DNA. Step 3: The viral dsDNA is transported into the nucleus and integrated into one of many possible sites in the host-cell chromosomal DNA. For simplicity, only one host-cell chromosome is depicted. Step 4: The integrated viral DNA (provirus) is transcribed by the host-cell RNA polymerase, generating viral mRNAs (dark red) and viral genomic RNA molecules (bright red). The host-cell machinery translates the viral mRNAs into glycoproteins and nucleocapsid proteins. Step 5: Progeny virions then assemble and are released by budding, as illustrated in Figure 5-42.

Some retroviruses contain cancer-causing genes (*oncogenes*), and cells infected by such retroviruses are oncogenically transformed into tumor cells. Studies of oncogenic retroviruses (mostly viruses of birds and mice) have revealed a great deal about the processes that lead to transformation of a normal cell into a cancer cell (see Chapter 25).

Among the known human retroviruses are human T-cell lymphotrophic virus (HTLV), which causes a form of leukemia, and human immunodeficiency virus (HIV-1), which causes acquired immune deficiency syndrome (AIDS). Both of these viruses can infect only specific cell types, primarily certain cells of the immune system and, in the case of HIV-1, some central nervous system neurons and glial cells. Only those cells have cell-surface receptors that interact with viral envelope proteins. Unlike most other retroviruses, HIV-1 eventually kills its host cells. The eventual death of large numbers of immune-system cells results in the defective immune response characteristic of AIDS.

Retroviruses, such as HIV-1, make a DNA replica of the information in viral RNA. The enzyme responsible for this copying reaction is reverse transcriptase, which functions similarly to DNA polymerase but uses the viral RNA as a template. Reverse transcriptase, encoded as a viral gene, lacks a proofreading activity and is much more error prone than cellular DNA polymerases. In this case, a low fidelity of replication works to the advantage of the virus because frequent

mutations in the viral capsid genes allow the virus to continually alter the antigens on the surface of the virion and thus evade the human immune system.

Some DNA viruses can also integrate into a host-cell chromosome. One example is the human papillomaviruses (HPVs), which most commonly cause warts and other benign skin lesions. These viruses replicate without integrating into a host-cell chromosome. However, the genomes of certain HPV serotypes occasionally integrate into the chromosomal DNA of infected cervical epithelial cells. Unlike retroviruses and temperate bacteriophages, this integration is not carried out by viral proteins, but rather is the consequence of host-cell DNA repair processes. Integration is a dead-end for HPV. The integrated viral DNA cannot replicate and generate progeny virions, but oncogenic (cancer-causing) viral proteins can be expressed from the integrated viral genome, initiating development of cervical cancer. A vaccine for the types of HPV associated with cervical cancer has been developed and can protect against the initial infection by these viruses, and consequently, against development of cervical cancer.

KEY CONCEPTS OF SECTION 5.7

Viruses: Parasites of the Cellular Genetic System

- Viruses are small packages of genetic information that can move from one cell to another and that depend on the host cell for their gene expression and replication.

Viral genomes may be either DNA (DNA viruses) or RNA (RNA viruses) and may be either single or double stranded.

- The capsid, which surrounds the viral genome, is composed of multiple copies of one or a small number of virus-encoded proteins. Some viruses also have an outer envelope, which is similar to the plasma membrane of the host cell but contains viral transmembrane proteins.
- Most animal and plant DNA viruses require host-cell nuclear enzymes to carry out transcription of the viral genome into mRNA and production of progeny genomes. In contrast, most RNA viruses encode enzymes that can transcribe the RNA genome into viral mRNA and produce new copies of the RNA genome.
- Host-cell ribosomes, tRNAs, and translation factors are used in the synthesis of all viral proteins in infected cells.
- Lytic viral infection entails adsorption, penetration, synthesis of viral proteins and progeny genomes (replication), assembly of progeny virions, and release of hundreds to thousands of virions, leading to death of the host cell (see Figure 5-41). Release of enveloped viruses occurs by budding through the host-cell plasma membrane (see Figure 5-42).
- Retroviruses are enveloped animal viruses containing a single-stranded RNA genome. After a host cell is penetrated, reverse transcriptase, a viral enzyme carried in the virion, converts the viral RNA genome into double-stranded DNA, which is integrated into chromosomal DNA by an integrase enzyme that enters the cell inside the virion (see Figure 5-43).
- Unlike infection by other retroviruses, HIV-1 infection eventually kills host cells, causing the defects in the immune response characteristic of AIDS.
- Tumor viruses, which contain oncogenes, may have an RNA genome (e.g., human T-cell lymphotrophic virus) or a DNA genome (e.g., human papillomaviruses). Integration of the genomes of these viruses into a host-cell chromosome can cause transformation of the host cell into a tumor cell.

End of Chapter

Achieve

Visit Achieve to access study tools and learn more about the content of this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools including videos, animations, and quizzes

Key Terms

alternative splicing
aminoacyl-tRNA
anticodon
base pairs
capsid
codon
complementary
denaturation (melting)
deoxyribonucleic acid (DNA)
DNA polymerase
double helix
elongation factors (EFs)
eukaryotic translation initiation factors (eIFs)
excision-repair systems

gene conversion
genetic code
GTPase superfamily
helicase
Holliday structure
homologous recombination
hybridization
isoform
lagging strand
leading strand
messenger RNA (mRNA)
mutation
nonhomologous end joining
nucleocapsid
Okazaki fragments
open reading frame
phosphodiester bond
point mutation
polymer
polyribosome
precursor mRNA (pre-mRNA)
primase
primer
promoter
purine
pyrimidine
reading frame
recombination
release factors (RFs)
replication fork

retrovirus
reverse transcriptase
ribosomal RNA (rRNA)
ribosome
RNA polymerase
RNA splicing
transcription
transfer RNA (tRNA)
translation
viral envelope
virion
virus
xeroderma pigmentosum

Review the Concepts

1. What are Watson-Crick base pairs? Why are they important?
2. Preparing plasmid DNA (double stranded, circular) for sequencing involves annealing a complementary, short, single-stranded oligonucleotide DNA primer to one strand of the plasmid template. This is routinely accomplished by heating the plasmid DNA and primer to 90° C and then slowly bringing the temperature down to 25° C. Why does this protocol work?
3. What difference between RNA and DNA helps to explain the greater stability of DNA? What implications does this have for the function of DNA?
4. What are the major differences in the synthesis and structure of prokaryotic and eukaryotic mRNAs?

5. While investigating the function of a specific growth factor receptor gene from humans, researchers found that two types of proteins are synthesized from this gene. A larger protein containing a membrane-spanning domain recognizes growth factors at the cell surface, stimulating a specific downstream signaling pathway. In contrast, a related, smaller protein is secreted from the cell and binds available growth factor circulating in the blood, thus inhibiting the downstream signaling pathway. Speculate on how the cell synthesizes these disparate proteins.
6. How would a mutation in the poly(A)-binding protein gene affect translation? How would an electron micrograph of polyribosomes from such a mutant differ from the normal pattern?
7. What characteristic of DNA results in the requirement that some DNA synthesis be discontinuous? How are Okazaki fragments and DNA ligase used by the cell?
8. Eukaryotes have repair systems that prevent mutations due to copying errors and exposure to mutagens. What are the three excision-repair systems found in eukaryotes, and which one is responsible for correcting thymine-thymine dimers that form as a result of UV light damage to DNA?
9. DNA-repair systems are responsible for maintaining genomic fidelity in normal cells despite the high frequency with which mutational events occur. What type of DNA mutation is generated by (a) UV radiation and (b) ionizing radiation? Describe the system responsible for repairing each of these types of mutations in mammalian cells. Postulate why a loss of function in one or more DNA repair systems is present in many cancers.

10. What is the name given to the process that can repair DNA damage *and* generate genetic diversity? Briefly describe the similarities and differences of the two processes.
11. a. Which of the following DNA strands, the top or bottom, would serve as a template for RNA transcription if the DNA molecule were to unwind in the indicated direction?

    5′ ACGGACTGTACCGCTGAAGTCATGGACGCTCGA 3′

    3′ TGCCTGACATGGCGACTTCAGTACCTGCGAGCT 5′

    $\longrightarrow$

    Direction of DNA unwinding

    b. What would be the resulting RNA sequence (written $5' \rightarrow 3'$)?
12. Contrast prokaryotic and eukaryotic gene characteristics.
13. You have learned about the events surrounding DNA replication and the central dogma. Identify the steps associated with these processes that would be adversely affected in the following scenarios.
    a. Helicases unwind the DNA, but stabilizing proteins are mutated and cannot bind to the DNA.
    b. The mRNA molecule forms a hairpin loop on itself via complementary base pairing in an area spanning the AUG start site.
    c. The cell is unable to produce functional $\mathrm{tRNA_i^{Met}}$.
14. Use the accompanying key to determine the amino acid sequence of the polypeptide produced from the following DNA sequence. Intron sequences are highlighted. Note: Not all amino acids in the key will be used.

    5′ TTCTAAACGCATGAAGCACCGTCTCAGAGCCAGTGA 3′

    3′ AAGATTTGCGTACTTCGTGGCAGAGTCTCGGTCACT 5′

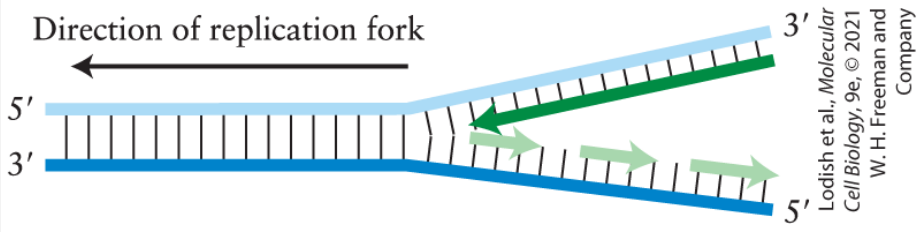
$\longrightarrow$

Direction of DNA unwinding

Asn = AAU Cys = TCG Gly = CAG His = CAU Lys = AAG

Met = AUG Phe = UUC Ser = AGC Tyr = UAC Val = GUC; GUA

15.



a. Look at the figure above. Explain why it is necessary for Okazaki fragments to be formed as the lagging strand is produced (instead of a continuous strand).
b. If the DNA polymerase in the figure above could bind only to the lower template strand, under what condition(s) would it be able to produce a leading strand?

16. The DNA repair systems preferentially target the newly synthesized strand. Why is this important?
17. Identify the specific types of point mutations below (you are viewing the direct DNA version of the RNA sequence).

    Original sequence: 5′ AUG TCA GGA CGT CAC TCA GCT 3′

    Mutation A: 5′ AUG TCA GGA CGT CAC TGA GCT 3′

    Mutation B: 5′ AUA TCA GGA CGT CAC TCA GCT 3′
18. The genome of a retrovirus can integrate into the host-cell genome. A number of retroviruses can infect certain types of human cells.
    a. What gene is unique to retroviruses, and why is the protein encoded by this gene absolutely necessary for maintaining the retroviral life cycle?

b. List two retroviruses that infect specific cell types, briefly describe the medical implications resulting from these infections, and describe why only certain cells are infected.

19. a. Detail the key differences between lytic and nonlytic viral infection and provide an example of each.

    b. Which of the following processes occurs in both lytic and nonlytic viral infections?

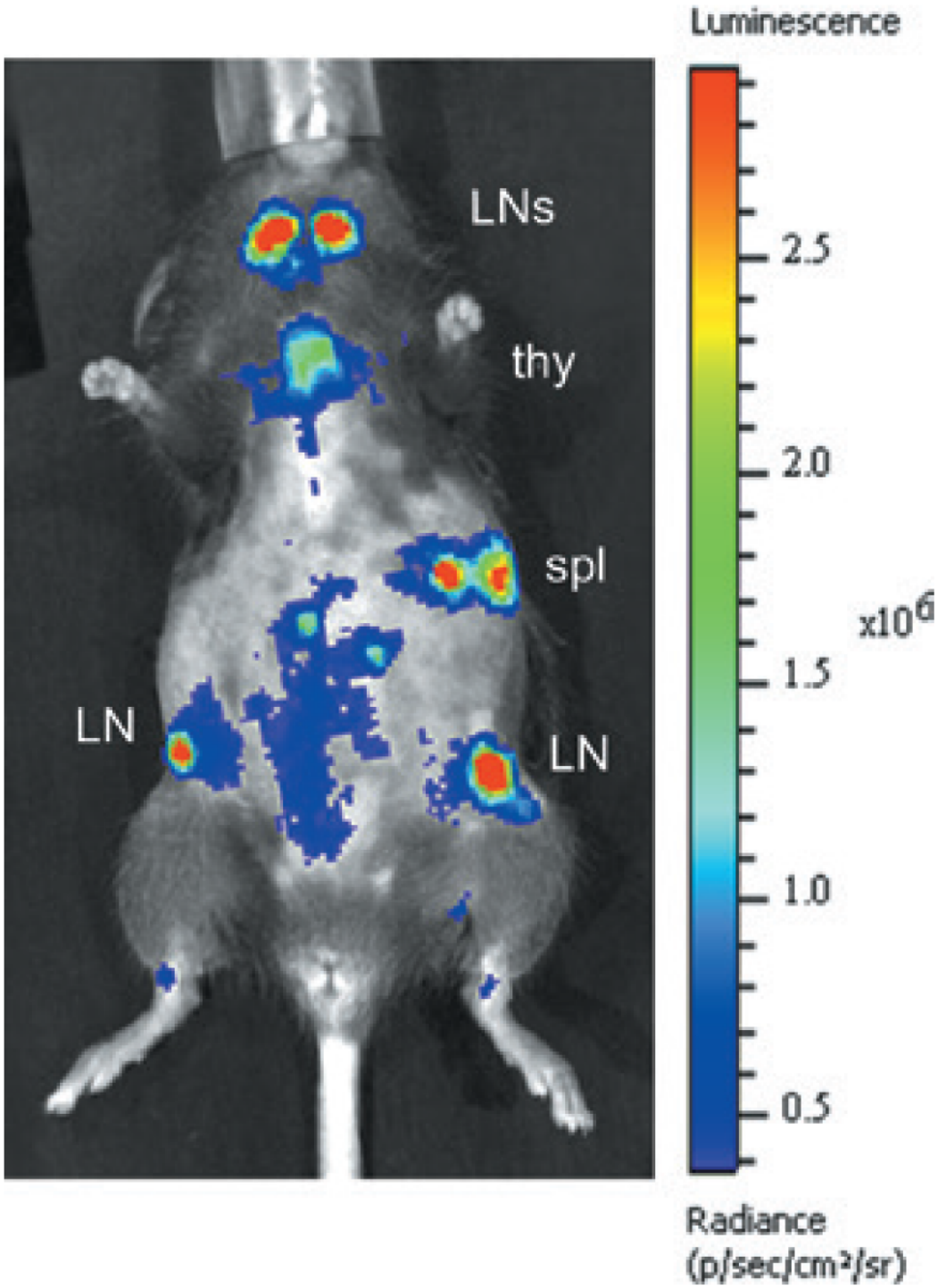
        i. Infected cell ruptures to release viral particles.

        ii. Viral mRNAs are transcribed by the host-cell translation machinery.

        iii. Viral proteins and nucleic acids are packaged to produce virions.

CHAPTER 6

Molecular Genetic Techniques



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman

Genetic modifications make it possible to locate a specific cell type within the body of a living mouse. Scientists can label cells of the immune system by generating a transgenic mouse carrying the gene for insect luciferase expressed from the CD2 promoter, which is specifically expressed in T cells. After injection of the bioluminescent substrate for luciferase,

the light emitted through the skin is imaged for 30 seconds. T cells can be seen located in the lymph nodes (LNs), the thymus (thy), and the spleen (spl).

[From J. W. Kleinovink et al., 2019. A dual-color bioluminescence reporter mouse for simultaneous in vivo imaging of T cell localization and function. *Frontiers in Immunology* **9**:3097. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6333049]

OUTLINE

6.1 Using Genetic Analysis of Mutations to Identify and Study Genes
6.2 DNA Cloning and Characterization
6.3 Using Sequence Information to Identify Genes and Deduce Their Function
6.4 Locating and Identifying Genes That Specify Human Traits
6.5 Using Cloned DNA Fragments to Study Gene Expression
6.6 Altering the Function of Specific Genes by Design

In the field of molecular cell biology, we seek to understand the biological behavior of cells in terms of the underlying chemical and molecular mechanisms. Most often, cell biologists investigating some molecular process focus on the function of particular proteins or sets of proteins. They set out to answer a set of fundamental questions: What is the function of the protein in the context of a living cell? What is the biochemical function of the purified protein? Where is the protein located within the cell? When and where is the protein expressed? How is it related to other proteins? And how did the function of the protein arise during evolution? The answers to these questions almost always come from the study and manipulation of the gene that encodes the particular protein. In the first part of this chapter, we consider strategies for identifying and isolating genes of interest. We then consider strategies for

determining the cellular function of the gene product (Figure 6-1). These strategies rely on experimental methods for manipulating genes.

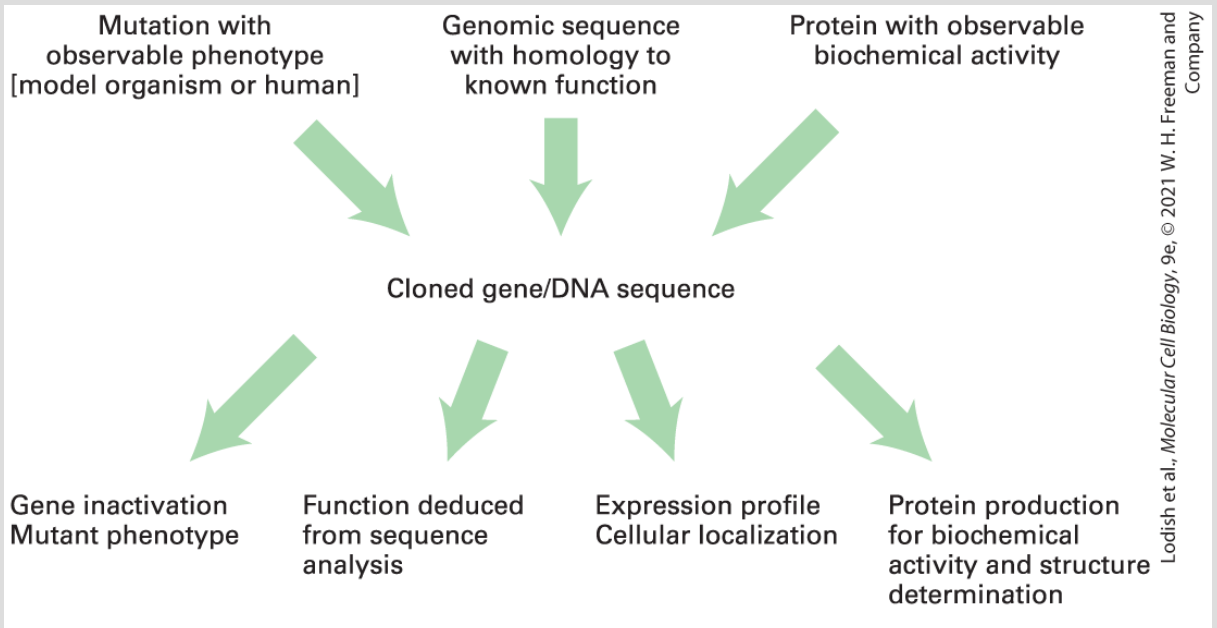


FIGURE 6-1 Genetic and genomic analysis is central to understanding cell function at the molecular level. The general strategy for understanding the function of the constituent parts of a cell is to first identify and isolate a gene. This process begins with a mutant organism, a purified protein, or the identification of a protein-coding sequence by analysis of genomic sequence databases. The actual gene is isolated either from a DNA library of clones or by amplifying the specific gene sequence from genomic DNA. The gene sequence can be used to identify other similar gene sequences in genomic databases, providing insight into the general function of the gene and how it evolved. Once a cloned gene is isolated, it can be manipulated to produce large quantities of the protein it encodes for biochemical experiments and to design probes for studies of where and when the protein is expressed in an organism. Alternatively, a cloned gene can be inactivated by one of various techniques and used to generate mutant cells or organisms. By examining the phenotypic consequences of mutations that inactivate a particular gene, geneticists are able to connect knowledge about the sequence, structure, and biochemical activity of the encoded protein to its function in the context of a living cell or multicellular organism.

Classically, proteins with important cellular functions were discovered either by purifying a protein with a relevant biochemical activity or by isolating an appropriate mutant. An excellent example showing how these two approaches can mutually reinforce one another was the discovery of NSF, a key protein required for vesicle trafficking in the secretory pathway, described in Chapter 14. The mammalian NSF protein was purified using a biochemical assay for fusion of transport vesicles with their target membrane. Meanwhile, a mutation in the yeast gene known as *sec18* was identified in a genetic screen for mutants with a defect in protein secretion. Eventually the genetic and biochemical approaches were shown to have converged on the same gene when the yeast gene defective in the *sec18* mutant was shown to be highly similar in sequence to the mammalian gene that encodes the NSF protein.

Today the complete genome sequences of the human and all of the major experimental organisms are known, and the pace at which genes can be identified and studied has accelerated enormously. Because gene sequences and the sequences of the encoded proteins can be obtained from genomic sequences, almost the entire repertoire of encoded proteins is already known for most experimental organisms. Thus gene sequences of interest can now be readily and automatically identified from the chromosomal location of a mutation or just a short segment of the amino acid sequence of a purified protein. Moreover, genes worthy of study can be identified entirely from sequence information by comparing the

genome sequences from different organisms. Using this approach, investigators can study the function of a human gene that is implicated in cancer, for example, by manipulating the corresponding gene of similar sequence in the mouse or even in the fruit fly or in yeast if a similar gene is present.

6.1 Using Genetic Analysis of Mutations to Identify and Study Genes

As described in Chapter 5, the information encoded in the DNA sequence of genes specifies the sequence — and therefore the structure and function — of every protein molecule in a cell. The power of genetics as a tool for studying cells and organisms lies in the ability of researchers to make a change in some gene and then evaluate the effect on a living organism. Geneticists expose organisms to an agent that induces mutations, then search for mutant organisms that are defective in a process of interest. Genetic analyses of mutants defective in a particular process can reveal (1) genes required for the process to occur, (2) the order in which gene products act in the process, and (3) whether and how the proteins encoded by different genes interact with one another. Before we look more closely at genetic studies of this type, let's first review some basic genetic terms used throughout our discussion.

The different forms or variants of a gene are referred to as **alleles**. Geneticists commonly refer to the numerous naturally occurring genetic variants that exist in populations, particularly human populations, as alleles. The term **mutation** is usually reserved for a recent change in an allele, such as after treatment of an

experimental organism with a **mutagen**, an agent that causes a heritable change in the DNA sequence.

Strictly speaking, the particular set of alleles for all the genes carried by an individual constitutes its **genotype**. However, this term is most often used in a more restricted sense to denote the alleles of a particular gene or genes under examination. For experimental organisms, the term **wild type** is often used to designate a standard genotype for use as a reference in breeding experiments. Thus the normal, nonmutant allele is usually designated as the wild type. Because of the enormous allelic variation that naturally exists in human populations, the term *wild type* usually denotes an allele that is present at a much higher frequency than any of the other possible alternatives.

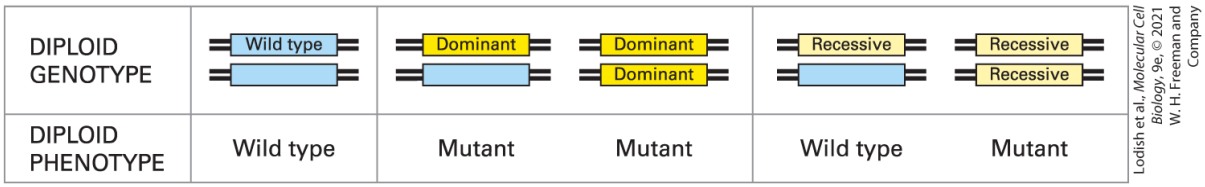
Geneticists draw an important distinction between the *genotype* and the *phenotype* of an organism. The term **phenotype** refers to all the physical attributes or traits of an individual that are the consequence of a given genotype. In practice, however, the term *phenotype* is usually used to denote the consequences that result from the particular alleles that are under experimental study. Readily observable phenotypic characteristics are critical in the genetic analysis of mutations.

Recessive and Dominant Mutant Alleles Generally Have Opposite Effects on Gene Function

A fundamental genetic difference between experimental organisms is whether their cells carry two copies of each chromosome or only one copy of each chromosome. The former are referred to as *diploid*; the latter as *haploid*. Most complex multicellular organisms (e.g., fruit flies, mice, humans) are diploid, whereas many simple, unicellular organisms are haploid. Some organisms, notably the yeast *Saccharomyces cerevisiae*, can exist in either haploid or diploid states. The normal cells of some organisms, both plants and animals, carry more than two copies of each chromosome and are thus designated *polyploid*. Moreover, cancer cells begin as diploid cells, but through the process of transformation into cancer cells can gain extra copies of one or more chromosomes and are thus designated as *aneuploid*. Our discussion of genetic techniques and analysis relates to diploid organisms, including diploid yeasts.

Although many different alleles of a gene might be present in a population, any individual diploid organism will carry two copies of each gene and thus at most can have two different alleles. A diploid individual with two different alleles is **heterozygous** for a gene, whereas a diploid individual that carries two identical alleles is **homozygous** for a gene. A mutant allele is **recessive** if both alleles must be mutant in order for the mutant phenotype to be observed;

that is, the individual must be homozygous for the mutant allele to show the mutant phenotype. In contrast, a mutant allele is **dominant** if the mutant phenotype can be observed in a heterozygous individual carrying one mutant and one wild-type allele (Figure 6-2).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 6-2 Effects of dominant and recessive mutant alleles on phenotype in diploid organisms. A single copy of a dominant mutant allele is sufficient to produce a mutant phenotype, whereas both copies of a recessive mutant allele must be present to cause a mutant phenotype. Recessive mutations usually cause a loss of function; dominant mutations usually cause a gain of function or an altered function.

Whether a mutant allele is recessive or dominant provides valuable information about the function of the affected gene and the nature of the causative mutation. Recessive alleles usually result from a mutation that inactivates the affected gene, leading to a partial or complete *loss of function*. Such recessive mutations may remove part of the gene or the entire gene from the chromosome, disrupt expression of the gene, or alter the structure of the encoded protein, thereby altering its function. Conversely, dominant alleles are often the consequence of a mutation that causes some kind of *gain of function*. Such dominant mutations may increase the activity of the

encoded protein, confer a new function on it, or lead to a new spatial or temporal pattern of expression.

In some rare cases, dominant mutations are associated with a loss of function. For instance, some genes are *haploinsufficient*, in that removing or inactivating one of the two alleles of such a gene leads to a mutant phenotype because not enough gene product is made. In other rare instances, a dominant mutation in one allele may lead to a structural change in the protein that interferes with the function of the wild-type protein encoded by the other allele. This type of mutation, referred to as a *dominant-negative*, produces a phenotype similar to that obtained from a loss-of-function mutation.

Some alleles can exhibit both recessive and dominant properties. In such cases, statements about whether an allele is dominant or recessive must specify the phenotype. For example, the allele of the hemoglobin gene that produces sickle cell disease, designated Hb^s, has more than one phenotypic consequence. Individuals who are homozygous for this allele (Hb^s/Hb^s) have debilitating anemia caused by sickle-cell disease, but heterozygous individuals (Hb^s/Hb^a) do not have the disease. Therefore, Hb^s is *recessive* for the trait of sickle-cell disease. On the other hand, heterozygous (Hb^s/Hb^a) individuals are more resistant to malaria than are homozygous (Hb^a/Hb^a) individuals, revealing that Hb^s is *dominant* for the trait of malaria resistance. ■

As described in Chapter 5, spontaneous mutations continually arise as a result of chemical and radiation damage to DNA and of errors in DNA replication. In genetic screens of experimental organisms, it is usually desirable to increase the frequency of mutations by controlled application of a chemical mutagen. A mutagen commonly used in experimental organisms is ethylmethane sulfonate (EMS), which can chemically modify guanine bases in DNA, ultimately leading to the conversion of a G·C base pair into an A·T base pair. The alteration of only a single base pair in the sequence of a gene is known as a **point mutation**. As will be described in Section 6.3, a point mutation can cause either a loss of function or a gain of function in a gene depending on the type of mutation and where it occurs in a gene sequence.

Segregation of Mutations in Breeding Experiments Reveals Whether They Are Dominant or Recessive

To see how Mendelian genetics is used to test whether an allele is dominant or recessive, we must first review the type of cell division that gives rise to gametes (sperm and egg cells in animals). Whereas the body (somatic) cells of most multicellular organisms divide by mitosis, the germ cells that give rise to gametes undergo meiosis. Like somatic cells, before meiosis germ cells are diploid, containing two homologs of each morphological type of chromosome. The two

homologs are descended from different parents, and thus their genes may exist in different allelic forms. Figure 6-3 depicts the major events in mitotic and meiotic cell division. In mitosis, DNA replication is always followed by cell division, yielding two diploid daughter cells. In meiosis, *one* round of DNA replication is followed by *two* separate cell divisions, yielding haploid (1*n*) cells known as gametes that contain only one chromosome of each homologous pair. The apportionment, or **segregation**, of the replicated homologous chromosomes to daughter cells during the first meiotic division is random, and different chromosomes segregate independently of one another, yielding gametes with different mixes of paternal and maternal chromosomes.

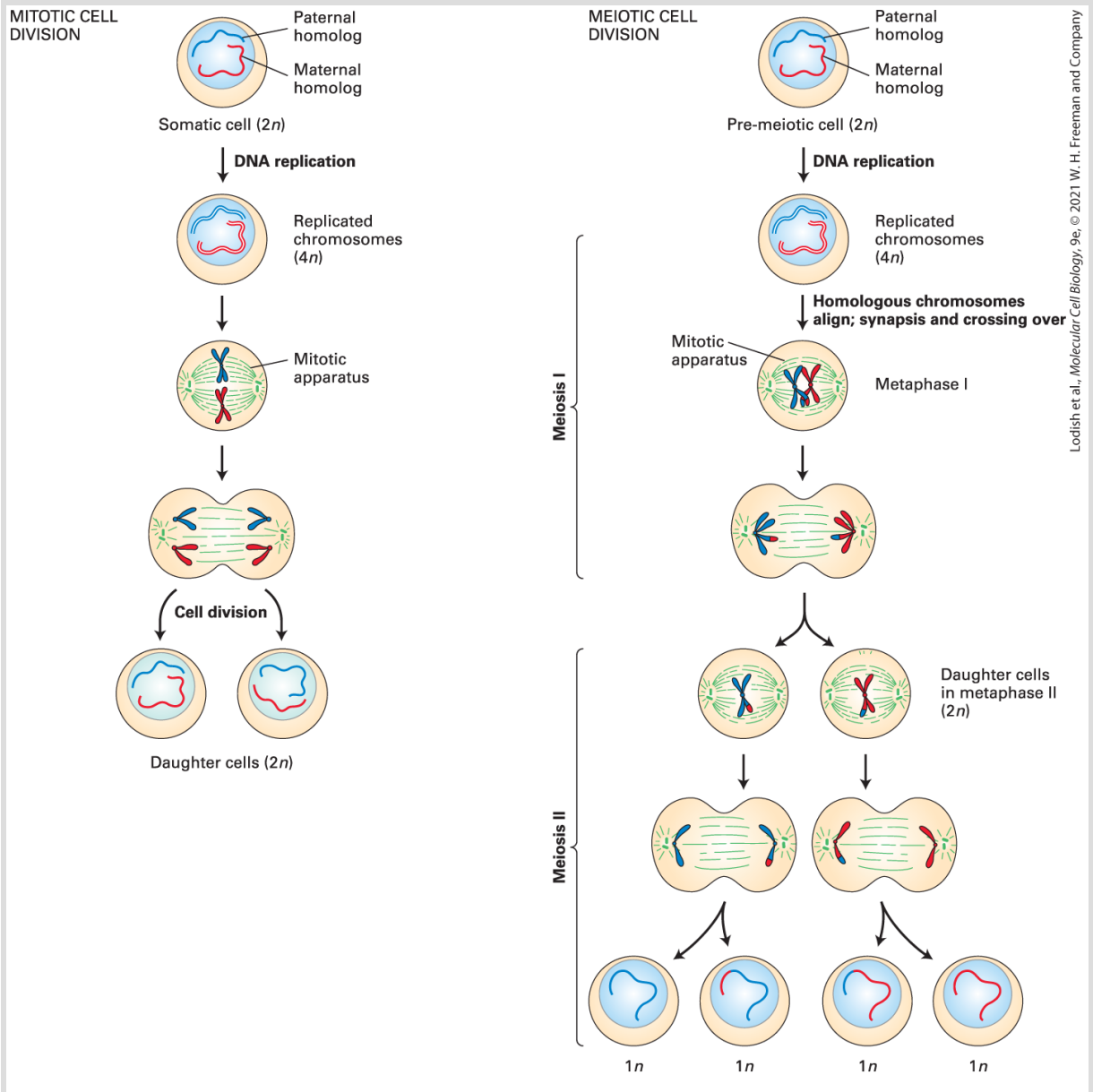
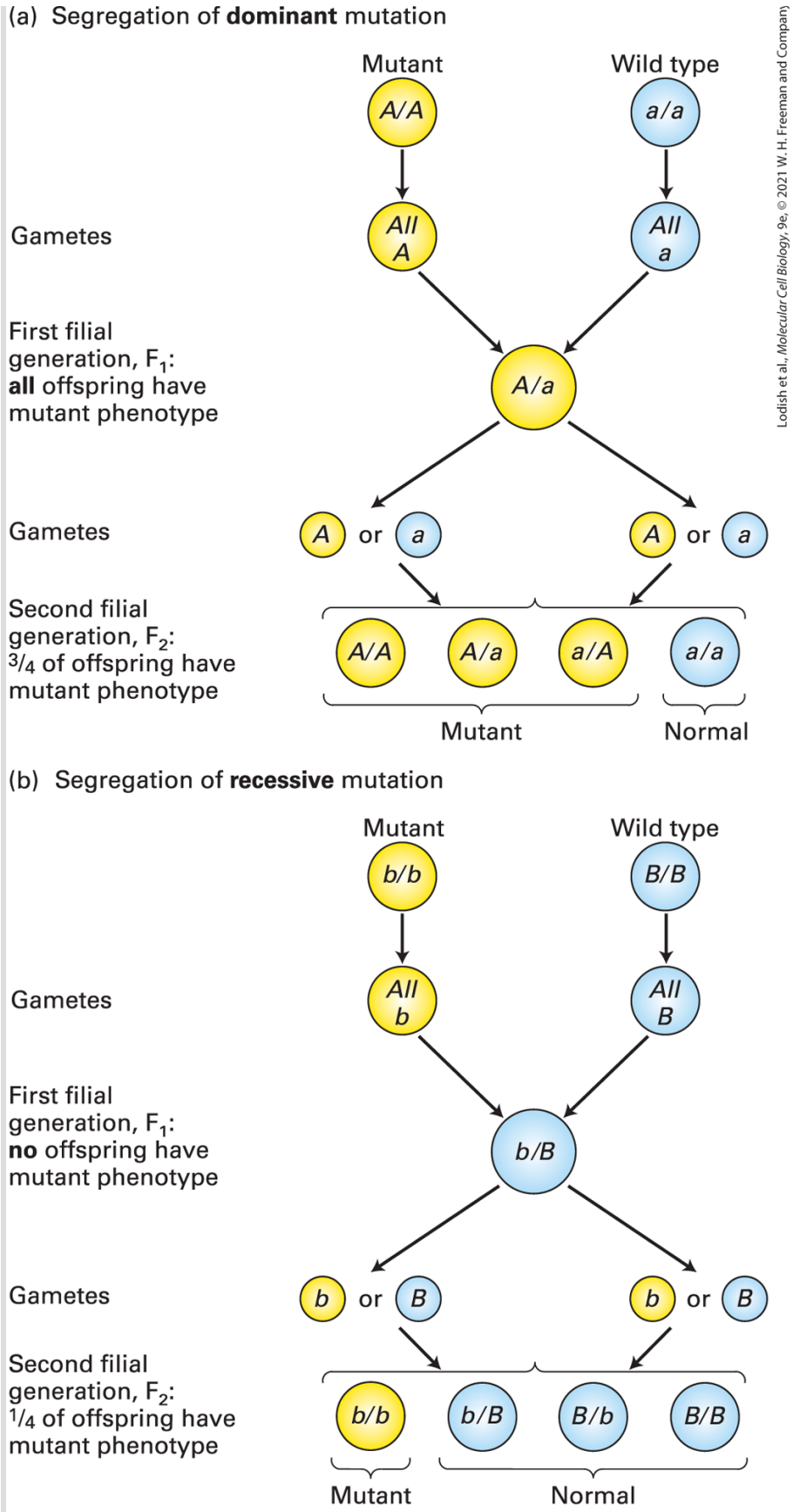


FIGURE 6-3 Comparison of mitosis and meiosis. Both somatic cells and pre-meiotic germ cells have two copies of each chromosome (2*n*), one inherited from the mother and one from the father. In mitosis, the replicated chromosomes, each composed of two sister chromatids, align at the cell center in such a way that both daughter cells receive a maternal and a paternal homolog of each morphological type of chromosome. During the first *meiotic* division, however, each replicated chromosome pairs with its homologous partner at the cell center; this pairing off is referred to as *synapsis*, and crossing over between homologous chromosomes is evident at this stage. One replicated chromosome of each morphological type then goes into each daughter cell. The resulting cells undergo a second division without intervening DNA replication, so that one of the sister chromatids of each morphological type

is apportioned to the daughter cells. In the second meiotic division, the alignment of chromatids and their equal segregation into daughter cells is the same as in mitotic division. The alignment of pairs of homologous chromosomes in metaphase I is random with respect to other chromosome pairs, resulting in a mix of paternally and maternally derived chromosomes in each daughter cell.

The key to determining whether an allele is dominant or recessive is to construct a heterozygote by breeding experiments. As a way to avoid unwanted complications in this process, geneticists usually strive to begin with strains of organisms that are homozygous for the genes under examination. In such *true-breeding* strains, every individual will receive the same allele from each parent, and therefore the composition of alleles will not change from one generation to the next. When a true-breeding mutant strain is mated to a true-breeding wild-type strain, all the first filial (F_1) progeny will be heterozygous (Figure 6-4). If the F_1 progeny exhibit the mutant trait, then the mutant allele is dominant; if the F_1 progeny exhibit the wild-type trait, then the mutant allele is recessive. Further crossing between F_1 individuals will also reveal different patterns of inheritance according to whether the mutation is dominant or recessive. When F_1 individuals that are heterozygous for a dominant allele are crossed among themselves, three-fourths of the resulting F_2 progeny will exhibit the mutant trait. In contrast, when F_1 individuals that are heterozygous for a recessive allele are crossed among themselves, only one-fourth of the resulting F_2 progeny will exhibit the mutant trait.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 6-4 Segregation patterns of dominant and recessive mutations in crosses between true-breeding strains of diploid organisms. All the offspring in the first (F_1) generation are heterozygous. If the mutant allele is dominant, the F_1 offspring will exhibit the mutant phenotype, as in part (a). If the mutant allele is recessive, the F_1 offspring will exhibit the wild-type phenotype, as in part (b). Crossing of the F_1 heterozygotes among themselves also produces different segregation ratios for dominant and recessive mutant alleles in the F_2 generation. Diploid organisms with the mutant phenotype are yellow and diploid organisms with a normal phenotype are blue.

As noted earlier, the yeast *S. cerevisiae*, an important experimental organism, can exist in either a haploid or a diploid state. In these unicellular eukaryotes, crosses between haploid cells can determine whether a mutant allele is dominant or recessive. Haploid yeast cells, which carry one copy of each chromosome, can be of two different mating types, known as **a** and α. Haploid cells of opposite mating type can mate to produce **a**/α diploids, which carry two copies of each chromosome. If a new mutation with an observable phenotype is isolated in a haploid strain, the mutant strain can be mated to a wild-type strain of the opposite mating type to produce **a**/α diploids that are heterozygous for the mutant allele. If these diploids exhibit the mutant trait, then the mutant allele is dominant, but if the diploids exhibit the wild-type trait, then the mutant allele is recessive. When **a**/α diploids are placed under starvation conditions, the cells undergo meiosis, each giving rise to a tetrad of four haploid spores, two of type **a** and two of type α. A heterozygous diploid cell yields two spores carrying the mutant allele and two carrying the wild-type allele (Figure 6-5). Under appropriate

conditions, yeast spores will germinate, producing vegetative haploid strains of both mating types.

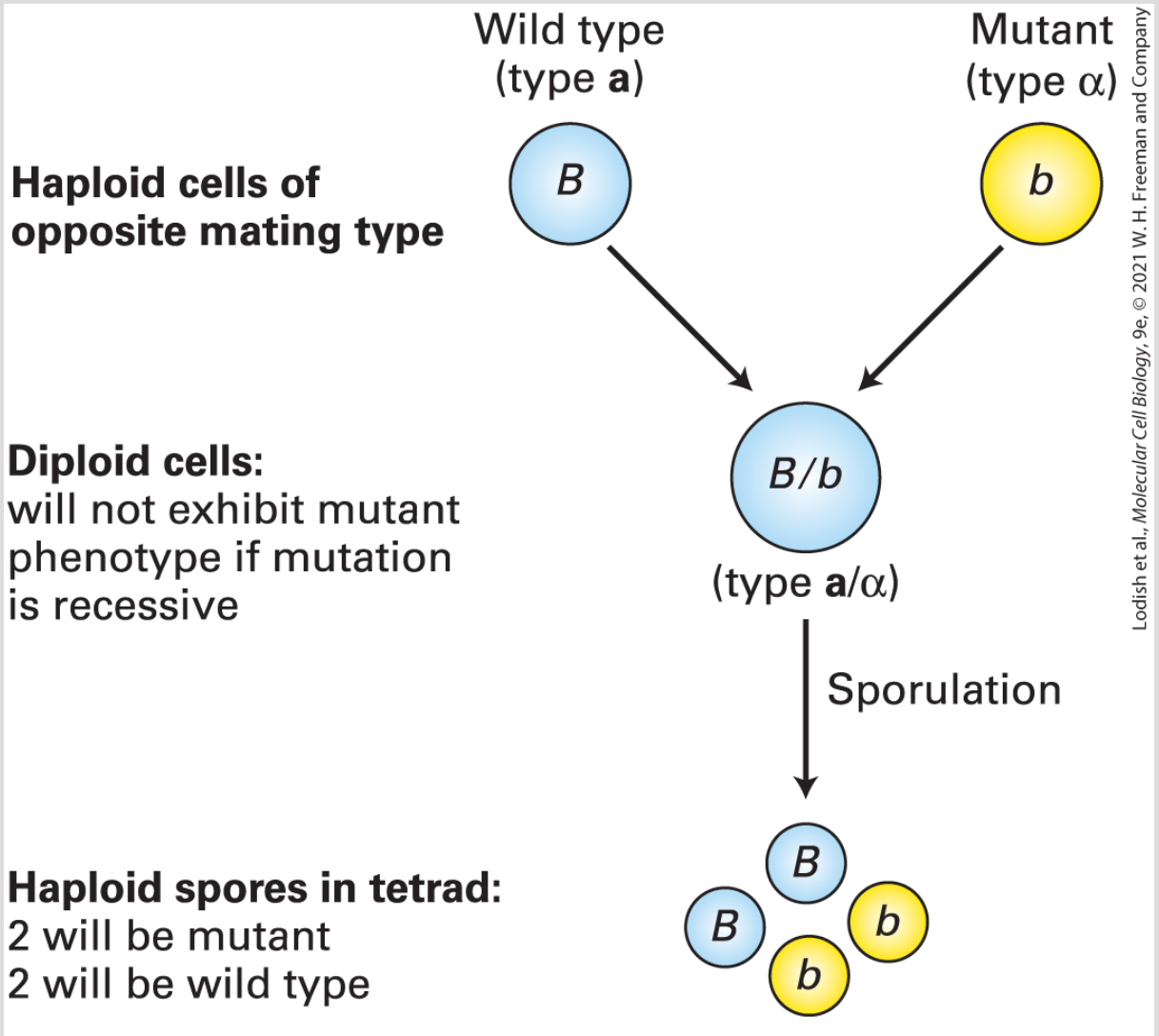


FIGURE 6-5 Segregation of alleles in yeast. Haploid *Saccharomyces* cells of opposite mating type (i.e., one of mating type **a** and one of mating type α) can mate to produce an **a**/α diploid. If one haploid carries a dominant wild-type allele and the other carries a recessive mutant allele of the same gene, the resulting heterozygous diploid will express the dominant trait. Under certain conditions, a diploid cell will form a tetrad of four haploid spores. Two of the spores in the tetrad will express the recessive trait and two will express the dominant trait. Yeast with the mutant trait are shown in yellow and yeast with a normal phenotype are shown in blue.

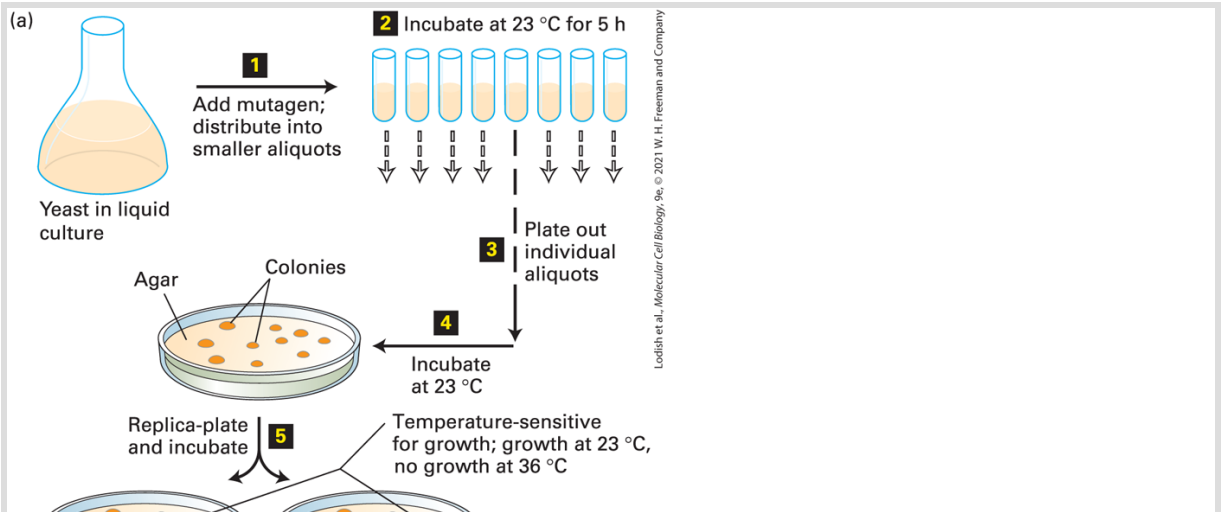
Conditional Mutations Can Be Used to Study Essential Genes in Yeast

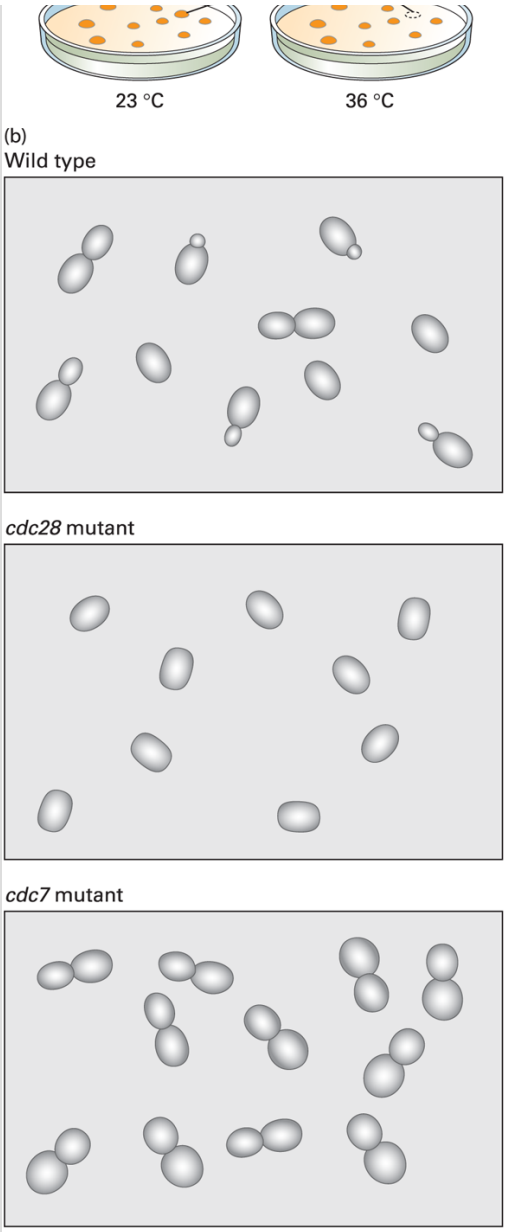
Procedures called *genetic screens* are used to identify and isolate mutants. These procedures depend on whether the experimental organism is haploid or diploid and, if the latter, whether the mutation is recessive or dominant. Genes that encode proteins essential for life are among the most interesting and important ones to study. Since phenotypic expression of a mutation in an essential gene leads to death of the individual, clever genetic screens are needed to isolate and maintain organisms with a lethal mutation.

In haploid yeast cells, essential genes can be studied through the use of *conditional mutations*, so called because the mutant phenotype is exhibited under certain conditions but not others. Among the most common conditional mutations are *temperature-sensitive mutations*, which are useful in organisms, such as bacteria and lower eukaryotes, that can grow at a range of temperatures. For instance, a missense mutation may cause the resulting mutant protein to be less stable at certain temperatures, such that the protein is fully functional at one temperature (e.g., 23 °C) but begins to denature and is thus inactive at another temperature (e.g., 36 °C), whereas the normal protein would be fully stable and functional at both temperatures. A temperature at which the mutant phenotype is observed is called *nonpermissive*; a *permissive* temperature is one at which the mutant phenotype is not observed even though the

mutant allele is present. Thus mutant strains can be maintained at a permissive temperature and then grown at a nonpermissive temperature for analysis of the mutant phenotype.

An example of a particularly important screen for temperature-sensitive mutants in the yeast *S. cerevisiae* comes from the studies of L. H. Hartwell and colleagues in the late 1960s and early 1970s. They set out to identify genes important in regulation of the cell cycle (during which a cell synthesizes proteins, replicates its DNA, and then undergoes mitotic cell division). Exponential growth of a single yeast cell for 20–30 cell divisions forms a visible yeast colony on solid agar medium. Because mutants with a complete block in the cell cycle would not be able to form colonies, conditional mutants were required to study mutations that affect this basic cellular process. To screen for such mutants, the researchers first exposed yeast cells to mutagens and then identified mutant yeast cells that could grow normally at 23 °C, but could not form a colony when placed at 36 °C (Figure 6-6a).





EXPERIMENTAL FIGURE 6-6 Haploid yeast cells carrying temperature-sensitive lethal mutations can be maintained at permissive temperature and analyzed at nonpermissive temperature. (a) Genetic screen for temperature-sensitive cell division cycle (cdc) mutants in *S. cerevisiae*. Yeast cells that grow and form colonies at 23 °C (permissive temperature) but not at 36 °C (nonpermissive temperature) may carry a lethal mutation that blocks cell division. See L. H. Hartwell, 1967, *J. Bacteriol.* **93**:1662. (b) The subset of temperature-sensitive mutants that have a block in the cell cycle can be identified by an arrest at the nonpermissive temperature at a uniform stage of the cell cycle. Shown here are representations of wild-type yeast and two different temperature-sensitive mutants after

incubation at the nonpermissive temperature for 6 hours. Wild-type cells, which continue to grow, can be seen with all different sizes of buds, reflecting different stages of the cell cycle. In contrast, *cdc28* mutants arrest at a point before emergence of a new bud and therefore appear as unbudded cells, while the *cdc7* mutants, which arrest just before separation of the mother cell and bud (emerging daughter cell), appear as cells with large buds.

Once temperature-sensitive mutants were isolated, further analysis revealed that some indeed were defective in cell division. In *S. cerevisiae*, cell division occurs through a budding process, and the size of the bud, which is easily visualized by light microscopy, indicates a cell's position in the cell cycle. Each of the mutants that could not grow at 36 °C was examined by microscopy after several hours at the nonpermissive temperature. Examination of many different temperature-sensitive mutants revealed a set that exhibited a distinct block in the cell cycle. These mutants were therefore designated *cdc (c*ell *d*ivision *c*ycle) *mutants*. Importantly, these yeast mutants did not simply fail to grow, as they might have if they carried a mutation affecting general cellular metabolism. Rather, at the nonpermissive temperature, the mutants grew normally for part of the cell cycle, but then arrested at a particular stage of the cell cycle, so that many cells at that stage were seen (Figure 6-6b). Most cdc mutations in yeast are recessive; that is, when haploid cdc strains are mated to wild-type haploids, the resulting heterozygous diploids are neither temperature sensitive nor defective in cell division.

Recessive Lethal Mutations in Diploids Can Be Identified by Inbreeding and Maintained in Heterozygotes

In diploid organisms, phenotypes resulting from recessive mutations can be observed only in individuals that are homozygous for the mutant alleles. Yet exposure of a diploid organism to a mutagen typically changes only one of the two copies of a given gene, yielding heterozygous mutants. Thus genetic screens must include inbreeding steps to generate progeny that are homozygous for the mutant allele. The geneticist H. Muller developed a general and efficient procedure for carrying out such inbreeding experiments in the fruit fly *Drosophila*. By using such procedures, recessive lethal mutations in *Drosophila* and other diploid organisms can be maintained in heterozygous individuals and their phenotypic consequences analyzed in homozygotes.

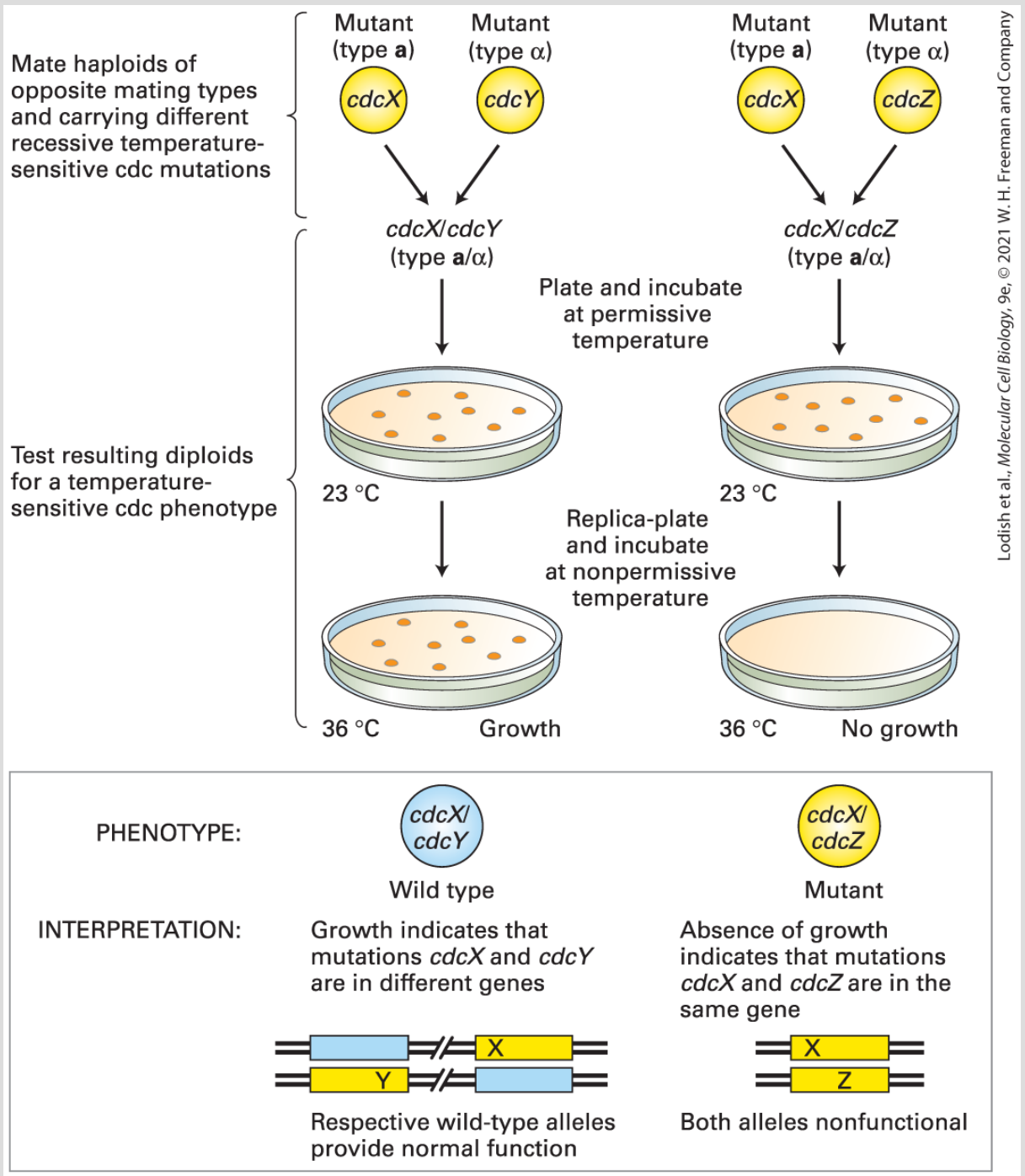
The Muller approach was used to great effect by Christiane Nüsslein-Volhard and Eric Wieschaus, who systematically screened for recessive lethal mutations affecting embryogenesis in *Drosophila*. The two researchers screened for dead homozygous embryos carrying recessive lethal mutations, then examined the dead embryos under the microscope for specific morphological defects. Current understanding of the molecular mechanisms underlying the development of multicellular organisms is based, in

large part, on the detailed picture of embryonic development revealed by characterization of these *Drosophila* mutants.

Complementation Tests Determine Whether Different Recessive Mutations Are in the Same Gene

Researchers using the classical genetic approach to studying a particular cellular process often isolate multiple recessive mutations that produce the same phenotype. A common test for determining whether these mutations are in the same gene or in different genes exploits the phenomenon of *genetic complementation*; that is, restoration of the wild-type phenotype by mating two different mutants. If two recessive mutations, *a* and *b*, are in the *same* gene, then a diploid organism carrying one *a* allele and one *b* allele will exhibit the mutant phenotype because neither allele provides a functional copy of the gene. In contrast, if mutations *a* and *b* are in *separate* genes, then heterozygotes carrying a single copy of each mutant allele will not exhibit the mutant phenotype because a wild-type allele of each gene is also present. In this case, the mutations are said to *complement* each other. Complementation analysis cannot be performed on dominant mutations because the phenotype conferred by the mutant allele is displayed even in the presence of a wild-type allele of the gene.

Complementation analysis of a set of mutants exhibiting the same phenotype can distinguish the individual genes in a set of functionally related genes, all of which must function to produce a given phenotypic trait. For example, the screen for cdc mutations in the yeast *S. cerevisiae* described previously yielded many recessive, temperature-sensitive mutants that appeared to be arrested at the same cell cycle stage. To determine how many genes were affected by these mutations, Hartwell and his colleagues performed complementation tests on all of the pair-wise combinations of their cdc mutants, following the general protocol outlined in Figure 6-7. These tests organized more than 100 cdc mutations into about 20 different *CDC* genes. The subsequent molecular characterization of the *CDC* genes and their encoded proteins, as described in detail in Chapter 19, has provided a framework for understanding how cell division is regulated in organisms ranging from yeast to humans.



EXPERIMENTAL FIGURE 6-7 Complementation analysis determines whether recessive mutations are in the same or different genes. Complementation tests in yeast are performed by mating haploid **a** and α cells carrying different recessive mutations to produce diploid cells. In the analysis of cdc mutations, pairs of different haploid temperature-sensitive cdc strains were systematically mated and the resulting diploids tested for growth at the permissive and nonpermissive temperatures. In this hypothetical example, the *cdcX*

and *cdcY* mutants complement each other and thus have mutations in different genes, whereas the *cdcX* and *cdcZ* mutants have mutations in the same gene.

Double Mutants Are Useful in Assessing the Order in Which Proteins Function

By careful analysis of mutant phenotypes associated with a particular cellular process, researchers can often deduce the order in which a set of genes and their protein products function. Two general types of processes are amenable to such analysis: (1) biosynthetic pathways in which a precursor material is converted via one or more intermediates to a final product, and (2) signaling pathways that regulate other processes and involve the flow of information rather than chemical intermediates.

Ordering of Steps in Biosynthetic Pathways

A simple example of the first type of process is the biosynthetic pathway for secreted proteins discussed in Chapter 14. In the secretory pathway, proteins move from their site of synthesis on the rough endoplasmic reticulum (ER) to the Golgi complex, then to secretory vesicles, and finally to the cell surface. The genetic analysis of this pathway in *S. cerevisiae* began, like the analysis of cell cycle mutants, with the isolation of temperature-sensitive mutants

that blocked different steps along the secretory pathway. These mutants could be categorized according to the type of organelle that accumulated at the nonpermissive temperature. For example, at the nonpermissive temperature, *sec12* mutants accumulated membranes that looked like ER membranes, whereas *sec4* mutants accumulated secretory vesicles. It was possible to establish the order in which the genes act along the pathway by analysis of double mutants. By constructing a double mutant that contained both *sec12* and *sec4*, it was conclusively shown that the *sec12* mutant acts earlier in the pathway than *sec4* because the double mutant accumulates ER membranes and not secretory vesicles. The general logic by which double mutants defective in two steps in a biosynthetic pathway are useful in ordering such pathways is shown in Figure 6-8a.

(a) Analysis of a biosynthetic pathway

A mutation in A accumulates intermediate 1.

A mutation in B accumulates intermediate 2.

PHENOTYPE OF DOUBLE MUTANT:	A double mutation in A and B accumulates intermediate 1.
INTERPRETATION:	The reaction catalyzed by A precedes the reaction catalyzed by B.

(b) Analysis of a signaling pathway

A mutation in A gives **repressed** reporter expression.

A mutation in B gives **constitutive** reporter expression.

PHENOTYPE OF

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

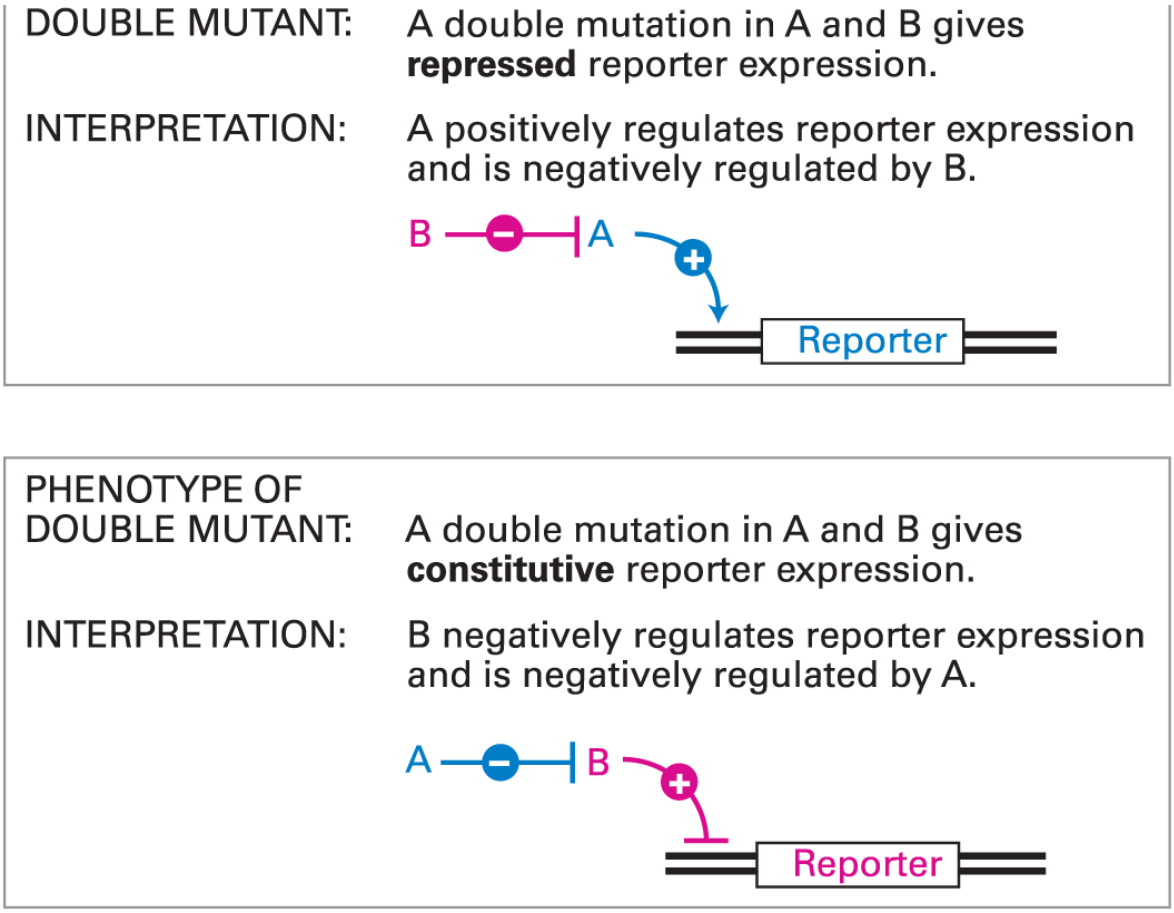


FIGURE 6-8 Analysis of double mutants can often order the steps in biosynthetic or signaling pathways. When mutations in two different genes affect the same cellular process but produce distinctly different phenotypes, the phenotype of the double mutant can often reveal the order in which the two genes must function. (a) In the case of mutations that affect the same biosynthetic pathway, a double mutant will accumulate the intermediate immediately preceding the step catalyzed by the protein that acts earlier in the wild-type organism. (b) Double-mutant analysis of a signaling pathway is possible if two mutations have opposite effects on expression of a reporter gene.

Ordering Steps of Signaling Pathways

As we will learn in later chapters, the expression of many eukaryotic genes is regulated by signaling pathways that are initiated by

extracellular hormones, growth factors, or other signals. Such signaling pathways may include numerous components, and double-mutant analysis can often provide insight into the functions and interactions of these components. The only prerequisite for obtaining useful information from this type of analysis is that the two mutations must have very different, or even opposite, effects on the output of the same regulated pathway as measured by expression of a reporter gene. Most commonly, one mutation represses expression of a particular reporter gene even when the signal is present, while another mutation results in the expression of that reporter gene even when the signal is absent (i.e., constitutive expression). As illustrated in Figure 6-8b, two simple regulatory mechanisms are consistent with such single mutants, but the double-mutant phenotype can distinguish between them. Thus the observed phenotype of the double mutant provides information about the order in which the proteins act and whether they are positive or negative regulators.

This general approach has enabled geneticists to delineate many of the key steps in a variety of different regulatory pathways, setting the stage for more specific biochemical assays.

Note that this technique differs from the complementation analysis just described in that both dominant and recessive mutants can be subjected to double-mutant analysis. When two recessive mutations

are tested, the double mutant created must be *homozygous* for both mutations.

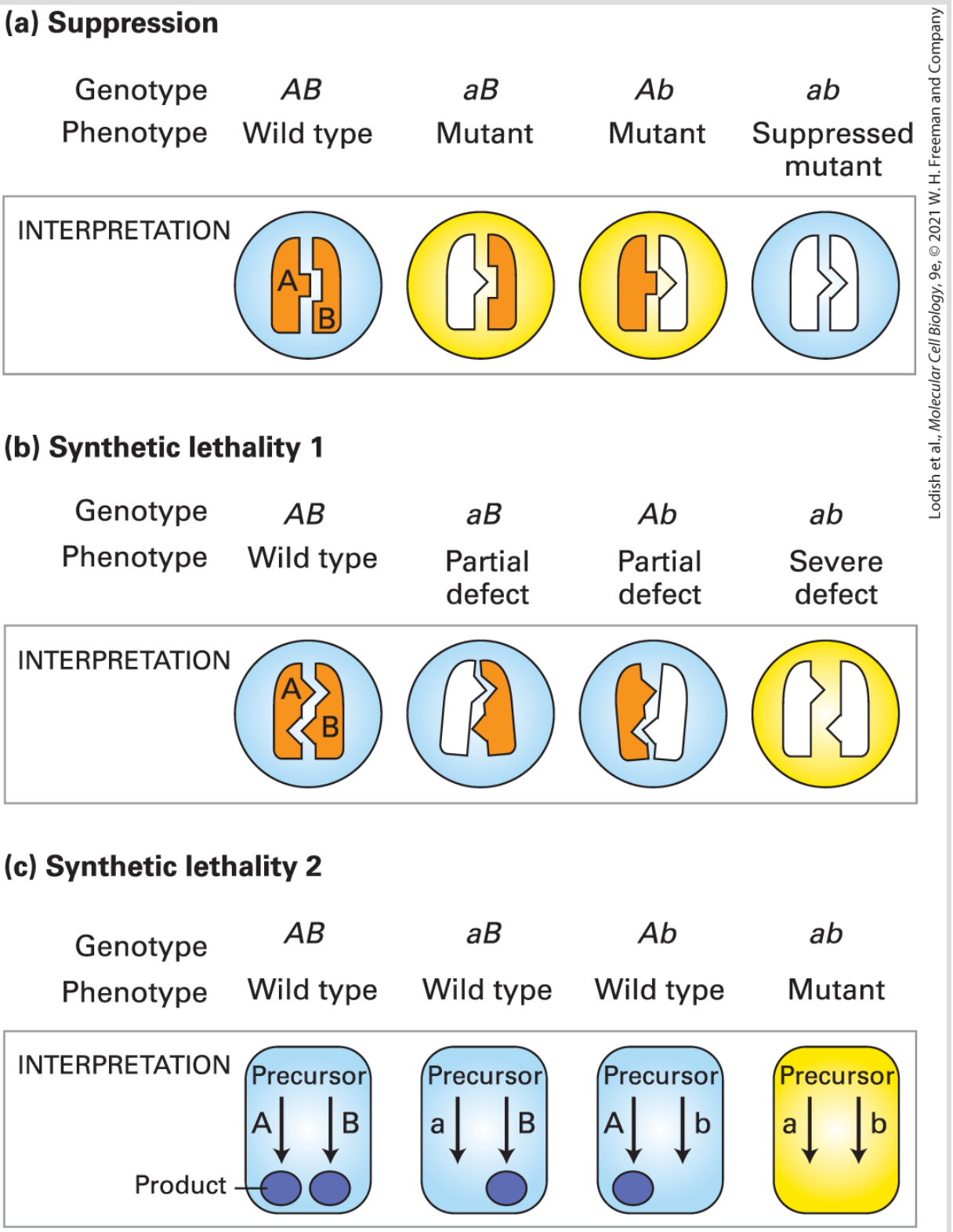
Genetic Suppression and Synthetic Lethality Can Reveal Interacting or Redundant Proteins

Two other types of genetic analysis can provide additional clues about how proteins that function in the same cellular process may interact with one another in the living cell. Both of these methods, which are applicable in many experimental organisms, involve the use of double mutants in which the phenotypic effects of one mutation are changed by the presence of a second mutation.

Suppressor Mutations

The first type of analysis is based on *genetic suppression*. To understand this phenomenon, suppose that point mutations lead to structural changes in one protein (A) that disrupt its ability to associate with another protein (B) involved in the same cellular process. Similarly, mutations in protein B lead to small structural changes that inhibit its ability to interact with protein A. Assume, furthermore, that the normal functioning of proteins A and B depends on their interacting. In theory, a specific structural change in protein A might be suppressed by compensatory changes in

protein B, allowing the mutant proteins to interact. In the rare cases in which such suppressor mutations occurred, strains carrying both mutant alleles would be normal, whereas strains carrying only one or the other mutant allele would have a mutant phenotype (Figure 6-9a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 6-9 Mutations that result in genetic suppression or synthetic lethality reveal interacting or redundant proteins. (a) The observation that double mutants with two

defective proteins (A and B) have a wild-type phenotype but that single mutants have a mutant phenotype indicates that the function of each protein depends on interaction with the other. (b) The observation that double mutants have a more severe phenotypic defect than single mutants is also evidence that two proteins (e.g., subunits of a heterodimer) must interact to function normally. (c) The observation that a double mutant is nonviable but that the corresponding single mutants have the wild-type phenotype indicates that two proteins function in redundant pathways to produce an essential product.

The observation of genetic suppression in yeast strains carrying a mutant actin allele (*act1-1*) and a second mutation in another gene (*sac6*) provided early evidence for a direct interaction in vivo between the proteins encoded by the two genes. Later biochemical studies showed that these two proteins — Act1 and Sac6 — do indeed interact in the construction of functional actin structures within the cell.

Synthetic Lethal Mutations

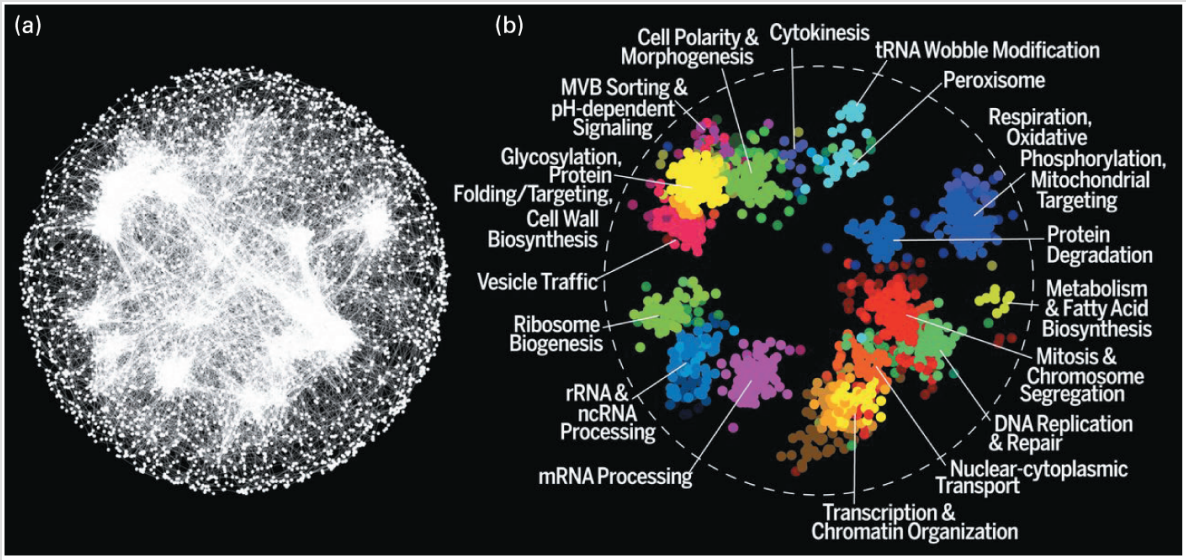
Another phenomenon, called *synthetic lethality*, produces a phenotypic effect opposite to that of suppression. In this case, the deleterious effect of one mutation is greatly exacerbated (rather than suppressed) by a second mutation in a related gene. One situation in which such synthetic lethal mutations can occur is illustrated in Figure 6-9b. In this example, a heterodimeric protein is partially, but not completely, inactivated by mutations in either one of its nonidentical subunits. However, in double mutants carrying specific mutations in the genes encoding both subunits, little interaction between subunits occurs, resulting in severe phenotypic

effects. Synthetic lethal mutations can also reveal nonessential genes whose encoded proteins function in redundant pathways for producing an essential cell component. As depicted in Figure 6-9c, if either pathway alone is inactivated by a mutation, the other pathway will be able to supply the needed product. However, if both pathways are inactivated at the same time, the essential product cannot be synthesized, and the double mutants will be nonviable.

Global Analysis of Double Mutant Combinations Can Reveal Networks of Gene Functions

In the preceding sections we have seen how evaluation of the phenotype of double mutants can provide information about whether two genes participate in the same process or in redundant processes. We have also seen how double mutants can be used to determine the order of function for genes that participate in the same biosynthetic or regulatory pathway. The consequences of double mutants can usually be interpreted in more than one way, and the phenotype of any particular double mutant rarely provides a definitive conclusion about how the participating genes interact. However, when pair-wise combinations of double mutants from large sets of genes are examined together, convincing patterns of gene interactions can emerge.

The first comprehensive analysis of double mutants on a genome-wide scale was performed in *S. cerevisiae,* which has a total of about 6000 genes. Mutations in the approximately 1000 essential genes were obtained as temperature-sensitive alleles as described previously for *cdc* mutants, and the 5000 genes that are not essential were deleted, as described in Section 6.6. Using an automated system to perform genetic crosses, almost every possible double mutant was constructed and its phenotype compared to that of each of the single-component mutants. Computer algorithms determined the type and strength of genetic interaction based on the logic outlined previously. These algorithms produced a two-dimensional map of genetic interactions that clusters genes according to how closely related their functions are (Figure 6-10). Remarkably, when looking at the genes for which something is already known about their function, this interaction map, based only on information from the phenotype of double mutants, closely matches the known organization of metabolic and cell biologic processes. Importantly, for the many genes whose function is still unknown, their placement on the interaction map provides important clues to their function that may guide future biochemical and genetic studies. Methods using CRISPR-Cas9 now enable mutations to be produced in mammalian cells, as described in Section 6.6. These methods should allow similar large-scale gene interaction networks to be mapped in human cells.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 6-10 Gene Interaction Network for *S. cerevisiae*. The approximately 6000 genes of *S. cerevisiae* were tested in every possible pair-wise combination (a total of 2.3×10^7 double mutants were constructed). The relationships between genes were then arranged based on the type and strength of their genetic interactions by a computer algorithm. (a) In this global map of gene interactions, each point represents a single *S. cerevisiae* gene and the length of the segments connecting genes represents the degree to which genes display similar genetic interactions, with genes that exhibit the most closely related functions clustered most closely together. (b) A sample of genes of known function, color coded according to their cellular function, are placed on the global map of gene interactions. Importantly, genes known to be involved in the same process are in close proximity on the genetic interaction map.

[Reprinted by permission from American Association for the Advancement of Science, from M. Costanzo et al., 2016. A global genetic interaction network maps a wiring diagram of cellular function. *Science*, **353**(6306):aaf1420. https://doi.org/10.1126/science.aaf1420; permission conveyed through Copyright Clearance Center, Inc.]

KEY CONCEPTS OF SECTION 6.1

Genetic Analysis of Mutations to Identify and Study Genes

- Diploid organisms carry two copies (alleles) of each gene, whereas haploid organisms carry only one copy.
- Recessive mutations often lead to a loss of function, which is masked if a wild-type allele of the gene is present. For the mutant phenotype to be observed, both alleles must carry the mutation.
- Dominant mutations lead to a mutant phenotype in the presence of a wild-type allele of the gene. The phenotypes associated with dominant mutations often represent a gain of function.
- In meiosis, a diploid cell undergoes one DNA replication and two cell divisions, yielding four haploid cells in which maternal and paternal chromosomes and their associated alleles are randomly assorted (see Figure 6-3).
- Dominant and recessive mutations exhibit characteristic segregation patterns in genetic crosses (see Figure 6-4).
- In haploid yeast, temperature-sensitive mutations are particularly useful for identifying and studying genes essential to survival.
- The number of functionally related genes involved in a process can be defined by complementation analysis (see Figure 6-7).
- The order in which genes function in a biosynthetic or signaling pathway can be deduced from the phenotype of double mutants defective in two steps in the affected process.
- Functionally significant interactions between proteins can be deduced from the phenotypic effects of allele-specific suppressor mutations or synthetic lethal mutations.
- By evaluating combinations of double mutants for all of the genes in an organism, it is possible to construct a global map of the network of functional gene interactions within a cell. Such a map has been constructed for the 6000 genes of *S. cerevisiae*, and it should be possible to construct an equivalent map of the 20,000 mammalian genes.

6.2 DNA Cloning and Characterization

Detailed studies of the structure and function of a gene at the molecular level require large quantities of the individual gene in pure form. Researchers use **DNA cloning** as a way to prepare large numbers of identical DNA molecules. DNA cloning uses a variety of techniques, often referred to as *recombinant DNA technology*. **Recombinant DNA** is simply any DNA molecule composed of sequences derived from different organisms.

The key to cloning a DNA fragment of interest is to link it into a **vector** DNA molecule that can replicate within a host cell. Such a vector could be a bacterial plasmid or a modified viral genome, for example. The DNA fragment is inserted into the vector, forming a single recombinant DNA molecule. That recombinant molecule is introduced into a host cell, and the inserted DNA is replicated along with the vector, generating a large number of identical DNA molecules. The basic scheme can be summarized as follows:

$$\begin{array}{c}\text{Vector} + \text{DNA fragment}\\ \downarrow \\ \text{Recombinant DNA}\\ \downarrow \\ \text{Replication of recombinant DNA within host cells}\\ \downarrow \\ \text{Isolation, sequencing, and manipulation}\\ \text{of purified DNA fragment}\end{array}$$

Although investigators have devised numerous experimental variations, this flow diagram indicates the essential steps in DNA cloning.

In this section, we first describe methods for isolating a specific sequence of DNA from a sea of other DNA sequences. This process often involves cutting the genome into fragments and then placing each fragment in a vector so that the entire collection can be propagated as recombinant molecules in separate host cells. While many different types of vectors exist, our discussion will mainly focus on plasmid vectors in *Escherichia coli* host cells, which are commonly used. Various techniques can then be employed to identify the sequence of interest from this collection of DNA fragments. Once a specific DNA fragment is isolated, the exact sequence of nucleotides in the fragment can be determined. We end with a discussion of the polymerase chain reaction (PCR). This powerful and versatile technique can be used in many ways to

generate large quantities of a specific sequence and to manipulate DNA in the laboratory. The various uses of cloned DNA fragments are discussed in subsequent sections.

Restriction Enzymes and DNA Ligases Allow Insertion of DNA Fragments into Cloning Vectors

A major objective of DNA cloning is to obtain discrete, small regions of an organism's DNA that constitute specific genes. In addition, only relatively small DNA molecules can be inserted into any of the available vectors. For these reasons, the very long DNA molecules that compose an organism's genome must be cleaved into fragments that can be inserted into the vector DNA. Two types of enzymes — restriction enzymes and DNA ligases — are used to produce such recombinant DNA molecules.

Cutting DNA Molecules into Small Fragments

Restriction enzymes are endonucleases produced by bacteria that typically recognize specific 4–8-bp sequences, called *restriction sites*, and cleave both DNA strands at these sites. Restriction sites commonly are short *palindromic* sequences; that is, the restriction-site sequence is the same on each DNA strand when read in the 5′ to 3′ direction (Figure 6-11).

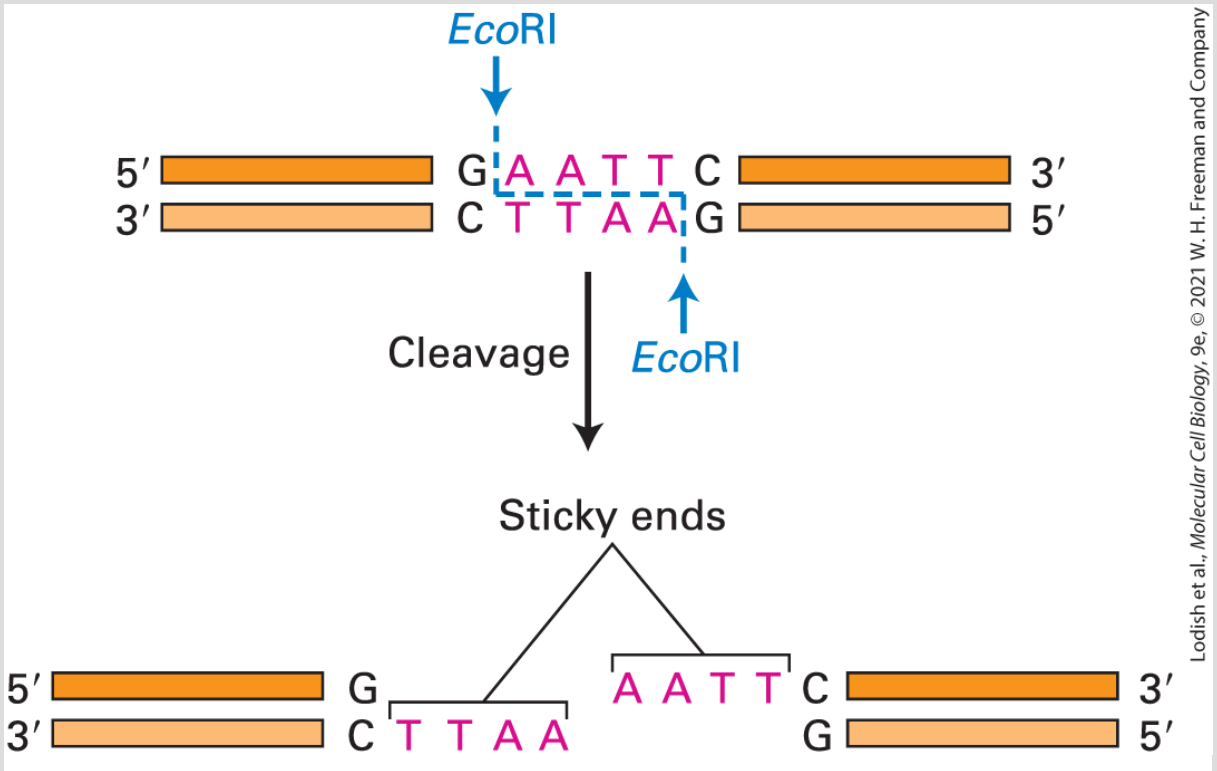


FIGURE 6-11 Cleavage of DNA by the restriction enzyme *Eco*RI. This restriction enzyme from *E. coli* makes staggered cuts at the specific 6-bp palindromic sequence shown, yielding fragments with single-stranded, complementary 4-base sticky ends. Many other restriction enzymes also produce fragments with sticky ends.

For each restriction enzyme, bacteria also produce a *modification enzyme*, which protects a host bacterium's own DNA from cleavage by modifying the host DNA at or near each potential cleavage site. The modification enzyme adds a methyl group to one or two bases, usually within the restriction site. When a methyl group is present there, the restriction endonuclease is prevented from cutting the DNA. Together, the restriction endonuclease and the modification enzyme form a restriction-modification system that protects the host DNA while it destroys incoming foreign DNA (e.g.,

bacteriophage DNA or DNA taken up during transformation) by cleaving it at all the available restriction sites.

Many restriction enzymes make staggered cuts in the two DNA strands at the corresponding restriction site, generating fragments that have a single-stranded *tail* at both ends (see Figure 6-11). The tails on the fragments generated at a given restriction site are complementary to those on all other fragments generated by the same restriction enzyme. At room temperature, these *sticky ends* can transiently base-pair with those on other DNA fragments generated with the same restriction enzyme.

The DNA isolated from an individual organism has a specific sequence that, purely by chance, contains a specific set of restriction sites. Thus a given restriction enzyme will cut the DNA from a particular source into a reproducible set of fragments called *restriction fragments*. The frequency with which a restriction enzyme cuts DNA, and thus the average size of the resulting restriction fragments, depends largely on the length of the recognition site. For example, a restriction enzyme that recognizes a 4-bp site will cleave DNA an average of once every 4^4, or 256, base pairs, whereas an enzyme that recognizes an 8-bp sequence will cleave DNA an average of once every 4^8 base pairs (65 kb). Hundreds of different restriction enzymes have been identified from different species of bacteria. Thus DNA molecules can be cut at a large number of

different sequences corresponding to the recognition sites of these enzymes.

Inserting DNA Fragments into Vectors

DNA fragments with either sticky ends or blunt ends can be inserted into vector DNA with the aid of *DNA ligase*. During normal DNA replication, DNA ligase catalyzes the end-to-end joining (ligation) of short fragments of DNA. For purposes of DNA cloning, purified DNA ligase is used to covalently join together the ends of a restriction fragment with the ends of the restriction enzyme–cut vector DNA by forming the standard covalent phosphodiester bonds of DNA (Figure 6-12). DNA ligase from bacteriophage T4 can ligate complementary sticky ends as well as two blunt DNA ends. However, blunt-end ligation is inherently inefficient and requires a higher concentration of both DNA and DNA ligase than does ligation of sticky ends.

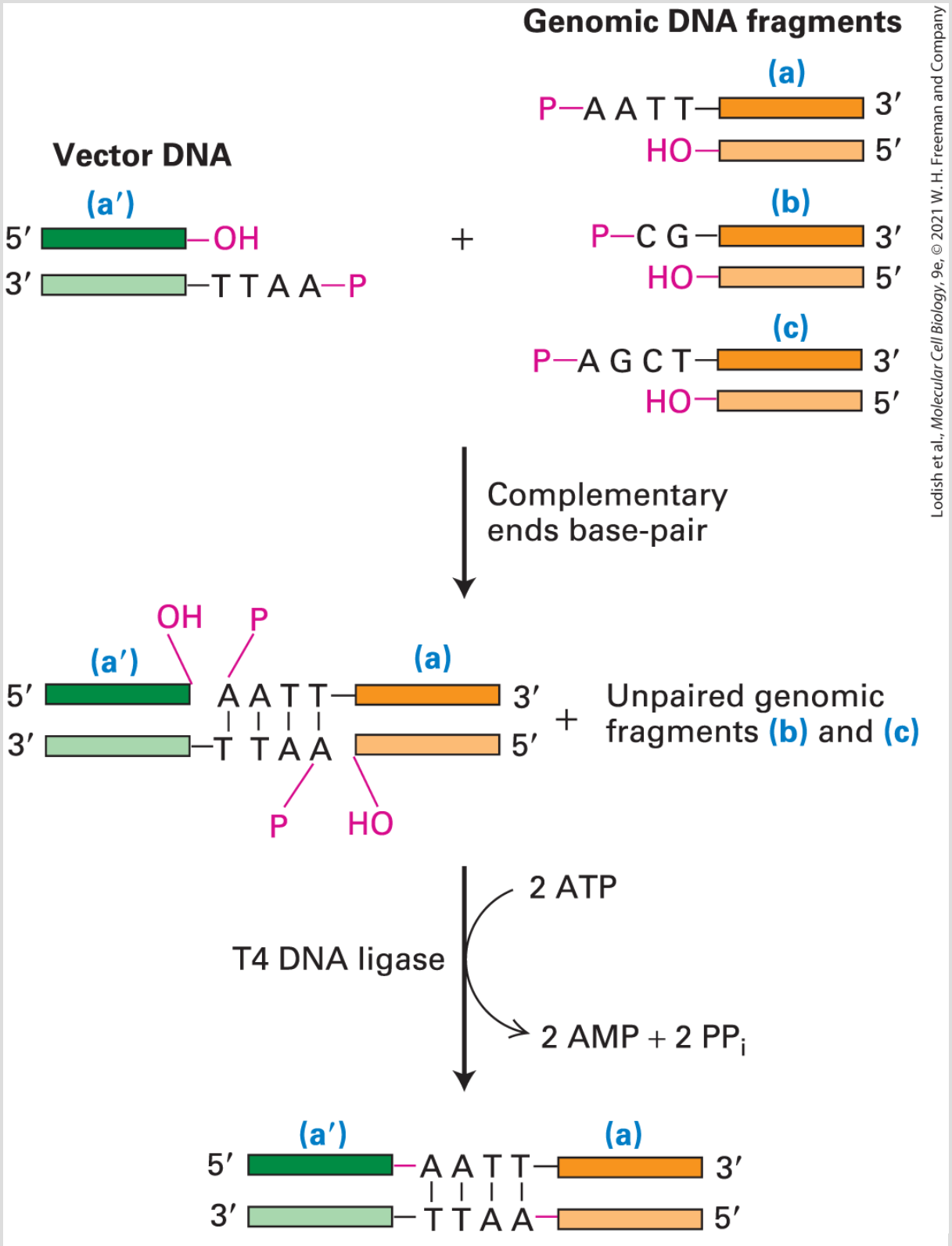


FIGURE 6-12 Ligation of restriction fragments with complementary sticky ends. In this example, vector DNA cut with *Eco*RI is mixed with a sample containing restriction fragments

produced by cleaving genomic DNA with several different restriction enzymes. The short base sequences composing the sticky ends of each fragment type are shown. The sticky end on the cut vector DNA (a') base-pairs only with the complementary sticky ends on the *Eco*RI fragment (a) in the genomic sample. The adjacent $3'$ hydroxyl and $5'$ phosphate groups (red) on the base-paired fragments are then covalently joined (ligated) by T4 DNA ligase.

Isolated DNA Fragments Can Be Cloned into *E. coli* Plasmid Vectors

Let's take a closer look at the plasmids used as vectors. **Plasmids** are circular, double-stranded DNA (dsDNA) molecules that replicate separately from a cell's chromosomal DNA. These extrachromosomal DNAs, which occur naturally in bacteria and in lower eukaryotic cells (e.g., yeast), exist in a symbiotic relationship with their host cell. Like the host-cell chromosomal DNA, plasmid DNA is duplicated before every cell division. During cell division, copies of the plasmid DNA segregate to each daughter cell, ensuring that the plasmid continues to be propagated through successive generations of the host cell.

The plasmids most commonly used in recombinant DNA technology are those that replicate in *E. coli*. Investigators have engineered these plasmids to optimize their use as vectors in DNA cloning — for instance, by removing unneeded portions from naturally occurring *E. coli* plasmids. Removal of unneeded portions yields plasmid vectors about 1.2–3 kb in length that contain three regions essential

for DNA cloning: a replication origin (ORI); a marker that permits selection of plasmids that have the incorporated gene (*selectable marker*), usually a drug-resistance gene; and a region in which exogenous DNA fragments can be inserted (Figure 6-13).

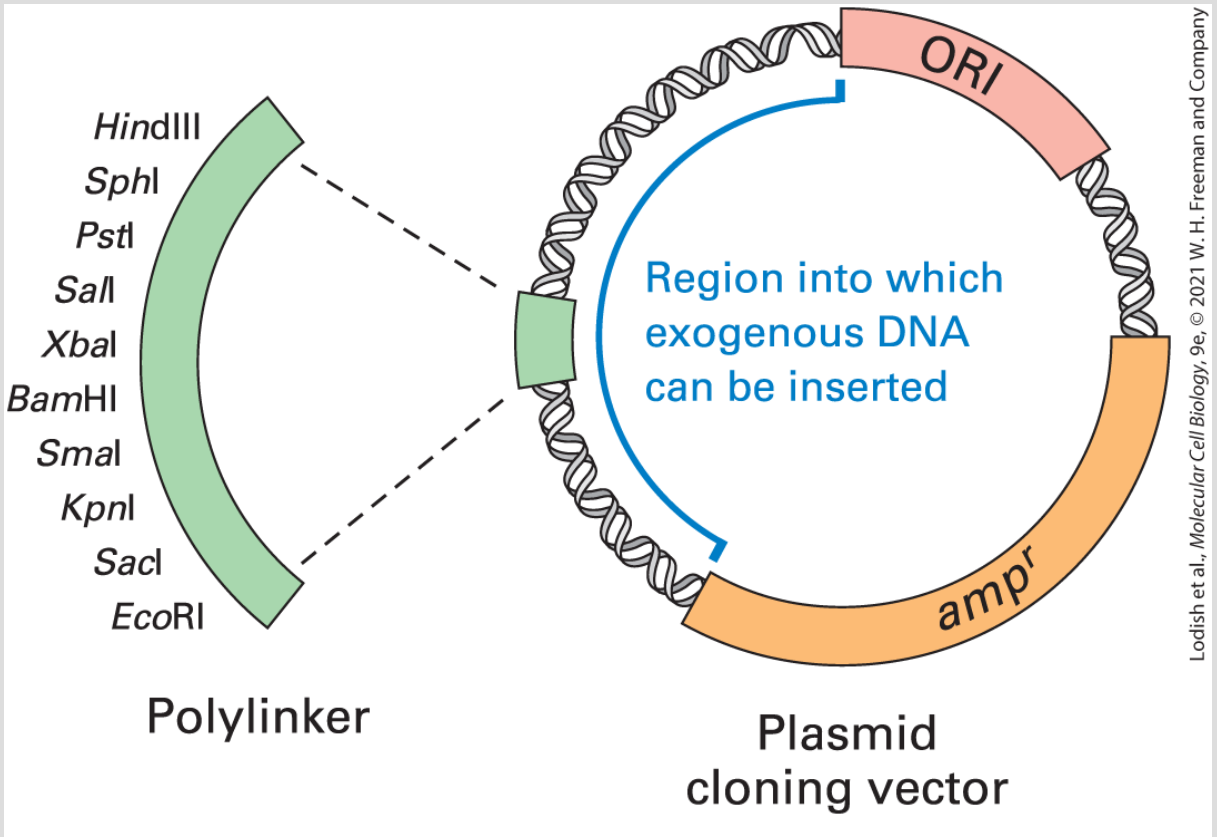
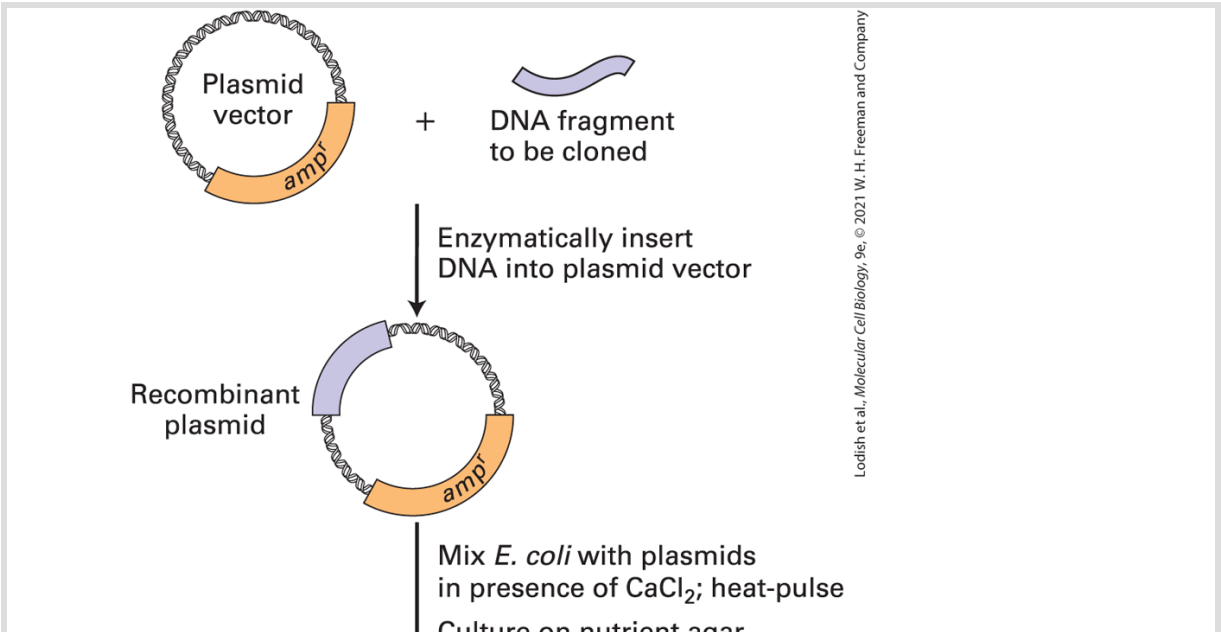
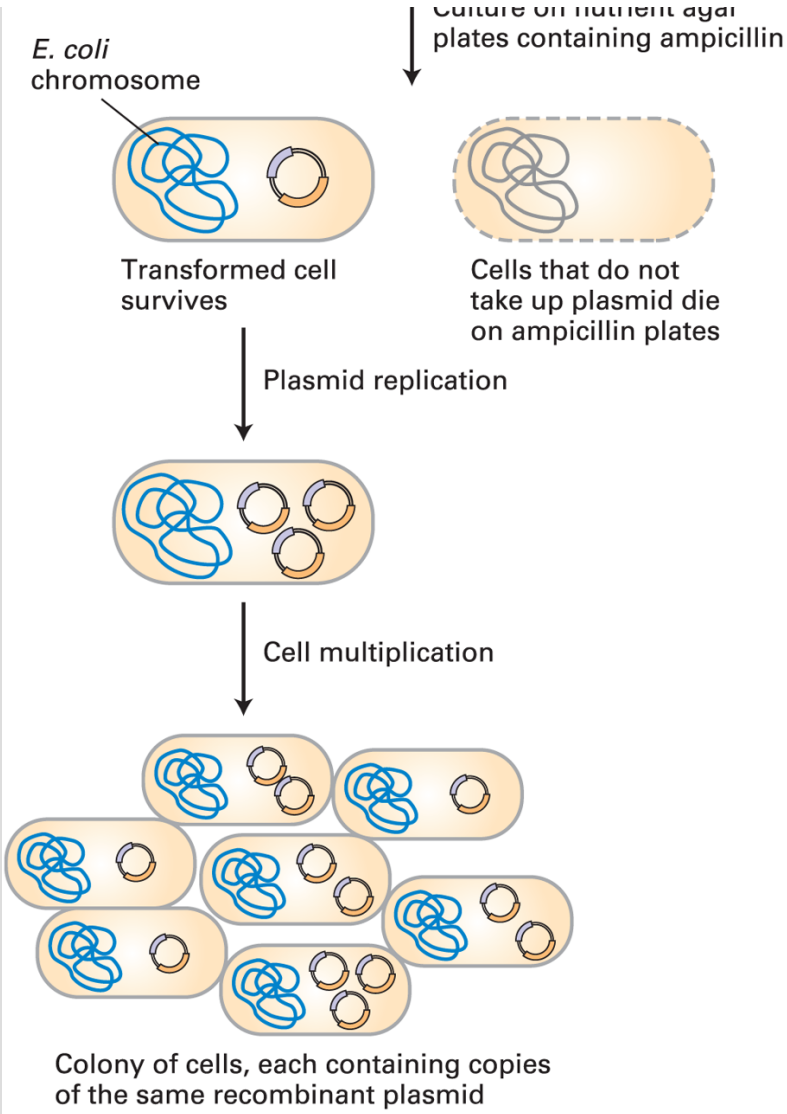


FIGURE 6-13 Basic components of a plasmid cloning vector that can replicate within an *E. coli* cell. Plasmid vectors contain a replication origin (ORI) sequence where DNA replication can be initiated by host-cell enzymes. They also contain a gene that functions as a selectable marker, such as amp^r, which encodes the enzyme β-lactamase and confers resistance to ampicillin. Inclusion of a synthetic polylinker containing the recognition sequences for several different restriction enzymes increases the versatility of a plasmid vector since the vector is designed so that each site in the polylinker is unique on the plasmid. Exogenous DNA can be inserted into the polylinker without disturbing the ability of the plasmid to replicate or express the amp^r gene.

Figure 6-14 outlines the general procedure for cloning a DNA fragment using *E. coli* plasmid vectors. When *E. coli* cells are mixed with recombinant vector DNA and subjected to a stress such as heat shock, a small fraction of the cells will take up the plasmid DNA, a process known as **transformation**. Typically, 1 cell in about 10,000 incorporates a *single* plasmid DNA molecule and thus becomes transformed. The rare transformed cells can be easily selected by use of a selectable marker. For instance, if the plasmid carries a gene that confers resistance to the antibiotic ampicillin, transformed cells can be selected by growing them in a medium containing ampicillin. All the antibiotic-resistant progeny cells that arise from the initial transformed cell will contain plasmids with the same inserted DNA. Since all the cells in a colony arise from a single transformed parent cell, they constitute a **clone** of cells, and the initial fragment of DNA inserted into the parental plasmid is referred to as *cloned DNA* or a *DNA clone*.





EXPERIMENTAL FIGURE 6-14 DNA cloning in a plasmid vector permits amplification of a DNA fragment. A fragment of DNA to be cloned is first inserted into a plasmid vector containing an ampicillin-resistance gene (amp^r), such as that shown in Figure 6-13. Only the few cells transformed by incorporation of a plasmid will survive on ampicillin-containing medium. In transformed cells, the plasmid DNA replicates and segregates into daughter cells, resulting in the formation of an ampicillin-resistant colony in which each cell contains the cloned DNA.

The versatility of an *E. coli* plasmid vector is increased by the addition of a *polylinker*, a synthetically generated sequence

containing one copy of each of several different restriction sites that are not present elsewhere in the plasmid sequence (see Figure 6-13). Typically, the polylinker is cleaved with two different restriction enzymes in preparation to accept a DNA fragment prepared with two different sticky ends formed by cleavage by the same two enzymes. Such a strategy eliminates unwanted by-products such as a cleaved plasmid vector that has reclosed and greatly increases the efficiency of cloning DNA fragments.

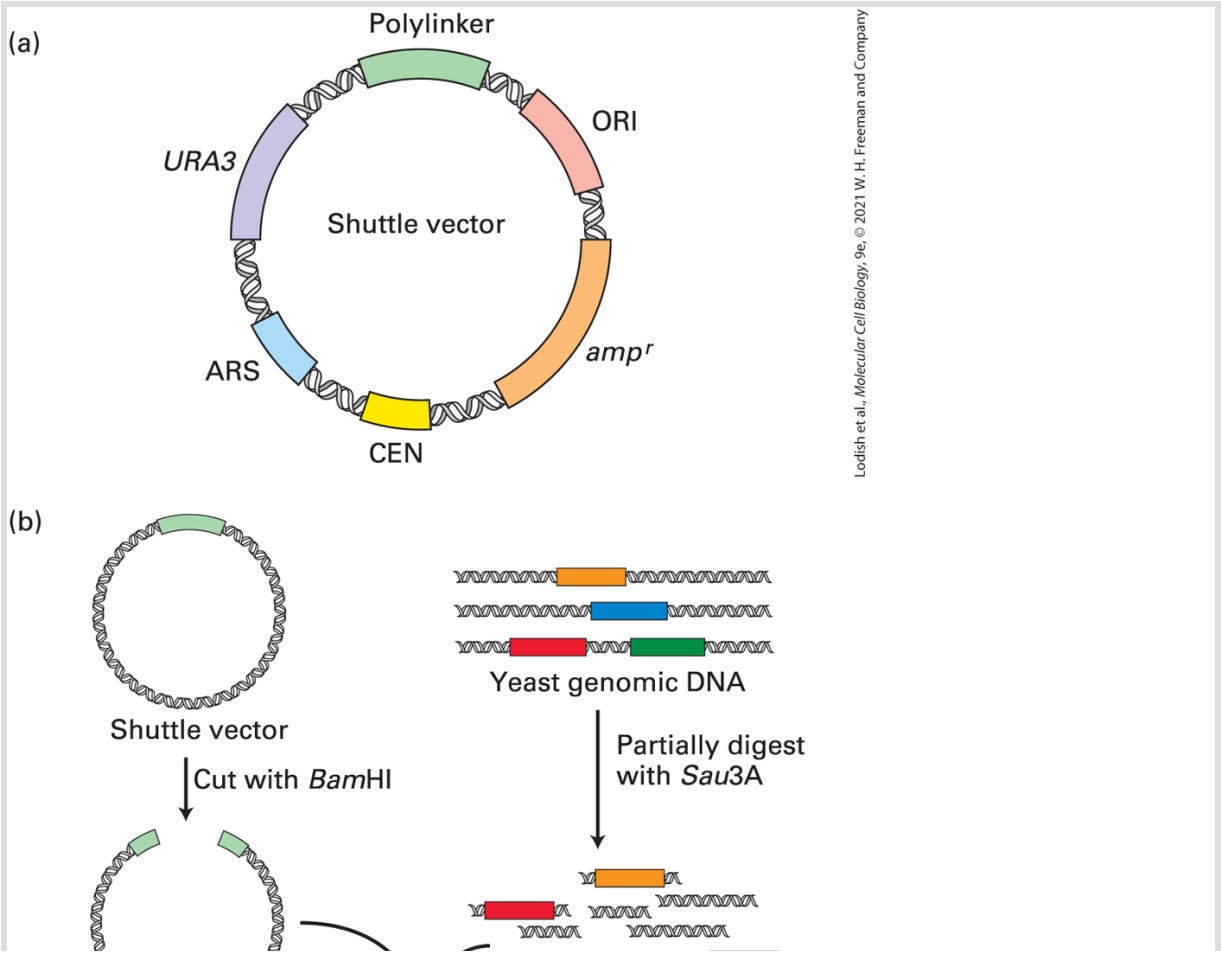
Plasmid cloning vectors are useful for propagating DNA fragments up to about 20 kb in length, but fragments longer than this cannot be reliably replicated within one cell-division cycle. For some purposes, such as the isolation and manipulation of large segments of the human genome, it is desirable to clone DNA segments as large as several megabases [1 megabase (Mb) = 1 million base pairs]. For this purpose, specialized plasmid vectors known as *BACs (b*acterial *a*rtificial *c*hromosomes) have been developed. One type of BAC uses a replication origin, derived from an *E. coli* plasmid, known as the *F factor*. The F factor, and cloning vectors derived from it, can be stably maintained at a single copy per *E. coli* cell even when it contains inserted sequences of up to about 2 Mb. Production of BAC libraries requires special methods for the isolation, ligation, and transformation of large segments of DNA because segments of DNA larger than about 20 kb are highly vulnerable to mechanical breakage even by standard manipulations such as pipetting.

Yeast Genomic Libraries Can Be Constructed with Shuttle Vectors and Screened by Functional Complementation

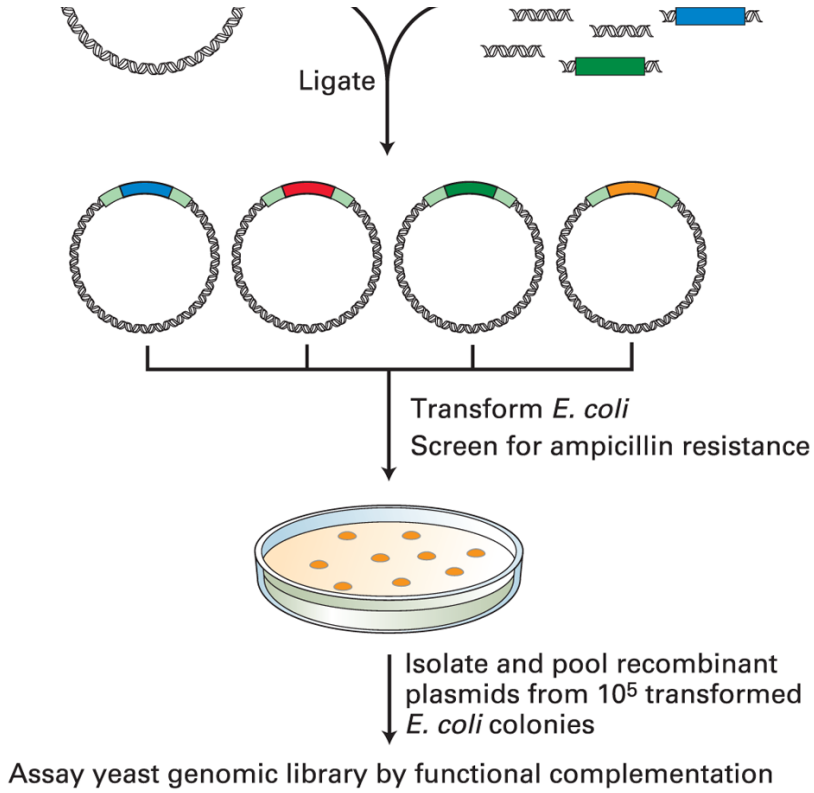
A collection of DNA molecules, each cloned into a vector molecule, is known as a **DNA library**. When genomic DNA from a particular organism is the source of the starting DNA, the set of clones that collectively represent all the DNA sequences in the genome is known as a *genomic library*. Once a genomic library is established, investigators need ways to isolate the cloned genes that are relevant to the function they are studying. For example, investigators may have identified an interesting recessive mutation in an experimental organism and would like to isolate the cloned wild-type copy of the gene from their genomic library. The method frequently used for this is referred to as **functional complementation**.

Let's look at a functional complementation assay in yeast. Yeast genes that are cloned in special *E. coli* plasmids can be introduced into mutant yeast cells to identify the wild-type gene that is defective in the mutant strain. Because genes of the yeast *Saccharomyces* do not contain multiple introns, they are sufficiently compact that an entire sequence of as many as 10 genes can be included in a genomic DNA fragment inserted into a plasmid vector.

A plasmid genomic library that is to be screened by functional complementation in yeast cells needs to be constructed so that it contains clones of all the yeast genes. The plasmid vector carrying the clones must be capable of replication in both *E. coli* cells and yeast cells: in *E. coli* cells to produce the clones and in yeast cells to screen for functional complementation. This type of vector, capable of propagation in two different hosts, is called a **shuttle vector**. The structure of a typical yeast shuttle vector is shown in Figure 6-15a. This vector contains the basic elements that permit cloning of DNA fragments in *E. coli* as well as sequences required for its propagation in yeast.



Lodish et al., Molecular Cell Biology, 9e, © 2021 W. H. Freeman and Company



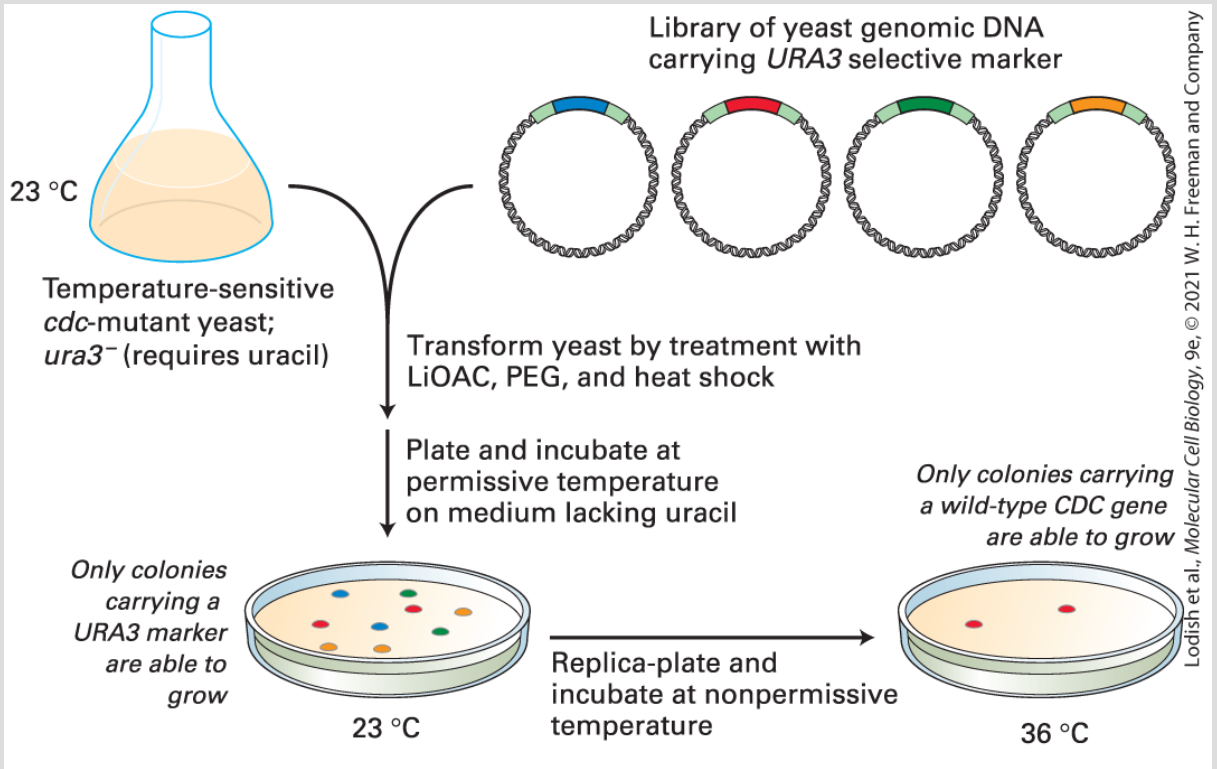
EXPERIMENTAL FIGURE 6-15 A yeast genomic library can be constructed in a plasmid shuttle vector that can replicate in yeast and in *E. coli*. (a) Components of a typical plasmid shuttle vector for cloning *Saccharomyces* genes. The presence of a yeast replication origin (ARS) and a yeast centromere (CEN) allows stable replication and segregation in yeast. Also included is a yeast selectable marker such as *URA3*, which allows a *ura3* mutant to grow on medium lacking uracil. Finally, the vector contains sequences for replication and selection in *E. coli* (ORI and amp^r) and a polylinker for easy insertion of yeast DNA fragments. (b) Typical protocol for constructing a yeast genomic library. Partial digestion of total yeast genomic DNA with *Sau*3A is adjusted to generate fragments with an average size of ~10 kb. The vector is prepared to accept the genomic fragments by digestion with *Bam*HI, which produces the same sticky ends as *Sau*3A. Each transformed clone of *E. coli* that grows after selection for ampicillin resistance contains a single type of yeast DNA fragment.

To increase the probability that all regions of the yeast genome will be successfully cloned and represented in the plasmid library, the genomic DNA is usually only partially digested to yield overlapping

restriction fragments of ~10 kb. These fragments are then ligated into a shuttle vector in which the polylinker has been cleaved with a restriction enzyme that produces sticky ends complementary to those on the yeast DNA fragments (Figure 6-15b). Because the 10-kb restriction fragments of yeast DNA are incorporated into the shuttle vectors randomly, at least 10^5 *E. coli* colonies, each containing a particular recombinant shuttle vector, are necessary to ensure that each region of yeast DNA has a high probability of being represented in the library at least once.

Figure 6-16 outlines how such a yeast genomic library can be screened to isolate the wild-type gene corresponding to one of the temperature-sensitive *cdc* mutations mentioned earlier in this chapter. The starting yeast strain is a double mutant that requires uracil for growth due to a $ura3^-$ mutation and is temperature sensitive due to a *cdc28* mutation identified by its phenotype (see Figure 6-6b). Recombinant plasmids isolated from the yeast genomic library are mixed with yeast cells under conditions that promote transformation of the cells with foreign DNA. Since transformed yeast cells carry a plasmid-borne copy of the wild-type *URA3* gene, they can be selected by their ability to grow in the absence of uracil. Typically, about 20 petri dishes, each containing about 500 yeast transformants, are sufficient to represent the entire yeast genome. This collection of yeast transformants can be maintained at 23 °C, a temperature permissive for growth of the *cdc28* mutant. The entire collection on 20 plates is then transferred to replica plates, which

are maintained at 36 °C, a nonpermissive temperature for *cdc* mutants. Yeast colonies that carry recombinant plasmids expressing a wild-type copy of the *CDC28* gene will be able to grow at 36 °C. Once temperature-resistant yeast colonies have been identified, plasmid DNA can be extracted from the cultured yeast cells and analyzed by DNA sequencing, a topic we take up shortly.



EXPERIMENTAL FIGURE 6-16 Screening of a yeast genomic library by functional complementation can identify clones carrying the normal form of a mutant yeast gene. In this example, a wild-type *CDC28* gene is isolated by complementation of a temperature-sensitive *cdc28* yeast mutant. The yeast genomic library prepared as shown in Figure 6-15 is transformed into a $ura3^-$, temperature-sensitive cdc mutant strain. The relatively few transformed yeast cells, which contain recombinant plasmid DNA, can grow in the absence of uracil at 23 °C. When these colonies are replica-plated and incubated at 36 °C (a nonpermissive temperature), only clones carrying a library plasmid that contains the wild-type copy of the *CDC* gene will survive. LiOAC = lithium acetate; PEG = polyethylene glycol.

cDNA Libraries Represent the Sequences of Protein-Coding Genes

Genomic libraries are ideal for representing the genetic content of relatively simple organisms such as bacteria or yeast, but present certain experimental difficulties for higher eukaryotes. First, the genes of eukaryotes usually contain extensive intron sequences and can therefore be too large to be inserted intact into plasmid vectors. Within a genomic library, the sequences of individual genes are often broken apart and carried in more than one clone. Moreover, the presence of introns and long intergenic regions in genomic DNA often makes it difficult to identify the important parts of a gene that actually encode protein sequences. For example, only about 1.5 percent of the human genome actually represents protein-coding gene sequences. Thus for many studies, cellular mRNAs, which lack the noncoding regions present in genomic DNA, are a more useful starting material for generating a DNA library. In this approach, DNA copies of mRNAs, called **complementary DNAs (cDNAs)**, are synthesized and cloned into plasmid vectors. A large collection of the resulting cDNA clones, representing all the mRNAs expressed in a cell type, is called a *cDNA library*.

The first step in preparing a cDNA library is to isolate the total mRNA from the cell type or tissue of interest. Because of their poly(A) tails, mRNAs are easily separated from the much more prevalent rRNAs and tRNAs present in a cell extract by use of a

matrix to which short strings of thymidylate (oligo-dTs) are linked. The crucial step for preparing cDNAs from cellular mRNAs is the synthesis of a DNA strand complementary to the mRNA. This synthesis is carried out by the enzyme *reverse transcriptase*, which is found in retroviruses, and is similar to DNA polymerase except that reverse transcriptase uses RNA as a template for synthesis of a complementary DNA strand. As with DNA polymerase, DNA synthesis by reverse transcriptase requires a DNA primer. For cDNA synthesis, oligo-dT can efficiently base-pair with the poly(A) tails of mRNAs and serve as a primer for synthesis of a complementary DNA strand (Figure 6-17). In several steps, the resulting cDNA-mRNA hybrid molecules are converted into double-stranded cDNA molecules and DNA sequences are appended onto the ends of the cDNA to enable cloning into an appropriate vector. Figure 6-17 outlines one of the many possible procedures for generating a cDNA library inserted into a plasmid vector.

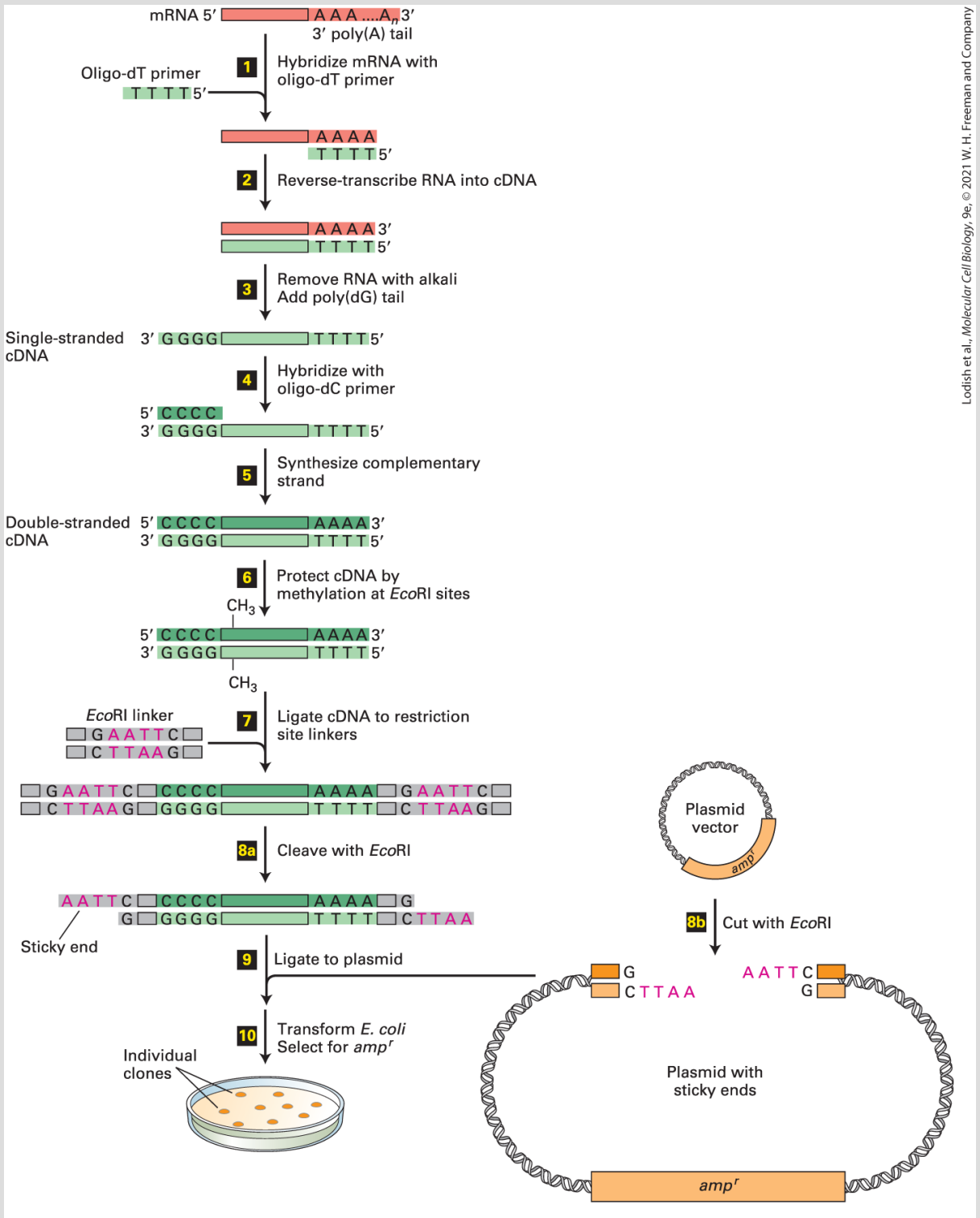


FIGURE 6-17 Generation and cloning of a cDNA copy of an mRNA. A mixture of mRNAs is the starting point for preparing recombinant plasmid clones, each containing a unique cDNA. The enzyme reverse transcriptase is used to synthesize a strand of DNA

complementary to each mRNA molecule, starting from an oligo-dT primer (steps 1 and 2). The resulting cDNA-mRNA hybrid molecules are converted in several steps into double-stranded cDNA molecules corresponding to all the mRNA molecules in the original preparation (steps 3 – 5). Each double-stranded cDNA contains an oligo-dC·oligo-dG double-stranded region at one end and an oligo-dT·oligo-dA double-stranded region at the other end. Methylation of the cDNA protects it from subsequent restriction enzyme cleavage (step 6).

To prepare double-stranded cDNAs for cloning, short double-stranded DNA molecules containing the recognition site for a particular restriction enzyme (called *linkers*) are ligated to both ends of the cDNAs using DNA ligase from bacteriophage T4 (step 7). As noted earlier, this ligase can join *blunt-ended*, double-stranded DNA molecules lacking sticky ends. The resulting molecules are then treated with the restriction enzyme specific for the attached linker, generating cDNA molecules with sticky ends (step 8a). In a separate procedure, plasmid DNA is treated with the same restriction enzyme to produce the appropriate sticky ends (step 8b).

The plasmid vector and the collection of cDNAs, all containing complementary sticky ends, are then mixed and joined covalently by DNA ligase (step 9). The resulting DNA molecules are introduced into *E. coli* cells to generate individual clones; each clone carries a cDNA derived from a single mRNA.

Only expressed gene sequences are represented in cDNA libraries, and screening cDNA libraries can be a highly efficient means to identify genes of interest in large complex genomes, such as those of mammals. However, there are two technical limitations to the utility of cDNA libraries. The first is that because of the inherent instability of RNA molecules full-length mRNAs may be partially degraded before the cDNA strand is synthesized. The longer the mRNA, the more likely that some degradation will occur; for very long mRNAs it is unlikely that the cDNA will include the 5′ end of the mRNA sequence.

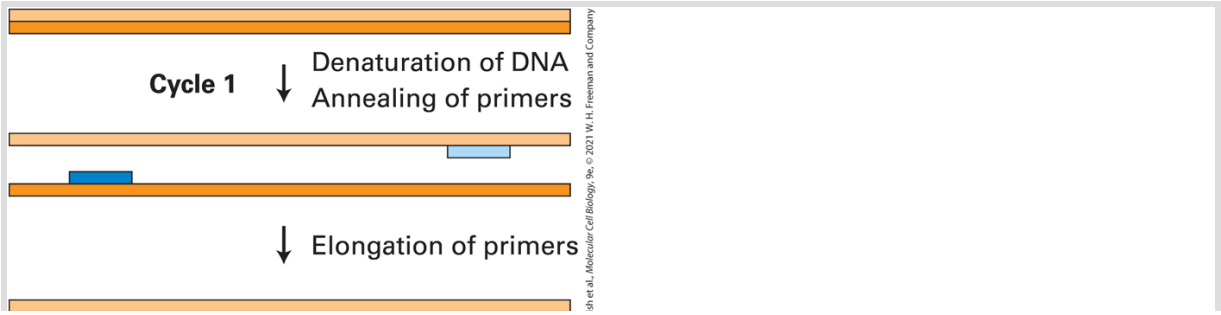
The second limitation is due to different genes being transcribed at very different rates. Because of this, cDNA clones corresponding to abundantly transcribed genes will be represented many times in a cDNA library, whereas cDNAs corresponding to infrequently transcribed genes will be extremely rare or not present at all. To have a reasonable chance of including clones corresponding to slowly transcribed genes, mammalian cDNA libraries must contain 10^6–10^7 individual recombinant clones.

The Polymerase Chain Reaction Amplifies a Specific DNA Sequence from a Complex Mixture

If the nucleotide sequences at the ends of a particular DNA region are known, the intervening fragment can be amplified directly by the **polymerase chain reaction (PCR)**. Here we describe the basic PCR technique and three situations in which it is used.

The PCR depends on the ability to alternately denature (melt) double-stranded DNA molecules and hybridize complementary single strands in a controlled fashion. As outlined in Figure 6-18, a typical PCR procedure begins with heat-denaturation of a DNA sample into single strands at 95 °C. Next two synthetic oligonucleotides complementary to the 3′ ends of the DNA segment of interest (the target sequence) are added in great excess to the

denatured DNA, and the temperature is lowered to 50–60 °C. These specific oligonucleotides, which are present at a very high concentration, hybridize to their complementary sequences in the DNA sample, whereas the long strands of the sample DNA remain apart because of their comparatively low concentration. The hybridized oligonucleotides then serve as primers for DNA chain synthesis in the presence of deoxynucleotides (dNTPs) and a temperature-resistant DNA polymerase such as that from *Thermus aquaticus*, a bacterium that lives in hot springs. This enzyme, called *Taq polymerase*, can remain active even after being heated to 95 °C and can extend the primers at temperatures up to 72 °C. When synthesis is complete, the whole mixture is reheated to 95 °C to denature the newly formed DNA duplexes. After the temperature is lowered again, another cycle of synthesis takes place because excess primer is still present. Repeated cycles of denaturation (heating) followed by hybridization and synthesis (cooling) quickly amplify the sequence of interest. At each cycle, the number of copies of the sequence between the primer sites is doubled; therefore, the target sequence increases exponentially — about 1-million-fold after 20 cycles — whereas all other sequences in the original DNA sample remain unamplified.



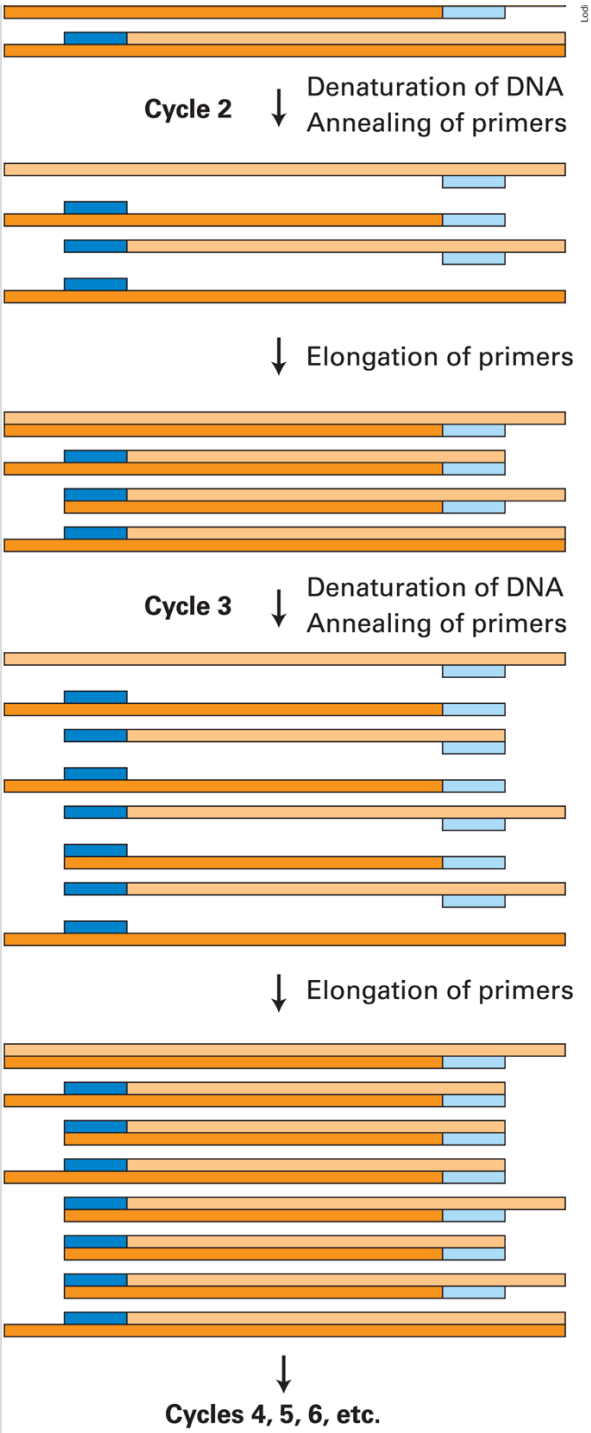


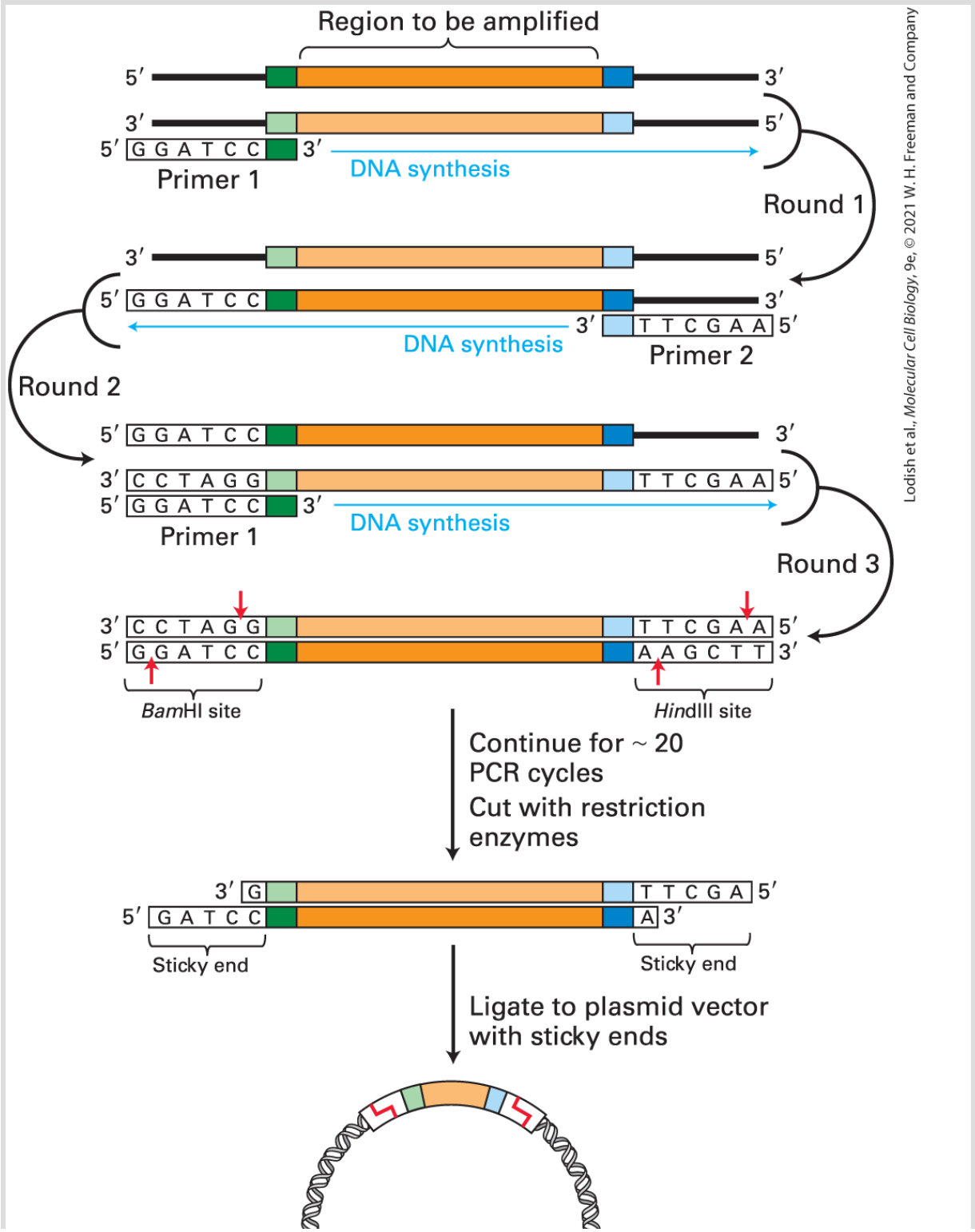
FIGURE 6-18 The polymerase chain reaction (PCR) is widely used to amplify DNA regions with known flanking sequences. To amplify a specific region of DNA, an investigator chemically synthesizes two different oligonucleotide primers complementary to sequences of approximately 18 bases flanking the region of interest (shown here as light blue and dark blue bars). The complete reaction is composed of a complex mixture of double-stranded

DNA (usually genomic DNA containing the target sequence of interest), a stoichiometric excess of both primers, the four dNTPs, and a heat-stable DNA polymerase known as *Taq* polymerase. During each PCR cycle, the reaction mixture is first heated to separate the strands and then cooled to allow the primers to bind to complementary sequences flanking the region to be amplified. *Taq* polymerase then extends each primer from its $3'$ end, generating newly synthesized strands that extend in the $3'$ direction to the $5'$ end of the template strand. During the third cycle, two double-stranded DNA molecules are generated equal in length to the sequence of the region to be amplified. In each successive cycle, the target sequence, which anneals to the primers, is duplicated, so it eventually vastly outnumbers all other DNA sequences in the reaction mixture. Successive PCR cycles can be automated by cycling the reaction at timed intervals between a high temperature for DNA melting and a lower temperature for the annealing and elongation parts of the cycle. A reaction that cycles 20 times will amplify the specific target sequence 1-million-fold.

Direct Isolation of a Specific Segment of Genomic DNA

For organisms in which all or most of the genome has been sequenced, PCR amplification starting with the total genomic DNA is often the easiest way to obtain a specific DNA region of interest for cloning. In this application of the PCR, the two oligonucleotide primers are designed to hybridize to sequences flanking the genomic region of interest and to include sequences that are recognized by specific restriction enzymes (Figure 6-19). For an oligonucleotide to be useful as a PCR primer, it must be long enough for its sequence to occur uniquely in the genome. For most purposes, this condition is satisfied by oligonucleotides containing about 20 nucleotides. Any given 20-nucleotide sequence occurs once in every $4^{20}(\sim 10^{12})$ nucleotides by chance, so it is usually possible to

identify two specific 20-nucleotide sequences flanking the target sequence that each occur only once in the genome.



EXPERIMENTAL FIGURE 6-19 A specific target sequence in genomic DNA can be amplified by PCR for use in cloning. Each primer for PCR is complementary to one end of the target sequence and includes the recognition site for a restriction enzyme that does not have a recognition site within the target region. In this example, primer 1 contains a *Bam*HI recognition sequence, whereas primer 2 contains a *Hin*dIII recognition sequence. (Note that for clarity, in any round, amplification of only one of the two strands — the one in brackets — is shown.) After amplification, the target segments are treated with appropriate restriction enzymes, generating fragments with sticky ends. These fragments can be incorporated into complementary plasmid vectors and cloned in *E. coli* by the usual procedure (see Figure 6-14).

After amplification of the target sequence for about 20 PCR cycles, the primers are cleaved with the appropriate restriction enzymes, producing fragments with sticky ends containing the target sequence. Each fragment is then ligated to a plasmid vector whose polylinker has been cleaved by the same restriction enzymes. The resulting recombinant plasmids, all carrying the identical genomic DNA segment, can then be cloned in *E. coli* cells. With certain refinements of the PCR, even DNA segments greater than 10 kb in length can be amplified and cloned in this way.

Note that this method does not require the cloning of large numbers of restriction fragments derived from genomic DNA and their subsequent screening to identify the specific fragment of interest. In effect, the PCR method inverts this traditional approach and so

avoids its most tedious aspects. The PCR method is useful for isolating gene sequences to be manipulated in a variety of useful ways described later. In addition, the PCR method can be used to isolate gene sequences from mutant organisms to determine how they differ from the wild type.

A variation on the PCR method is used to amplify a specific cDNA sequence made from cellular mRNAs. This method, known as *reverse transcriptase–PCR (RT-PCR)*, begins with the same procedure described previously for isolation of cDNA from a collection of cellular mRNAs (see Figure 6-17). Typically, an oligo-dT primer, which hybridizes to the 3′ poly(A) tail of the mRNA, is used as the primer for the first strand of cDNA synthesis by reverse transcriptase. A specific cDNA can then be isolated from this complex mixture of cDNAs by PCR amplification using two oligonucleotide primers designed to match sequences at the 5′ and 3′ ends of the corresponding mRNA. As described previously, these primers can be designed to include restriction sites that make it possible to insert the amplified cDNA into a suitable plasmid vector.

RT-PCR can be performed so that the starting amount of a particular cellular mRNA can be determined accurately. To carry out quantitative RT-PCR, the amount of double-stranded DNA sequence produced by each amplification cycle is determined as the amplification of a particular mRNA sequence proceeds. By extrapolation from these amounts, an estimate of the starting

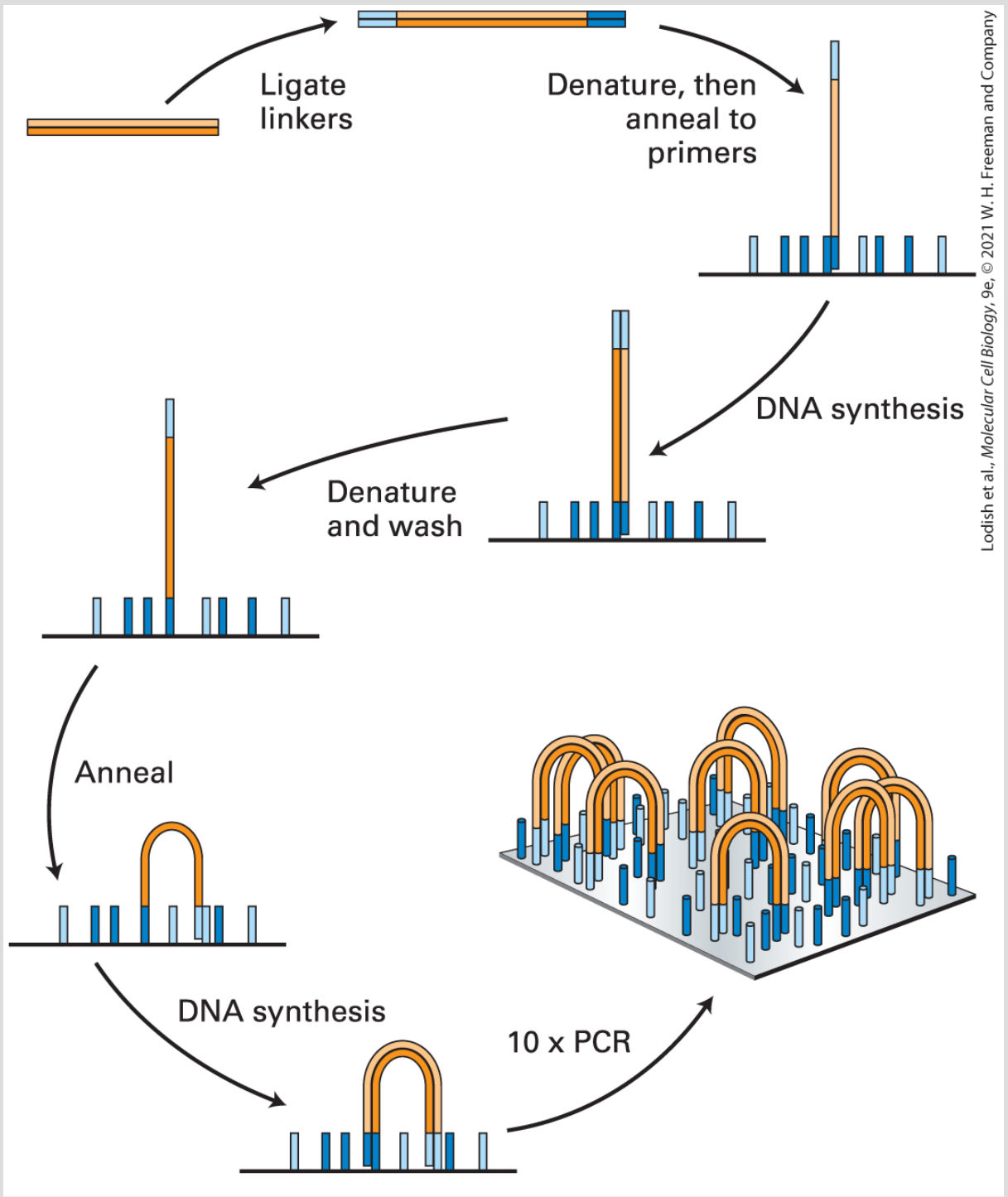
amount of the mRNA sequence can be obtained. Such quantitative RT-PCR analyses carried out on tissues or whole organisms using primers targeted to genes of interest provide one of the most accurate means to follow changes in gene expression.

Cloned DNA Molecules Can Be Sequenced Rapidly by Methods Based on PCR

A cloned DNA fragment is only completely characterized when its nucleotide sequence is known. The technology used to determine the sequence of a DNA segment represents one of the most rapidly developing fields in molecular biology. In the 1970s, Frederick Sanger and his colleagues developed the chain-termination procedure, which served as the basis for most DNA-sequencing methods for the next 30 years. This method is based on the selective incorporation of chain-terminating dideoxynucleotides into the daughter strands synthesized from the DNA fragment to be sequenced. Daughter-strand lengths from reactions containing each one of the four possible nucleotides provide the readout. Separation of the truncated daughter strands by gel electrophoresis, which can resolve strands that differ in length by a single nucleotide, can then reveal the lengths of all strands. From these collections of strands of different lengths, the nucleotide sequence of the original DNA fragment can be established. The Sanger method has undergone

many refinements and can now be fully automated, but because each DNA fragment of a few hundred nucleotides requires a separate, individual sequencing reaction, the overall rate at which new DNA sequences can be produced by this method is limited by the total number of reactions that can be performed at one time.

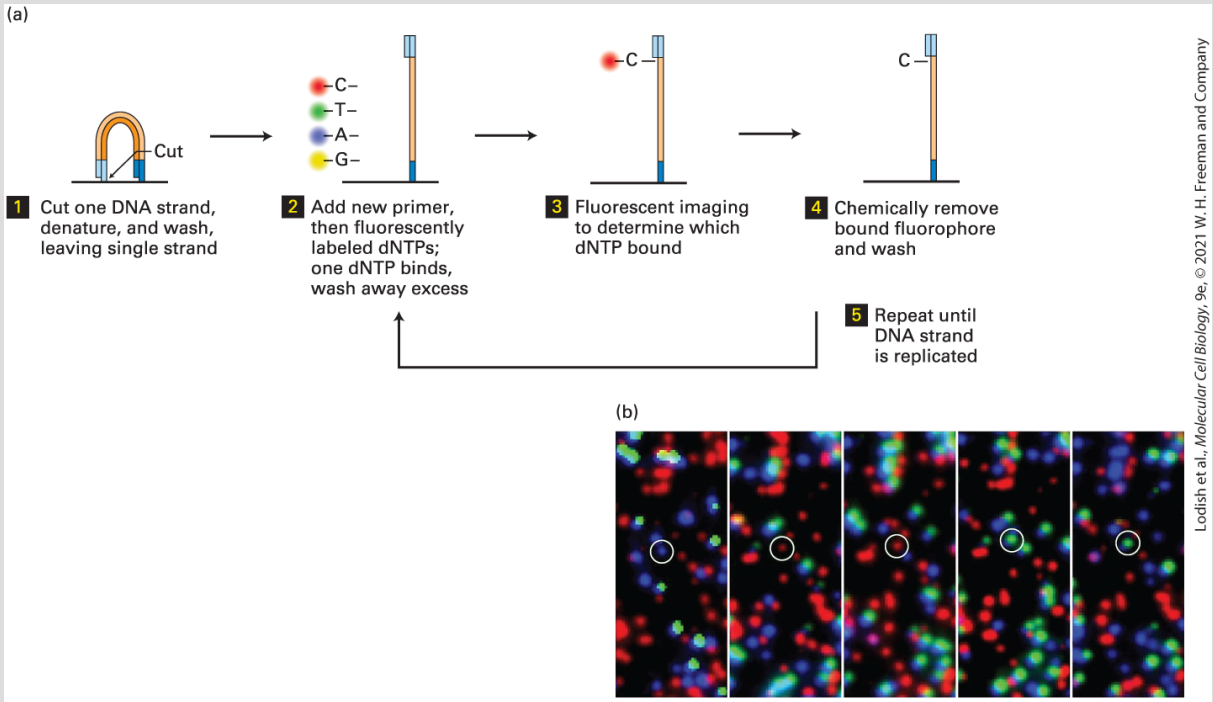
A breakthrough in sequencing technology occurred when methods were devised to allow a single sequencing instrument to carry out millions of sequencing reactions simultaneously. These sequencing methods localize the millions of reactions per square centimeter in tiny clusters on the surface of a solid substratum. In 2007, when these so-called *next generation* sequencers became commercially available, the capacity for new sequence production increased enormously, and since then, because of improvements in the technology, it has been further increasing at an amazing pace — doubling every few months. In one popular sequencing method, billions of different DNA fragments to be sequenced are prepared by ligating double-stranded linkers to their ends (Figure 6-20). Next the DNA fragments are amplified by PCR using primers that match the linker sequences. This reaction differs from the standard PCR amplification shown in Figure 6-19 in that the primers used are covalently attached to a solid substratum. Thus as the PCR amplification proceeds, one end of each daughter DNA strand is covalently linked to the substratum, and at the end of the amplification about a thousand identical PCR products are linked to the surface in a tight cluster.



EXPERIMENTAL FIGURE 6-20 Generation of clusters of identical DNA fragments attached to a solid support. A large collection of DNA fragments to be sequenced is ligated to double-stranded linkers, which become attached to each end of each fragment. The DNA is then amplified by PCR using primers matching the sequences of the linkers and that are covalently attached to a solid substratum. Ten cycles of amplification yield about a

thousand identical copies of each DNA fragment localized in a small cluster and attached at both ends to the solid substratum. These reactions are optimized to produce as many as 10^9 discrete, nonoverlapping clusters that are ready to be sequenced.

These clusters can then be sequenced by using a special microscope to image fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) as they are incorporated by DNA polymerase one at a time into a growing DNA chain (Figure 6-21). First, one strand is cut and washed out, leaving a single-stranded DNA template. Then sequencing is carried out on the thousand or so identical templates in each cluster, one nucleotide at a time. All four dNTPs are fluorescently labeled and added to the sequencing reaction. After they are allowed to anneal, the substratum is imaged and the color of each cluster is recorded. Next the fluorescent tag is chemically removed and a new dNTP is allowed to bind. This cycle is repeated about a hundred times, resulting in billions of hundred-nucleotide-long sequences. The entire procedure may take about one day and yield some 10^{11} bases of sequence information.

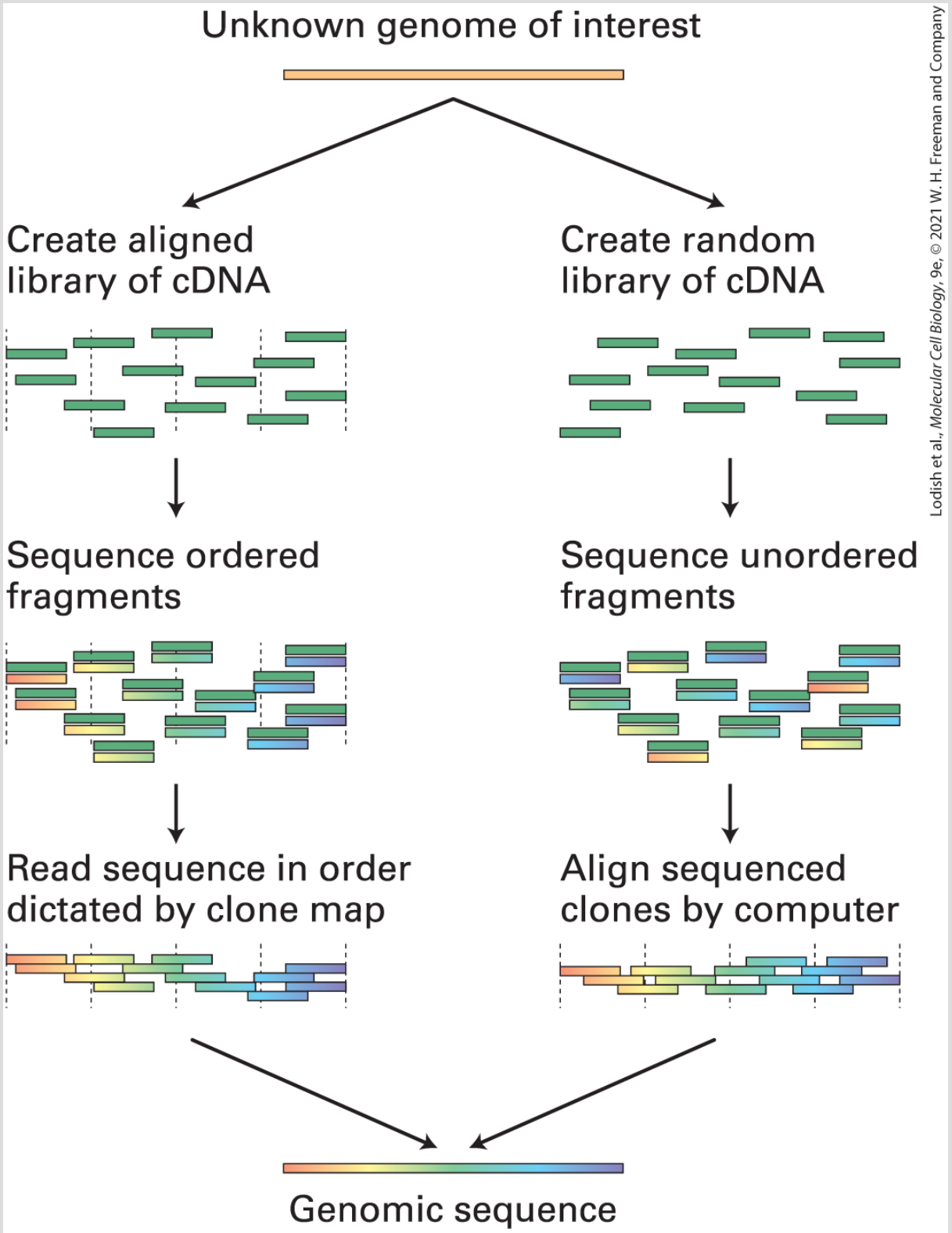


Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 6-21 Using fluorescently-tagged precursors for high-throughput sequence determination. (a) The reaction begins with the cleaving of one strand of the amplified, clustered DNA (see Figure 6-20). After denaturation, a single DNA strand remains attached to the substratum. A synthetic oligodeoxynucleotide is used as the primer for the polymerization reaction, which also contains dNTPs, each fluorescently tagged with a different color. The fluorescent tag is designed to block the $3'$ OH group on the dNTP so that once the fluorescent dNTP has been incorporated, further elongation is not possible. Because DNA polymerase will incorporate the same fluorescently labeled dNTP into each of the thousand or so identical DNA copies in a cluster, the entire cluster will be uniformly labeled with the same fluorescent color, which can be imaged in a special microscope. (b) Five images from the same field of view, each corresponding to an individual cycle of dNTP addition. Each colored dot represents a cluster of identical DNA fragments. After each image is made, the fluorescent tags are removed by a chemical reaction that leaves a new primer terminus available for the next cycle of dNTP addition. As can be seen for the circled colored dot, the color changes in each reaction cycle according to which nucleotide is added to the DNA fragments (from the five panels the sequence of the circled fragment would be read: A-C-C-T-T). A typical sequencing reaction may carry out a hundred cycles of NTP addition, allowing a hundred bases of sequence for each cluster to be determined. Thus a total sequencing reaction of this type may generate as much as 3×10^{11} bases of sequence information in about two days.

[Part (b) A. Loehr and A. W. Zaranek for the Harvard Personal Genome Project.]

In order to sequence a long, continuous region of genomic DNA, or even the entire genome of an organism, researchers usually employ one of the strategies outlined in Figure 6-22. The first method requires the isolation of a collection of cloned DNA fragments whose sequences overlap. Once the sequence of one of these fragments is determined, oligonucleotides based on that sequence can be chemically synthesized for use as primers in sequencing the adjacent overlapping fragments. In this way, the sequence of a long stretch of DNA is determined incrementally by sequencing of the overlapping, cloned DNA fragments that compose it. A second method, which is called *whole genome shotgun sequencing*, bypasses the time-consuming step of isolating an ordered collection of DNA segments that span the genome. This method involves simply sequencing random clones from a genomic library. A sufficient number of clones are chosen for sequencing so that on average, each segment of the genome is sequenced about 10 times. This degree of coverage ensures that each segment of the genome is sequenced more than once. The entire genomic sequence is then assembled using a computer algorithm that aligns all the segments using their regions of overlap. Whole genome shotgun sequencing is the fastest and most cost-effective method for sequencing long stretches of DNA, and most genomes, including the human genome, have been sequenced by this method.



EXPERIMENTAL FIGURE 6-22 Two strategies for assembling whole genome sequences. One method (*left*) depends on isolating and assembling a set of cloned DNA fragments that

span the genome. This can be done by matching cloned fragments by hybridization or by alignment of restriction-site maps. The DNA sequence of the ordered clones can then be assembled into a complete genomic sequence. The alternative method (*right*) depends on the relative ease of automated DNA sequencing and bypasses the laborious step of ordering a DNA library. By sequencing enough random clones from the library so that each segment of the genome is represented from 3 to 10 times, it is possible to reconstruct the genomic sequence by computer alignment of the very large number of sequence fragments.

KEY CONCEPTS OF SECTION 6.2

DNA Cloning and Characterization

- In DNA cloning, recombinant DNA molecules are formed in vitro by inserting DNA fragments into vector DNA molecules. The recombinant DNA molecules are then introduced into host cells, where they replicate, producing large numbers of recombinant DNA molecules.
- Restriction enzymes (endonucleases) typically cut DNA at specific 4–8-bp palindromic sequences, producing defined fragments that often have self-complementary, single-stranded tails (sticky ends).
- Two restriction fragments with complementary ends can be joined with DNA ligase to form a recombinant DNA molecule (see Figure 6-12).
- *E. coli* cloning vectors are small circular DNA molecules (plasmids) that include three functional regions: a replication origin, a selectable marker gene, and a site where a DNA fragment can be inserted. Transformed cells carrying a vector grow into colonies on the selection medium (see Figure 6-14).
- A genomic library is a set of clones carrying restriction fragments produced by cleavage of the entire genome.
- Shuttle vectors that can replicate in both yeast and *E. coli* can be used to construct a yeast genomic library. Specific genes can be isolated by their ability to complement the corresponding mutant genes in yeast cells (see Figure 6-16).
- In cDNA cloning, expressed mRNAs are reverse-transcribed into complementary DNAs, or cDNAs. A cDNA library is a set of cDNA clones prepared from the mRNAs isolated from a particular type of cell or tissue (see Figure 6-17).
- The polymerase chain reaction (PCR) permits exponential amplification of a specific segment of DNA from a single initial template DNA molecule if the sequence flanking the DNA region to be amplified is known (see Figure 6-19).

- PCR is a highly versatile method that can be programmed to amplify a specific genomic DNA sequence, a cDNA, or a sequence at the junction between a DNA transposon and flanking chromosomal sequences.
- DNA fragments up to about 100 nucleotides long can be sequenced by generating clusters of identical fragments by PCR and imaging fluorescently labeled nucleotide precursors incorporated by DNA polymerase (see Figures 6-20 and 6-21).
- Whole genome sequences can be assembled from the sequences of a large number of overlapping clones from a genomic library (see Figure 6-22).

6.3 Using Sequence Information to Identify Genes and Deduce Their Function

By using automated DNA-sequencing techniques and computer algorithms to piece together sequence data, researchers have determined vast amounts of DNA sequence, including nearly the entire genomic sequence of humans and of many experimental organisms. By now, the genomic sequences have been completely determined for all of the major model organisms and thousands of viruses, bacteria, archea, and eukaryotes including representatives of all of the 35 or so metazoan phyla.

The genome sequence of an organism contains the information that specifies the sequence of all of its protein and RNA components. In addition, regulatory sequences in the genome specify when and where each of these proteins or RNA molecules is expressed. Thus it should be possible, at least in principle, to learn almost all there is to know about an organism — what it is, how it functions, and how it evolved — from a properly informed reading of its genomic sequence. Useful biological information is extracted from sequence data, largely by computers that search for meaningful patterns in sequence data by comparing one sequence to another. The effort to extract biological information from genomes has led to the

emergence of an enormously powerful new field of biology: *bioinformatics*.

Although this field depends upon the highly technical application of advanced computer algorithms, the fundamental principle underlying all of bioinformatics is not complicated. Bioinformatic reasoning is based on the understanding that life on earth arose only once and all genome sequences, either existing or extinct, are derived from this single common ancestor through a continual process of random mutation and natural selection. Thus, any substantial similarity between two sequences from different species can be attributed to one of two causes. Either the organisms are closely related such that not enough time has passed for random differences to have accumulated by mutation, or the sequence is constrained to perform some important function and so mutational variations have been systematically weeded out by natural selection. This latter case is known as **conservation**, and we would say that a sequence is conserved if it has remained substantially unchanged in organisms that are evolutionarily so far apart that most sequences have lost recognizable similarity due to random mutations. Most bioinformatics reasoning rests on the following overarching concept: *the more a sequence shows conservation among distantly related organisms, the more important its function is likely to be*.

In this section, we first show how protein coding genes are identified in genomic sequences. Next we examine how functional

and evolutionary relationships between proteins are studied. Finally, we explore the relationship between the complexity of an organism and the number of protein-coding genes that it contains.

Most Genes Can Be Readily Identified Within Genomic DNA Sequences

The complete genomic sequence of an organism contains within it the information needed to deduce the sequence of every protein made by the cells of that organism. Different strategies are used to deduce protein sequences for prokaryotes and eukaryotes. In the genomes of prokaryotes, the protein-coding sequences are continuous reading frames that are not broken up by introns. Thus for prokaryotes the vast majority of protein-coding sequences can be found simply by scanning the genomic sequence for **open reading frames (ORFs)** of significant length. An ORF is defined as a continuous sequence of triplet codons that does not contain a stop codon.

ORFs that are never translated into amino acids may appear by random chance. How many codons must an ORF contain for us to be sure that it is part of a gene? This number can be calculated by utilizing the probability that a run of non-stop codons has occurred by chance. The probability of a non-stop codon in a random sequence is 61/64, which is simply the fraction of the 64 possible

triplet codons that are not stop codons. Thus the probability of a random sequence containing an ORF of greater than n codons is $(61/64)^n$, which is the joint probability of *not* encountering a stop codon n times in a row. Taking the average length of a prokaryotic protein to be about 300 amino acids, the calculated probability of an ORF greater than this length *arising by chance* is only 5.6×10^{-7}, and it is unlikely that even one ORF of this length would occur in an entire bacterial genomic sequence of 4×10^6 base pairs. In practice, prokaryotic genes can be reliably found by simply searching for ORFs greater than about 100 codons.

ORF analysis correctly identifies more than 90 percent of the genes in bacteria and in eukaryotes such as *S. cerevisiae* that have few introns. Some of the very shortest genes, however, are missed by this method, and occasionally long open reading frames that are not actually genes do arise by chance. Both types of misassignments can be corrected by more sophisticated analysis of the sequence and by genetic tests for gene function. Of the genes identified in this manner, about half were already known by some functional criterion such as mutant phenotype. Some of the proteins encoded by the remaining putative genes identified by ORF analysis have been assigned functions based on their sequence similarity to known proteins in other organisms.

More sophisticated algorithms are required to identify genes in organisms with a more complex genome structure. Scanning for

ORFs is a poor method for finding genes in higher eukaryotes because the coding sequences of most genes are broken up by multiple noncoding introns (humans have an average of 9 introns per gene), and the relatively short ORFs derived from exons are not easily distinguished from the many short ORFs that arise in the genome by chance. In a typical gene-sized segment of genomic DNA, more than a thousand fortuitous ORFs greater in length than the average exon would be expected. A much more reliable way to find eukaryotic genes is to consider the possible encoded amino acid sequences, not just the presence of an ORF. One way to do this is to have a computer generate a list of all of the thousands of peptide sequences produced by hypothetical translation of all ORFs of any length with a specific genome sequence. Each of these peptide sequences would then be compared to the entire database of protein sequences to look for significant matches to a known protein family. Finding multiple short ORFs from the same region of DNA that show similarity to sequential segments of a single known protein would provide strong evidence that the set of ORFs is the set of exons for a single human gene.

The best gene-finding algorithms combine all the available data that might suggest the presence of a gene at a particular genomic site. Relevant data include alignment of the sequence under examination to a full-length cDNA sequence; alignment to a partial cDNA sequence, generally 200–400 bp in length, known as an *expressed sequence tag (EST)*; fitting to models for exon, intron, and splice-site

sequences; and similarity to gene sequences from other organisms. Using these computer-based bioinformatic methods, computational biologists have identified approximately 21,000 protein-coding genes in the human genome.

A particularly powerful method for identifying human genes is to compare the human genomic sequence with that of other mammals, particularly the mouse. Humans and mice are sufficiently related to have most genes in common, although largely nonfunctional DNA sequences, such as introns and the regions between genes, tend to be very different because these sequences are not under strong selective pressure. In addition to conserved coding sequences of exons, the segments of the human and mouse genome that are important for gene structure and function, such as intron splice sites or transcription-control regions, can be identified because these sequences tend to be conserved.

Bioinformatic Principles Can Be Used to Deduce the Likely Functional Consequences of Mutations

Once a gene has been identified, comparison of a mutant gene sequence with that of wild type can be used to deduce the likely effect of the mutation on the gene product. Mutations that alter the coding sequence of a gene usually provide the most insight into

function of the expressed protein. The most severe types of mutations introduce a premature stop codon (*nonsense* mutations) or change the reading frame of a gene by the addition or deletion of bases (*frameshift* mutations). Nonsense mutations and frameshift mutations almost always cause a complete loss of gene function and are therefore often referred to as *null alleles. Point mutations* that result in substitution of one amino acid for another (*missense* mutations) can have dramatic effects on gene function or they can have no effect at all, and the principle of conservation of the functionally important parts of a gene can contribute to an educated guess as to the severity of the effect of a missense mutation. For example, a missense mutation that changed a highly conserved amino acid residue to a chemically dissimilar residue would likely have a large effect on protein function, whereas a change in a non-conserved position would likely have a mild effect or no effect at all. A point mutation that changes a codon to a different codon that codes for the same amino acid will almost never have an effect on gene function and is therefore considered to be a *silent* mutation.

In higher eukaryotes, point mutations outside the coding sequence of genes rarely have dramatic effects on phenotype. This is not because promoter sequences and sequences at splice sites are not important for gene function, but rather that the function of these cis-acting sequences is usually distributed redundantly among multiple nucleotide residues such that no single point mutation will have a large effect. The majority of point mutations outside gene-

coding sequences that have observable effects on gene function interfere with pre-mRNA splicing, and these mutations often lie in highly conserved residues that are pivotal to selecting the proper pre-mRNA splice site.

The Function and Evolutionary Origins of Genes and Proteins Can Be Deduced from Their Sequence

As discussed in Chapter 3, proteins with similar functions often contain similar amino acid sequences that correspond to important functional domains in the three-dimensional structure of the proteins. By comparing the amino acid sequence of the protein encoded by a newly cloned gene with the sequences of proteins of known function, an investigator can look for sequence similarities that provide clues to the function of the encoded protein. Because of the degeneracy in the genetic code, related proteins invariably exhibit much more sequence conservation than the genes encoding them. For this reason, protein sequences, rather than the corresponding DNA sequences, are usually compared.

The most widely used computer program for this purpose is known as BLAST (*b*asic *l*ocal *a*lignment *s*earch *t*ool). The BLAST algorithm divides the protein sequence to be examined (known as the *query sequence*) into shorter segments and then searches the database for

significant matches to any of the stored sequences. The matching program assigns a high score to identically matched amino acids and a lower score to matches between amino acids that are related (e.g., hydrophobic, polar, positively charged, negatively charged) but not identical. When a significant match is found for a segment, the BLAST algorithm searches locally to extend the region of similarity. After searching is completed, the program ranks the matches between the query protein and various known proteins according to their *p-values*. This parameter is a measure of the probability of finding such a degree of similarity between two protein sequences by chance. The lower the *p*-value, the greater the sequence similarity between two sequences. A *p*-value less than about 10^{-3} is usually considered significant evidence that two proteins share a common ancestor. Many alternative computer programs have been developed that can detect relationships between proteins that are more distantly related to each other than can be detected by BLAST. The development of such methods is currently an active area of bioinformatics research.

To illustrate the power of this sequence-comparison approach, let's consider the human gene *NF1*. Mutations in *NF1* are associated with the inherited disease neurofibromatosis 1, in which multiple tumors develop in the peripheral nervous system, causing large protuberances in the skin. After a cDNA clone of *NF1* was isolated and sequenced, the deduced sequence of the NF1 protein was checked against all other protein sequences in GenBank. A region of

NF1 protein was discovered to have considerable homology to a portion of the yeast protein called *Ira* (Figure 6-23). Previous studies had shown that Ira is a GTPase-activating protein (GAP) that modulates the GTPase activity of the monomeric G protein called *Ras* (see Figure 3-35). Its homology with Ira suggested that NF1 also regulated Ras activity. As we examine in detail in Chapter 16, GAP and Ras proteins normally function to control cell replication and differentiation in response to signals from neighboring cells. Functional studies on the normal NF1 protein, obtained by expression of the cloned wild-type gene, showed that it did indeed regulate Ras activity. These findings suggest that patients with neurofibromatosis express a mutant NF1 protein in cells of the peripheral nervous system, leading to abnormally high signaling through Ras protein. The enhanced rate of signaling in turn leads to excessive cell division and formation of the tumors characteristic of the disease.

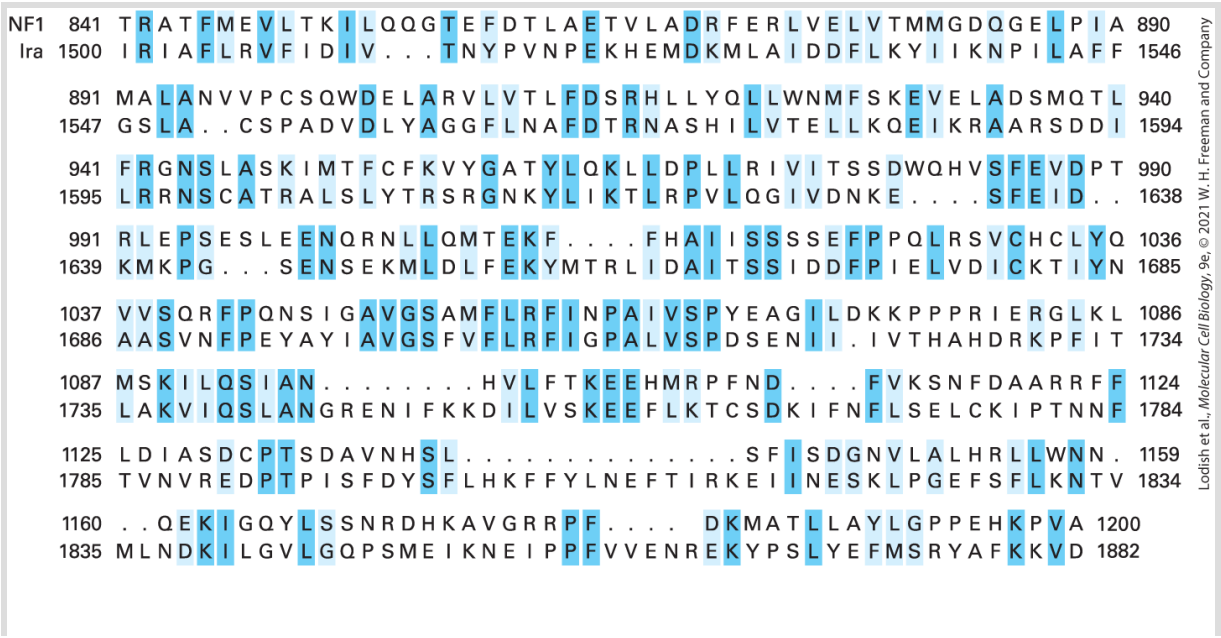


FIGURE 6-23 Comparison of the regions of human NF1 protein and *S. cerevisiae* Ira protein that show significant sequence similarity. The NF1 and the Ira sequences are shown on the top and bottom lines of each row, respectively, in the one-letter amino acid code (see Figure 2-14). Amino acids that are identical in the two proteins are highlighted in dark blue. Amino acids with chemically similar but nonidentical side chains are highlighted in light blue. Black dots indicate gaps in the upper and lower protein sequences, inserted by the algorithm in order to maximize the alignment of homologous amino acids. The BLAST *p*-value for these two sequences is 10^{-28}, indicating a high degree of similarity.

[Data from G. Xu et al., 1990, *Cell* **62**:599.]

Even when the BLAST algorithm finds no significant overall similarities, a query sequence may nevertheless share a short sequence with known proteins that is functionally important. Such short segments, referred to as **structural motifs**, recur in many different proteins. These recurring segments generally have similar functions. Several such motifs are described in Chapter 3 and illustrated in Figure 3-7. To search for these and other motifs in a new protein, researchers compare the query protein sequence with a database of known motif sequences.

Comparison of Related Sequences from Different Species Can Give Clues to Evolutionary Relationships Among Proteins

BLAST searches for related protein sequences may reveal that proteins belong to a protein family shared among related organisms. Consider, for example, the *tubulin* proteins, the basic subunits of microtubules, which are important components of the cytoskeleton (see Chapter 18). According to the simplified scheme in Figure 6-24a, the earliest eukaryotic cells are thought to have contained a single tubulin gene that was duplicated early in evolution. The two copies of the original tubulin gene subsequently diverged in sequence and became the ancestral versions of the α- and β-tubulin genes. As different species diverged from these early eukaryotic cells, each of these gene sequences further diverged, giving rise to the slightly different forms of α-tubulin and β-tubulin now found in each species.

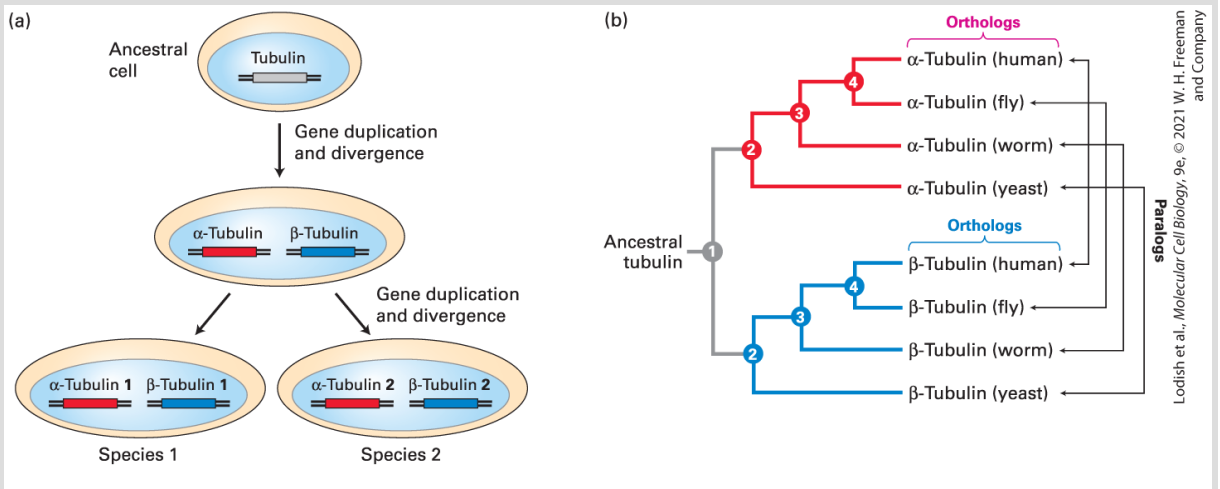


FIGURE 6-24 Generation of diverse tubulin sequences during the evolution of eukaryotes. (a) Probable mechanism giving rise to the tubulin genes found in existing species. It is possible to deduce that a gene-duplication event occurred before speciation because the α-tubulin sequences from different species (e.g., humans and yeast) are more alike than are the α-tubulin and β-tubulin sequences within a species. (b) A phylogenetic tree representing the relationship between the tubulin sequences. The branch points (nodes), indicated by

small numbers, represent common ancestral genes at the time that two sequences diverged. For example, node 1 represents the duplication event that gave rise to the α-tubulin and β-tubulin families, and node 2 represents the divergence of yeast from multicellular species. Braces and arrows indicate, respectively, the orthologous tubulin genes, which differ as a result of speciation, and the paralogous genes, which differ as a result of gene duplication. This diagram is simplified somewhat because flies, worms, and humans actually contain multiple α-tubulin and β-tubulin genes that arose from later gene duplication events.

All the different members of the tubulin family of genes (and proteins) are sufficiently similar in sequence to suggest that they arose from a common ancestral sequence. Thus all these sequences are considered to be *homologous*. More specifically, sequences that presumably diverged as a result of gene duplication (e.g., the α- and β-tubulin sequences) are described as *paralogous*. Sequences that arose because of speciation (e.g., the α-tubulin genes in different species) are described as *orthologous*. From the degree of sequence relatedness of the tubulins present in different organisms today, evolutionary relationships can be deduced, as illustrated in Figure 6-24b. Of the three types of sequence relationships, orthologous sequences are the most likely to share the same function.

Biological Complexity of an Organism Is Not Directly Related to the Number of Protein-Coding Genes in the Genome

The combination of genomic sequencing and gene-finding computer algorithms has yielded the complete inventory of protein-coding genes for a variety of organisms. Figure 6-25 shows the total number of protein-coding genes in several eukaryotic genomes that have been completely sequenced. The functions of about half the proteins encoded in these genomes are known or have been predicted on the basis of sequence comparisons. One of the surprising features of this comparison is that the number of protein-coding genes within different organisms does not seem proportional to our intuitive sense of their biological complexity. For example, the roundworm *C. elegans* apparently has more genes than the fruit fly *Drosophila*, which has a much more complex body plan and more complex behavior. And humans have only about 5 percent more protein-coding genes than *C. elegans*. When it first became apparent that humans appeared to have only a slightly more comlex genome than the simple roundworm, it was difficult to understand how such a small increase in the number of proteins could generate such a staggering difference in complexity.

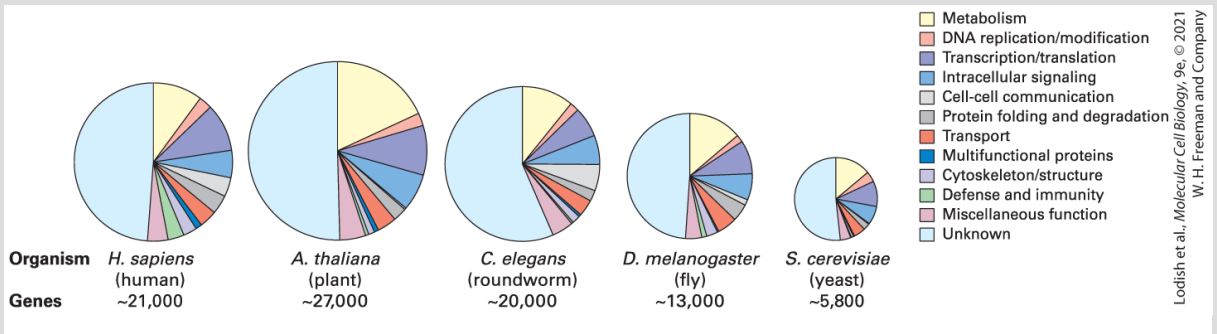


FIGURE 6-25 Comparison of the number and types of proteins encoded in the genomes of different eukaryotes. For each organism, the area of the entire pie chart represents the total number of protein-coding genes, all shown at roughly the same scale. In most cases, the

functions of the proteins encoded by about half the genes are still unknown (light blue). The functions of the remainder are known or have been predicted by sequence similarity to genes of known function.

[Data from ENCODE Project Consortium, 2012, *Nature* **489**:57; J. D. Hollister, 2014, *Chromosome Res.* **22**:103; L. W. Hillier et al., 2005, *Genome Res.* **15**:1651; FlyBase: FB2015_02 Release Notes, http://flybase.org/static_pages/docs/release_notes.html; Saccharomyces Genome Data Base 2015, http://www.yeastgenome.org/genomesnapshot.]

Clearly simple quantitative differences in the number of protein-coding genes in the genomes of different organisms are inadequate for explaining differences in biological complexity. However, several phenomena can generate more complexity in the expressed proteins of higher eukaryotes than is predicted from their genomes. First, a particular gene can yield multiple functional mRNAs, and therefore multiple proteins, through alternative splicing of a pre-mRNA (see Chapter 9). In humans, the mean number of alternatively spliced mRNAs expressed per gene is about six. Second, variations in the post-translational modification of many proteins may produce functional differences. Finally, stable modifications to chromatin structure that are not evident in gene sequences can have large effects on gene expression in particular cell types (see Chapter 9). These so-called epigenetic effects can greatly contribute to biological complexity. Evolution of the increasing biological complexity of multicellular organisms probably required increasingly complex regulation of cell replication and temporal and spatial regulation of gene expression

in the cells that make up the organisms, leading to increasing complexity of embryological development.

Researchers still have not determined the specific functions of many genes and proteins identified by analysis of genomic sequences. As they unravel the functions of individual proteins in different organisms and further detail their interactions with other proteins, the resulting advances will become immediately applicable to all homologous proteins in other organisms. No doubt, when the function of every protein is known, a more sophisticated understanding of the molecular basis of complex biological systems will emerge.

KEY CONCEPTS OF SECTION 6.3

Using Sequence Information to Identify Genes and Deduce Their Function

- All species are derived from a common ancestor and have evolved by random mutation and natural selection. Sequence patterns that are conserved among distantly related species are likely to be functionally important and to have been maintained by natural selection.
- A fundamental principle of bioinformatics is that the more highly conserved a sequence is, the more important its function is likely to be.
- For organisms such as bacteria and yeast whose genes are not interrupted by introns, protein-coding genes can be identified by computer searching of the entire bacterial and yeast genomic sequences for open reading frames (ORFs).
- For most higher eukaryotes, which have relatively short coding exons separated by introns, methods other than ORF identification are required to reliably identify genes. These methods include reconstructing spliced genes from cDNA sequences, searching for similarities between hypothetical translation of ORFs and known protein sequences, and identifying conserved gene features such as intron-splice sites.

- The function of a protein that has not been isolated can often be predicted on the basis of similarity of its amino acid sequence (the query sequence) to the sequences of proteins of known function.
- A computer algorithm known as BLAST rapidly searches databases of known protein sequences to find those with significant similarity to a query protein.
- Proteins with common functional motifs, which can often be quite short, may not be identified in a typical BLAST search. Such short sequences may be located by searches of structural motif databases.
- A protein family comprises multiple proteins all derived from the same ancestral protein. The genes encoding these proteins, which constitute the corresponding gene family, arose by an initial gene-duplication event and subsequent divergence during speciation (see Figure 6-24).
- Related genes and their encoded proteins expressed in one organism that derive from a gene duplication event are paralogous, such as the α- and β-tubulin proteins that combine to form an αβ dimer building block of the microtubule. Those that derive from mutations that accumulated during speciation are orthologous. Proteins that are orthologous usually have the same function in different organisms.
- Analysis of the complete genomic sequences of several different organisms indicates that biological complexity is not directly related to the number of protein-coding genes (see Figure 6-25).

6.4 Locating and Identifying Genes That Specify Human Traits

The function of human proteins and human genes is of primary interest to cell biologists. In many cases, cell biologists will study human genes that are orthologs of genes of known cell biological function that have been studied in model organisms because these orthologs are often genes with important cell biological functions in humans as well. However, in certain key cases human genes with important functions that were not identified in model organisms have been identified solely because mutations in these genes have an observable effect on human biology. Many of these genes have been identified because they cause, contribute to, or protect against human disease.

Table 6-1 lists several of the most commonly occurring inherited diseases. Although a "disease" gene may result from a new mutation that arose in the preceding generation, most cases of inherited diseases are caused by preexisting mutant alleles that have been passed from one generation to the next for many generations.

TABLE 6-1 • Common Inherited Human Diseases

Disease	Molecular and Cellular Defect	Incidence
Autosomal Recessive		
Sickle-cell disease	Abnormal hemoglobin causes deformation of	1/625 of

	red blood cells, which can become lodged in capillaries; also confers resistance to malaria.	sub-Saharan African origin
Cystic fibrosis	Defective chloride channel (CFTR) in epithelial cells leads to excessive mucus in lungs.	1/2500 of European origin
Phenylketonuria (PKU)	Defective enzyme in phenylalanine metabolism (tyrosine hydroxylase) results in excess phenylalanine leading to mental retardation, unless restricted by diet.	1/10,000 of European origin
Tay-Sachs disease	Defective hexosaminidase enzyme leads to accumulation of excess sphingolipids in the lysosomes of neurons, impairing neural development.	1/1000 Ashkenazi Jews
Autosomal Dominant		
Huntington's disease	Defective neural protein (huntingtin) may assemble into aggregates, causing damage to neural tissue.	1/10,000 of European origin
Hypercholesterolemia	Defective LDL receptor leads to excessive cholesterol in blood and early heart attacks.	1/250 French Canadians
X-Linked Recessive		
Duchenne muscular dystrophy (DMD)	Defective cytoskeletal protein (dystrophin) leads to impaired muscle function.	1/3500 males
Hemophilia A	Defective blood clotting factor VIII leads to uncontrolled bleeding.	1/5000 males

The typical first step in deciphering the underlying cause of any inherited human disease is to identify the affected gene and its encoded protein. Historically, researchers have used whatever phenotypic clues might be relevant to make guesses about the molecular basis of inherited diseases. An early example of successful guesswork was the hypothesis that sickle-cell disease, known to be a disease of blood cells, might be caused by defective hemoglobin. This idea led to the identification of a specific amino acid substitution in hemoglobin that causes the defective hemoglobin molecules to polymerize. These polymers in turn cause the sickle-like deformation of red blood cells in individuals who have inherited two copies of the Hb^s allele for sickle-cell hemoglobin.

Most often, however, the genes responsible for inherited diseases must be found without any prior knowledge or reasonable hypotheses about the identity of the affected gene or its encoded protein. In this section, we see how human geneticists can find the gene responsible for an inherited disease by following the segregation of the disease in families. That is, geneticists trace which family members had the disease and which did not over the course of a few generations. They then look for genetic markers present in the DNA sequence that appear only in the family members with the disease. In other words, the segregation of the disease within the family can be correlated with the segregation of

many other genetic markers. The location of the genetic markers, already known, eventually leads to identification of the chromosomal position of the affected gene. Using this information, along with knowledge of the sequence of the human genome, geneticists can ultimately pinpoint the affected gene and the disease-causing mutations.

Monogenic Diseases Show One of Three Patterns of Inheritance

We have already seen that one of the most useful aspects of genetic analysis in model organisms is that the phenotype of a mutation along with knowledge of whether the mutation is genetically recessive or dominant enables biologists to deduce the function of the normal protein. Of course in humans the same types of test crosses that are used to determine dominance or recessiveness in model organisms cannot be applied. Nevertheless, researchers can determine whether a human trait is dominant or recessive by examining the pattern of inheritance of a trait in families or human populations. Thus determining the mode of inheritance along with identifying the gene or genes that are affected are the primary objectives of human genetic studies.

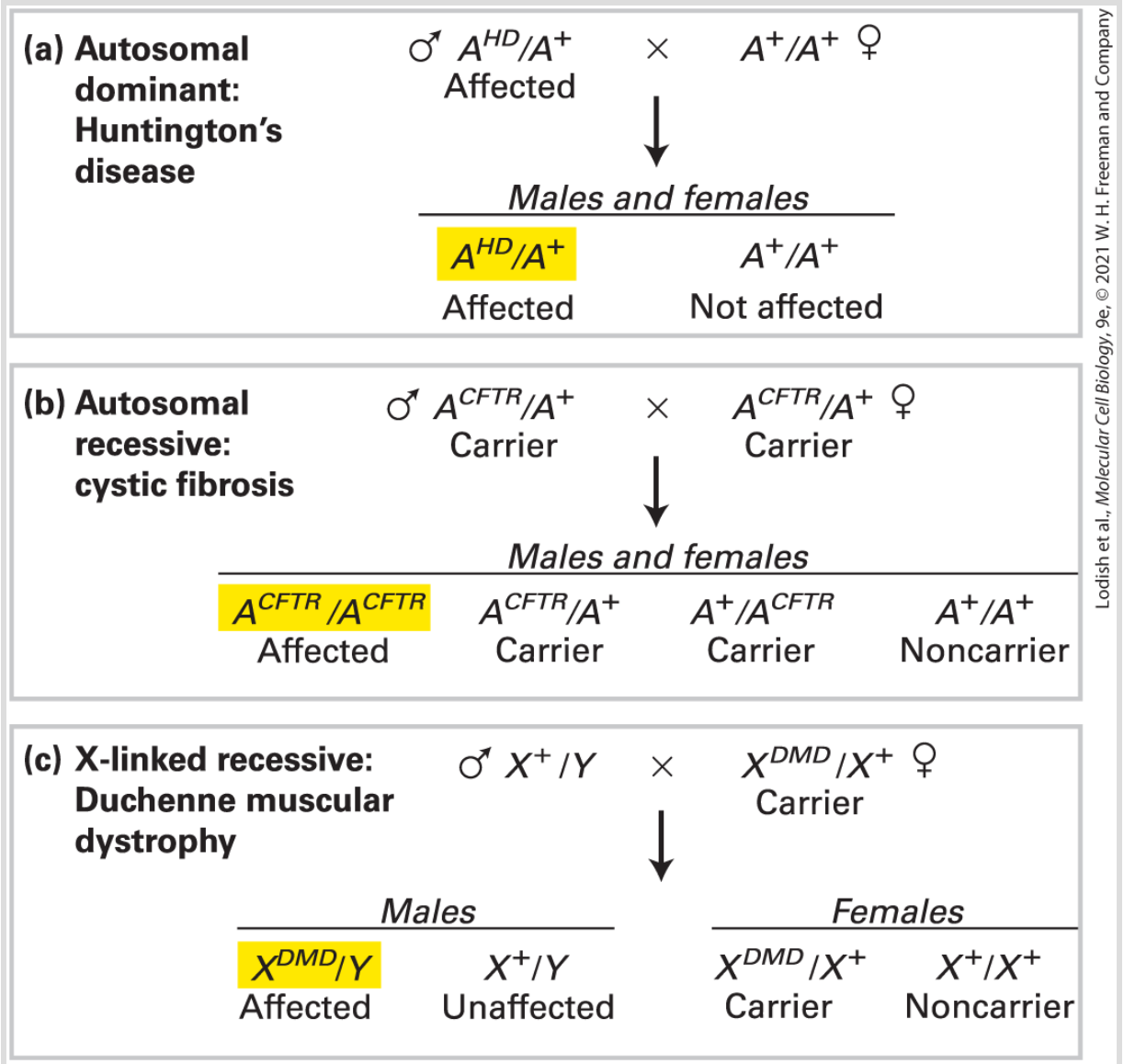
Human genetic diseases that result from mutation in one specific gene, referred to as *monogenic diseases*, display different inheritance

patterns depending on whether the alleles that cause them are dominant or recessive and whether they are located on an autosome or a sex chromosome. One characteristic pattern is that exhibited by a dominant allele in an autosome (i.e., one of the 22 human chromosomes that is not a sex chromosome). Because an *autosomal dominant* allele is expressed in the heterozygote, usually at least one of the parents of an affected individual will also have the disease. For example, in Huntington's disease, which is autosomal dominant, if either parent carries a mutant *HD* allele, each of his or her children (regardless of sex) has a 50 percent chance of inheriting the mutant allele and being affected (Figure 6-27a).

What are the molecular mechanisms underlying dominance? Huntington's disease, a neural degenerative disease, is caused by an expansion of the number of triplet CAG repeats in the coding sequence of a neuronal protein HTT. The mutant HTT protein that has an expanded number of glutamine residues has been shown to form protein aggregates. The fact that the mutant *HD* allele is inherited in a dominant fashion strongly implies that the symptoms of Huntington's disease result from a *gain* in propensity of the mutant HTT protein to form aggregates and not because of a *loss* of the normal HTT protein function. Diseases such as Huntington's disease caused by dominant alleles often strike in middle to late life, after reproductive age. If this were not the case, natural selection would have eliminated these alleles during human evolution.

A recessive allele in an autosome exhibits quite a different segregation pattern. Both parents must be heterozygous *carriers* of an autosomal recessive allele in order for their children to be at risk of the disease. Each child of heterozygous parents has a 25 percent chance of receiving both recessive alleles and thus being affected; a 50 percent chance of receiving one normal and one mutant allele and thus being a carrier; and a 25 percent chance of receiving two normal alleles. Related individuals (e.g., first or second cousins) have a relatively high probability of being carriers for the same recessive alleles. Thus children born to related parents are much more likely than those born to unrelated parents to be homozygous for, and therefore affected by, a rare autosomal recessive disorder.

A clear example of an autosomal recessive disease is cystic fibrosis, which results from a defective chloride-channel gene known as *CFTR* (Figure 6-26b). The presence of this defect disturbs the balance of ions across the epithelial cell layers of the lungs and intestine. Because cystic fibrosis is a recessive disease, we may deduce that the effects of the disease are due to insufficient activity of the chloride channels in the epithelial cells.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 6-26 Three common inheritance patterns for human monogenic diseases. Wild-type autosomes (A) and sex chromosomes (X and Y) are indicated by superscript plus signs. (a) In an autosomal dominant disorder such as Huntington's disease, only one mutant allele is needed to confer the disease. If either parent is heterozygous for the mutant *HD* allele, his or her children have a 50 percent chance of inheriting the mutant allele and getting the disease. (b) In an autosomal recessive disorder such as cystic fibrosis, two mutant alleles must be present to confer the disease. Both parents must be heterozygous carriers of the mutant *CFTR* allele for their children to be at risk of being affected or being carriers. (c) An X-linked recessive disease such as Duchenne muscular dystrophy is caused by a recessive mutation on the X chromosome and exhibits the typical sex-linked segregation pattern. Males born to mothers heterozygous for a mutant *DMD* allele have a 50 percent chance of

inheriting the mutant allele and being affected. Females born to heterozygous mothers have a 50 percent chance of being carriers.

The third common pattern of inheritance is that of an *X-linked recessive* allele. A recessive allele on the X chromosome will most often be expressed in males, who receive only one X chromosome from their mother, but not in females, who receive an X chromosome from both their mother and their father. The presence of an X-linked recessive allele leads to a distinctive sex-linked segregation pattern in which the disease is exhibited much more frequently in males than in females. For example, Duchenne muscular dystrophy (DMD), a muscle degenerative disease that specifically affects males, is caused by a recessive allele on the X chromosome. DMD exhibits the typical sex-linked segregation pattern in which mothers who are heterozygous, and therefore phenotypically normal, can act as carriers, transmitting the DMD allele, and therefore the disease, to 50 percent of their male progeny (Figure 6-26c).

DNA Polymorphisms Are Used as Markers for Linkage Mapping of Human Mutations

Once the mode of inheritance has been determined, the next step in finding a disease allele is to map its position with respect to known

genetic markers. Genetic mapping studies rely on exchanges of genetic information that occur during meiosis. As shown in Figure 6-27a, genetic **recombination** takes place in germ cells after the chromosomes of each homologous pair have replicated, but before the first meiotic cell division. At this time, homologous DNA sequences on maternally and paternally derived chromatids can be exchanged with each other in a process known as crossing over.

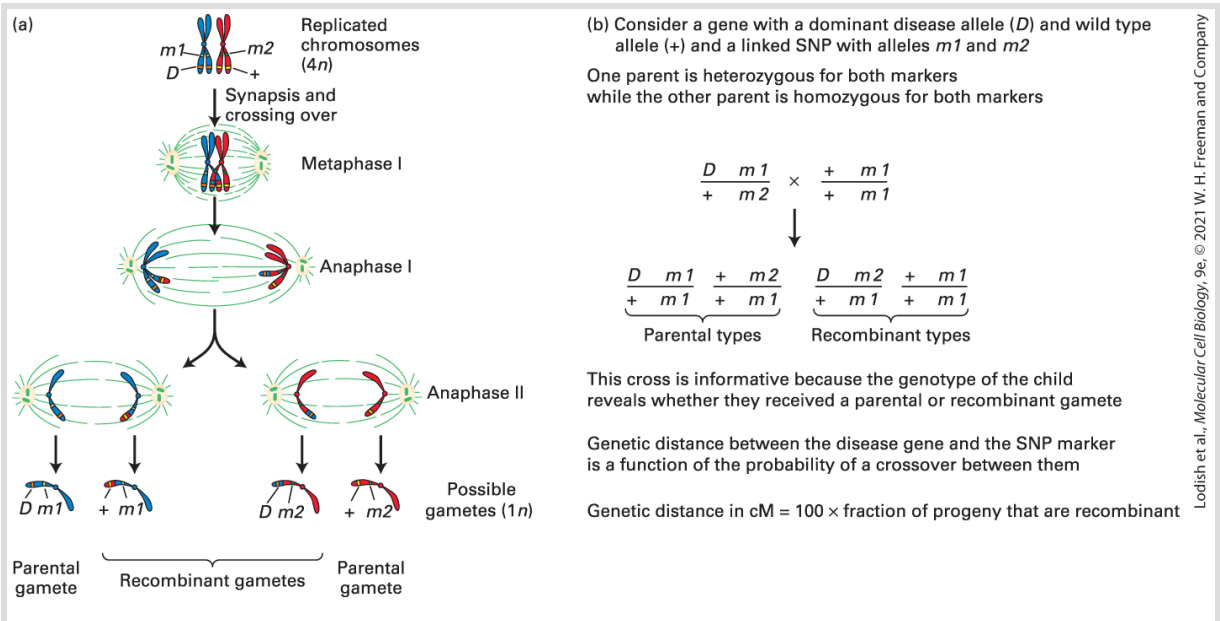


FIGURE 6-27 Recombination during meiosis can be used to map the positions of genes. (a) Consider the gametes produced by an individual that carries on one chromosome a dominant disease allele, designated *D* and linked SNP marker allele *m1,* and on the homologous chromosome the corresponding wild-type allele designated + and SNP marker *m2*. Because this arrangement of alleles is the same as in the parent of the individual these are known as parental types. If crossing over occurs within an interval between *the disease gene and the SNP marker*, then two *recombinant types* of gametes can be produced; one carries *D* and *m2* and one carries + and *m1*. The closer the SNP marker is to the disease gene, the less likely that a recombination event between them will occur, and the rarer recombinant types will be. If the markers were on different chromosomes or far apart on the same chromosome they would segregate independently of one another and recombinant types would occur with equal frequency as parental types. Note that in meiosis of the

human female, only one gamete is produced to form an egg cell. (b) Genetic mapping in humans involves scoring the genotypes of parents and offspring to determine the fraction of gametes that are recombinant types. For a cross to be informative, at least one of the parents must be heterozygous for both markers. In the example shown, the presence of the D allele in the gamete can be deduced from whether the child has the dominant trait and the presence of the m2 allele can be determined from the SNP genotype of the child. By scoring the gametes from many families in this way the genetic distance can be obtained by the fraction of recombinant types. By convention, the unit of genetic distance is a centimorgan (cM), defined as the distance between two positions along a chromosome that results in 1 recombinant individual in 100 progeny. Positions that show about as many recombinant types as parental types are said to be unlinked.

To see how genetic recombination during meiosis can be used as the basis of mapping the distances between genes, consider an individual with two different mutations, one inherited from each parent, that are located close to each other on the same chromosome. That individual can produce two different types of gametes according to whether a crossover occurs between the mutations during meiosis. If no crossover occurs between them, gametes known as *parental types*, which contain either one or the other mutation, will be produced. In contrast, if a crossover occurs between the two mutations, gametes known as *recombinant types* will be produced. In this example, recombinant chromosomes would contain either both mutations or neither one of them. Recombination events occur more or less at random along the length of chromosomes; thus the closer together two genes are, the less likely that recombination will happen to occur between them during meiosis. Thus *the less frequently recombination is observed to*

occur between two genes on the same chromosome, the closer together they must be. Two genes that are on the same chromosome and that are sufficiently close together that significantly fewer recombinant gametes than parental gametes are produced are considered to exhibit genetic **linkage**. If the number of recombinant gametes produced is not significantly less than the number of parental gametes, the two loci under consideration are considered to be *unlinked* and could be far apart on the same chromosome, or they could be on different chromosomes.

We have seen that by setting up the appropriate crosses, genetic recombination can be used in many different kinds of model organisms to determine the distance between two genetic markers that lie on the same chromosome. To find the map position of a new mutation, a process known as genetic mapping, it is necessary to carry out multiple pair-wise crosses with genetic markers of known position that span the entire genome. In the experimental organisms commonly used in genetic studies, numerous markers with easily detectable phenotypes are available for genetic mapping of mutations. For humans, there are not nearly enough phenotypic markers to carry out genetic mapping studies. Instead, recombinant DNA technology has made available a wealth of useful DNA-based molecular markers. Because most of the human genome does not encode proteins, a large amount of phenotypically inconsequential sequence variation exists between individuals. Indeed, it has been estimated that nucleotide differences between unrelated individuals

occur on an average of 1 of every 10^3 nucleotides. Since variations in DNA sequence, referred to as *DNA polymorphisms*, can be followed from one generation to the next by sequencing the DNA of individuals, they can serve as ideal genetic markers for linkage studies. Currently, a panel of as many as 10^6 different known polymorphisms whose locations have been mapped in the human genome is used for genetic linkage studies in humans.

Single-nucleotide polymorphisms (SNPs) constitute the most abundant type of DNA polymorphism and are therefore useful for constructing genetic maps of maximum resolution. Another useful type of DNA polymorphism consists of a variable number of repetitions of a two-, three-, or four-base sequence. Such polymorphisms, known as *short tandem repeats (STRs)* or *microsatellites*, presumably are formed by recombination or by slippage of either the template or newly synthesized strand during DNA replication. A useful property of STRs is that different individuals often have different numbers of repeats. The existence of multiple versions of an STR makes it more likely to produce an informative segregation pattern in a given pedigree and therefore to be of more general use in mapping the positions of disease genes. These polymorphisms can be detected by PCR amplification and DNA sequencing.

Human Linkage Studies Can Map Disease Genes with a Resolution of About 1 Mbp

Without going into all the technical considerations, let's see how a geneticist might map the allele conferring a particular dominant disease trait (e.g., familial hypercholesterolemia). The first step is to obtain DNA samples from all the members of a family containing individuals that exhibit the disease. The DNA from each individual, affected and unaffected, is then analyzed to determine that individual's genotype for a large number of known DNA polymorphisms (either STR or SNP markers can be used). The segregation pattern of each DNA polymorphism within the family is then compared with the segregation pattern of the disease under study. Polymorphisms that are not linked to the disease allele will not show any significant tendency to appear more frequently in individuals with the disease, whereas polymorphisms that are closely linked to it will almost always be present in individuals with the disease, because recombination events that would separate the disease allele and the polymorphism will be rare. Computer analysis of the segregation data is used to calculate the likelihood of linkage between each DNA polymorphism and the disease-causing allele. Under ideal circumstances, the segregation patterns from at least 10 individuals in a family are needed to establish statistically significant evidence for linkage.

In practice, segregation data are usually collected from different families exhibiting the same disease and pooled. The more families that can be examined, the greater the statistical significance of any evidence for linkage that can be obtained, and the greater the precision with which the distance between a linked DNA polymorphism and a disease allele can be measured. Most family studies have a maximum of about 100 individuals in whom linkage between a disease gene and a panel of DNA polymorphisms can be tested. This number of individuals sets the practical upper limit on the resolution of such a mapping study to about 1 centimorgan, or a physical distance of about 1 Mbp (1×10^6 bp).

Further Analysis Is Needed to Locate a Disease Gene in Cloned DNA

Although linkage mapping can usually locate a human disease gene to a region containing about 1 Mbp, as many as 10 different genes may be located in a region of this size. The ultimate objective of a mapping study is to locate the gene of interest within a cloned segment of DNA and then to determine the nucleotide sequence of this fragment. The relative scales of a chromosomal genetic map and physical maps corresponding to ordered sets of plasmid clones and the nucleotide sequence are shown in Figure 6-28.

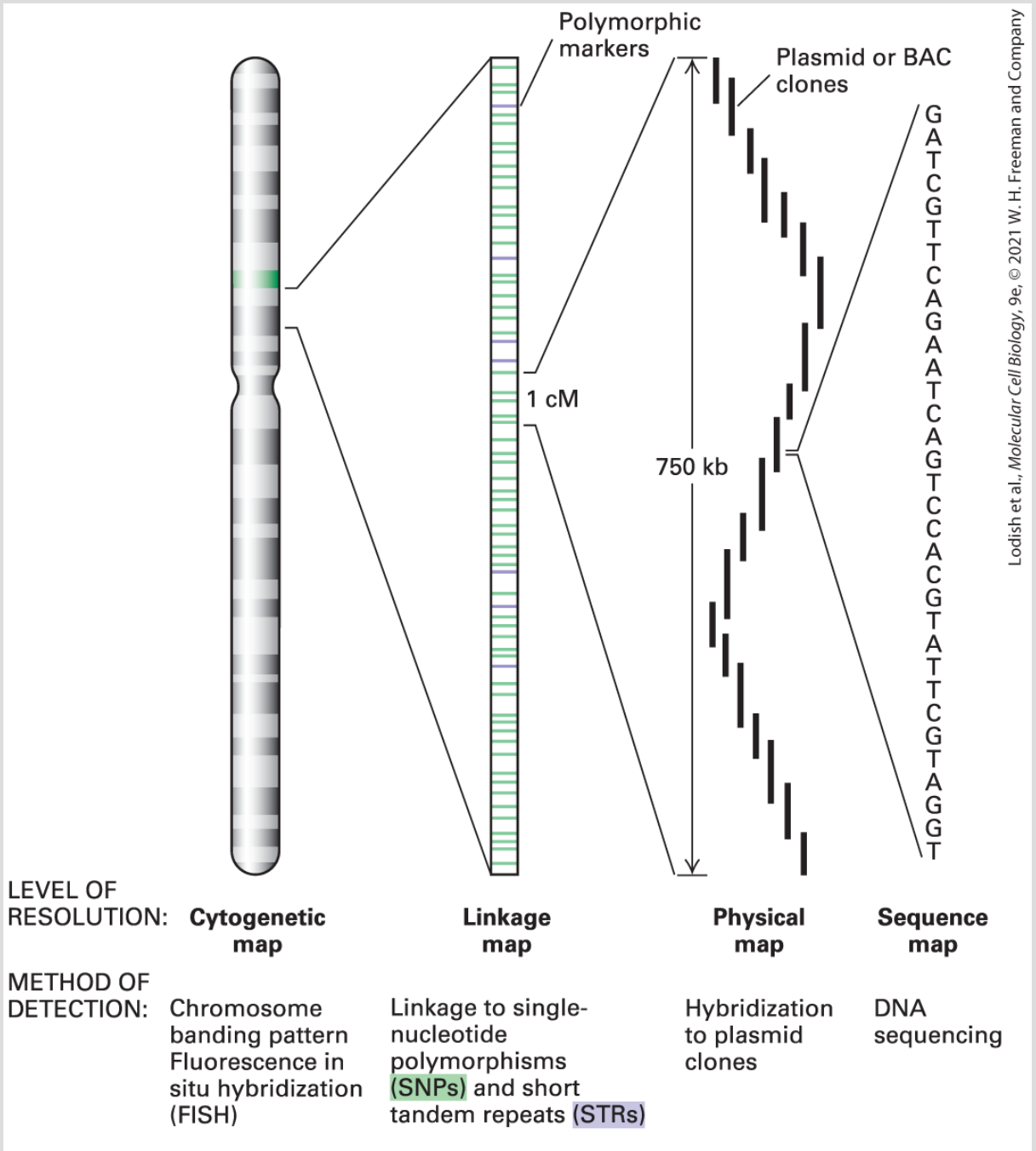


FIGURE 6-28 Relationship between genetic and physical maps of a human chromosome. The diagram depicts a human chromosome analyzed at different levels of detail. The chromosome as a whole can be viewed in the light microscope when it is in a condensed state that occurs at metaphase, and the approximate location of specific sequences can be determined by fluorescence in situ hybridization (FISH). At the next level of detail, genetic traits can be mapped relative to DNA-based genetic markers. These genetic mapping studies usually do not pinpoint the gene identity because they localize genes with an accuracy of

only 1 cM, which would correspond to a region of about 1 M bp. To identify the sequence of a gene segment local segments of the chromosome carried on bacterial artificial chromosomes (BACs) can be analyzed at the level of DNA sequences. Important genetic differences are ultimately defined by differences in the nucleotide sequence of the chromosomal DNA.

One strategy for further localizing a disease gene within the genome is to identify mRNA encoded by DNA in the region of the chromosome under study. The aim is to find an mRNA that is altered or missing in various individuals affected with a disease compared with wild-type mRNA. Such an mRNA would be an excellent candidate for encoding the protein whose disrupted function causes that disease. One strategy is to compare the expression of the genes in the candidate region in tissues from normal and affected individuals to discover tissues in which a particular disease gene is expressed in normal individuals but not in affected individuals. The rationale here is that a mutation that phenotypically affects muscle, but no other tissue, might be in a gene that is expressed only in muscle tissue. As discussed in Section 6.5, the level of expression of mRNA is generally determined by in situ hybridization of labeled DNA or RNA to tissue sections, by microarray analysis, or by massively parallel sequencing of cDNAs (RNAseq). Each of these methods allows researchers to compare gene expression in mutant and wild-type tissues. Although the sensitivity of in situ hybridization is lower than other methods, it is able to identify an mRNA that is expressed at low levels in a given tissue but at very high levels in a subclass of cells within that tissue.

In many cases, point mutations may give rise to disease-causing alleles and yet result in no obvious change in the level of expression of mRNAs or change in length. In that case, researchers undertake a search for point mutations in the DNA regions encoding the mRNAs. Now that highly efficient methods for sequencing DNA are available, researchers frequently identify point mutations by determining the sequences of candidate regions of DNA isolated from affected individuals. The overall strategy is to search for a coding sequence that consistently shows possibly deleterious alterations in DNA from individuals that exhibit the disease. A limitation of this approach is that the region near the affected gene may carry naturally occurring polymorphisms unrelated to the gene of interest. Such polymorphisms, not functionally related to the disease, can lead to misidentification of the DNA fragment carrying the gene of interest. For this reason, the more mutant alleles available for analysis, the more likely that a gene will be correctly identified.

Most Inherited Diseases Result from Multiple Genetic Defects

Until now we have considered monogenic diseases; that is, a clearly discernible disease state is produced by a defect in a single gene. Monogenic diseases caused by mutation in one specific gene exhibit one of the characteristic inheritance patterns shown in Figure 6-26. Researchers have already mapped the genes associated with

hundreds of common monogenic diseases, including those listed in Table 6-1, using DNA-based markers as described previously.

However, many other inherited diseases show more complicated patterns of inheritance. The underlying genetic cause of such diseases is much more difficult to identify. One frequently encountered type of added complexity is *genetic heterogeneity*. In such cases, mutations in any one of several different genes can cause the same disease. For example, retinitis pigmentosa, which is characterized by degeneration of the retina usually leading to blindness, can be caused by mutations in any one of more than 60 different genes. In human linkage studies, data from multiple families must usually be combined to determine whether a statistically significant linkage exists between a disease gene and known molecular markers. Genetic heterogeneity such as that exhibited by retinitis pigmentosa can confound such an approach because any statistical trend in the mapping data from one family tends to be canceled out by the data obtained from another family with an unrelated causative gene.

Human geneticists used two different approaches to identify the many genes associated with retinitis pigmentosa. The first approach relied on mapping studies of exceptionally large families. Each family contained so many affected individuals that it was possible to provide statistically significant evidence for linkage between known DNA polymorphisms and a single one of the many potential

causative genes. From these studies geneticists found that several of the mutations that cause retinitis pigmentosa lie within genes that encode proteins that are abundant in the retina. Following up on this clue, when screening other individuals with retinitis pigmentosa geneticists concentrated their attention on those genes that are highly expressed in the retina. This approach of using additional information to focus screening efforts on a subset of candidate genes led to identification of additional, rare causative mutations in many different genes encoding retinal proteins.

Identifying Component Genetic Risk Factors of Complex Traits

The most common human disorders not caused by foreign infectious agents are heart disease, obesity, hypertension, diabetes, most cancers, and a variety of mental disorders. Although diet, behavior, and other environmental conditions clearly contribute to these diseases, the existence of a strong genetic component is suggested by the propensity of these diseases to run in families. The relative contribution of genetics, known as *heritability*, can be formally measured by calculating the proportion of genetically identical monozygotic twins who have been raised apart, but nevertheless both have the disease. For the diseases listed, most of the incidence of disease is heritable and the contribution of genetics is measured to be from 60 to 80 percent. In all but a few rare

instances, the disease shows a complex pattern of inheritance that cannot be explained by a single causative allele — thus the genetic component of these diseases is known as a *complex trait*. Alleles from multiple genes must be involved in the production of complex traits, and for this reason these genes and their alleles are referred to as *risk factors*. The systematic mapping of risk factors for complex traits in humans is one of the most important and challenging problems in human genetics, and we already know that for each of the diseases listed alleles of tens and even hundreds of different genes may be risk factors.

The most promising method of studying complex traits is to seek a statistical correlation between inheritance of a particular region of a chromosome and the propensity to have a disease, using a procedure known as a **genome-wide association study (GWAS)**. To see how a GWAS would work, consider the complex trait of type II diabetes. Very large samples are needed to obtain statistically significant results because of the enormous heterogeneity of possible genetic causes for type II diabetes. A typical experiment might include a case group of say 5000 individuals with type II diabetes would be compared to a socioeconomically matched control group of 5000 individuals with normal blood glucose homeostasis. The basic concept behind a GWAS is quite simple. The SNPs represented in these 10,000 individuals are surveyed to find those that have a small but significantly greater tendency to appear in individuals that have type II diabetes than in those that do not.

The power of this approach lies in computer algorithms that scan data from large numbers of individuals to identify small but significant correlations between a disease and inheritance of a particular region of the genome. A conceptually straightforward method to find all of the SNPs represented in the 10,000 case and control individuals would be to obtain and survey their complete genome sequences. At present, large-scale sequencing projects of this type are not yet practical. Rather, a typical GWAS will survey a set of 10^6 of the most common SNPs.

Even rare disease-causing alleles not included in the set of 10^6 common SNPs can be identified by GWAS by taking advantage of the phenomenon of linkage disequilibrium. When an allele that predisposes an individual to a disease arises by mutation, the allele will reside within the context of whatever particular set of SNPs was carried by the ancestral chromosome. Even many generations later, genetic recombination between the disease gene and nearby SNPs will be rare, and the disease-causing allele will tend to remain associated with the particular set of SNPs in the neighborhood of its chromosomal location. This lingering association of nearby genetic markers is known as *linkage disequilibrium* because not enough time has passed for genetic recombination to have caused the association between markers to be randomized or, in other words, to have come to equilibrium (Figure 6-29). Thus a risk factor for type II diabetes that itself is too rare to be included in the standard set of 10^6 SNPs

may nevertheless be found by associations with SNPs that are in linkage disequilibrium.

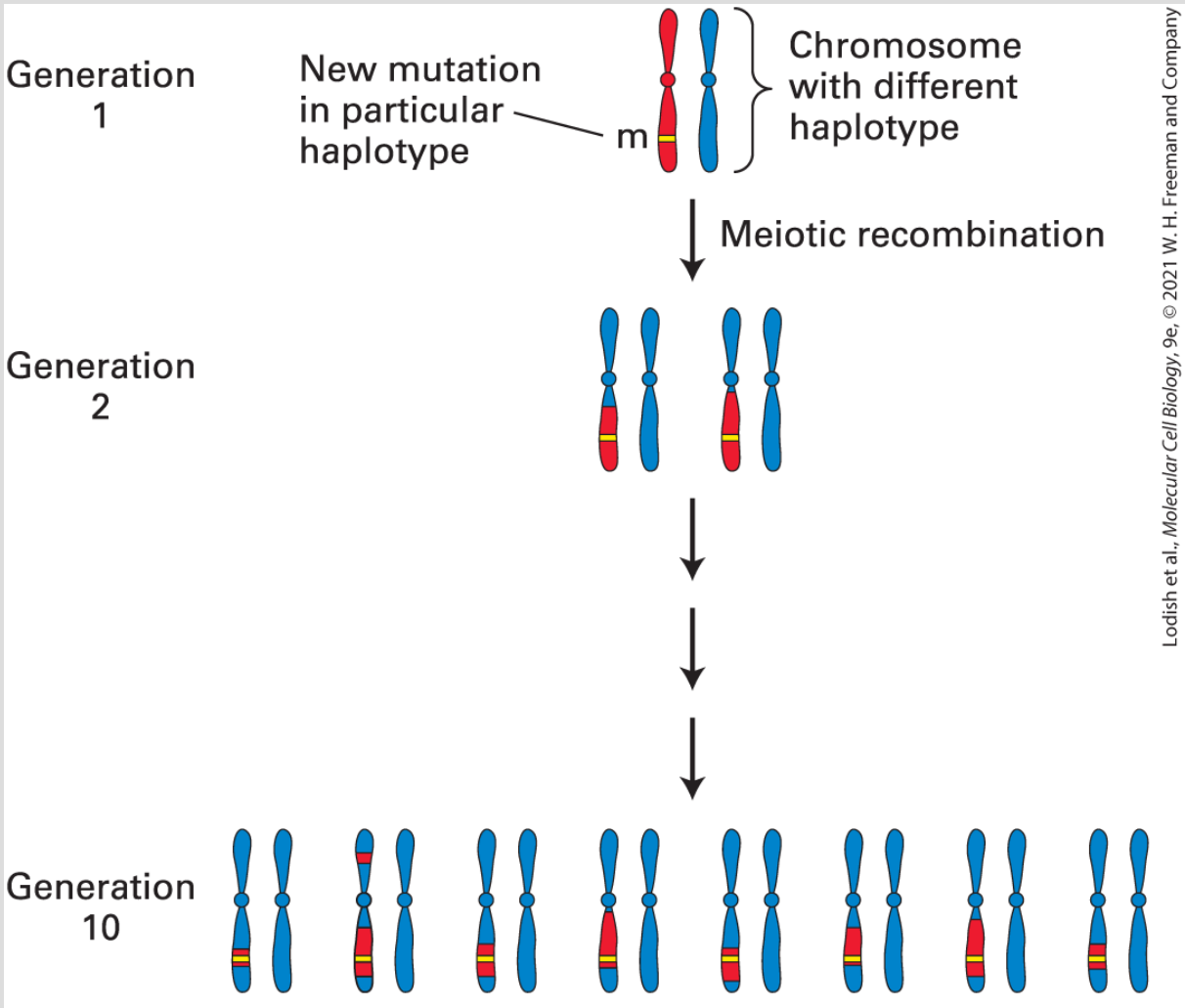


FIGURE 6-29 Linkage disequilibrium preserves the association between closely spaced SNPs. A new disease mutation arises in the context of an ancestral chromosome among a set of polymorphisms known as the *haplotype* of that chromosome (indicated by red shading; the blue segments of chromosomes represent general haplotypes derived from the general population and not from the ancestral haplotype in which the mutation originally arose). After many generations, chromosomes that carry the disease mutation will also carry segments of the ancestral haplotype that have not been separated from the disease mutation by recombination. The regions closest to the disease mutation are the most likely to be the ancestral haplotype. This phenomenon is known as *linkage disequilibrium*. The

position of the disease mutation can be located by scanning chromosomes that contain it for highly conserved polymorphisms corresponding to the ancestral haplotype.

GWAS can be a powerful tool to identify candidate genes that cause a predisposition to disease, but when all of the risk factors found by a typical GWAS are aggregated, usually they account for less than half of the total known heritability. Much of this missing heritability may eventually be accounted for by very rare SNPs or by factors that make such a small incremental contribution to the risk of disease that no significant association can be observed. As whole genome sequencing information becomes available for very large human cohorts, many more of these elusive risk factors should be found.

Medically Important Genes Can Be Identified by Alleles That Protect Against Disease

Until now, our discussion of human genes has centered on genes and alleles that cause inherited disease or increase its risk. The same technological advances that have accelerated the search for alleles that cause disease have enabled hunts for an equally important class of alleles that protect against disease. To illustrate the power of protective alleles, we will use the example of *PCSK9*, a gene recently found to play a key role in cholesterol homeostasis.

As described in Chapter 17, cholesterol moves through the bloodstream in lipoprotein particles known as LDL. High LDL levels in the blood can pose a risk of heart disease because excessive LDL can be deposited in coronary arteries as atherosclerotic plaque. *PCSK9* was first identified as a sequenced human gene that encodes a secreted protein with homology to a class of proteases. When it was discovered that a rare class of mutations that lead to increased LDL genetically mapped to *PCSK9*, a role in regulating cholesterol level was suspected. These mutated *PCSK9* alleles showed a dominant pattern of inheritance, implying that a gain of function leads to increased LDL. By extension of this logic, a loss of PCSK9 activity should decrease LDL levels. Sequencing the *PCSK9* gene in a large number of individuals to find frameshift and nonsense alleles confirmed that individuals with loss of function alleles did indeed have decreased LDL. Subsequent biochemical experiments have shown that the protein product of *PCSK9* negatively regulates the LDL receptor, which is responsible for transporting LDL from the blood into the liver by endocytosis. Inactivation of PCSK9 protein has been shown to dramatically decrease LDL levels and PCSK9 inhibitors have become a major target for cholesterol-lowering drugs. This great success in applying basic principles of human genetics to the discovery of the role of PCSK9 in cholesterol homeostasis has spurred interest in finding similar protective alleles for diabetes, obesity, and hypertension.

Identification of Causative Mutations in Cancer Cells

The cost of sequencing DNA has fallen so low that the entire genomes of cancer cells have been sequenced and compared with the genomes of normal cells from the same individuals in order to determine all the mutations that have accumulated in those persons' tumor cells. As discussed in Chapter 25, transformation of a normal cell into a cancer cell that divides without control usually requires multiple mutations. A major goal of sequencing cancer genomes is to identify all of the possible *driver* mutations implicated in transformation, but this approach is greatly complicated by the fact that transformed cells are usually defective in DNA repair and are prone to accumulate a large number of *passenger* mutations that do not contribute to cancer. To distinguish driver mutations from passenger mutations, investigators compared a large number of cancer genome sequences in an important application of the bioinformatic principles outlined in Section 6.3. The primary criterion for distinguishing driver mutations is that they are commonly present in all cancers, as well as commonly present in tumors from different patients with the same type of cancer (e.g., breast or colon cancer). A second criterion is based on the types of mutations that are found. A mutation in a driver gene would be expected to produce a significant change in gene function and may be a nonsense or frameshift mutation for a loss of function or a

missense mutation for a gain of function. In contrast, a mutation in a passenger gene is as likely as not to be a silent mutation.

KEY CONCEPTS OF SECTION 6.4

Locating and Identifying Genes That Specify Human Traits

- Inherited diseases and other traits in humans caused by a single gene show three major patterns of inheritance: autosomal dominant, autosomal recessive, and X-linked recessive (see Figure 6-26).
- Genes for human diseases and other traits can be mapped by determining their co-segregation with markers whose locations in the genome are known. The closer a gene is to a particular marker, the more likely they are to co-segregate (see Figure 6-27).
- Mapping of human genes with great precision requires thousands of molecular markers distributed along the chromosomes. The most useful markers are differences in the DNA sequence (polymorphisms) between individuals in noncoding regions of the genome, including single-nucleotide polymorphisms (SNPs) and short tandem repeats (STRs).
- Linkage mapping can often locate a human disease gene to a chromosomal region that includes as many as 10 genes. To identify the gene of interest within this candidate region typically requires expression analysis and comparison of DNA sequences between wild-type and disease-affected individuals.
- Some inherited diseases can result from mutations in different genes in different individuals (genetic heterogeneity).
- The most common inherited diseases depend on the presence of mutant alleles of multiple genes in the same individual. This type of disease is known as a complex trait and the alleles that contribute to the propensity for these diseases are known as risk factors. Risk factors can be mapped by finding a statistical correlation between the disease and a particular chromosomal location in a genome-wide association study.
- Genetic association studies can also identify alleles that protect against disease. Searching for alleles that protect against major inherited diseases is one of the most promising future directions of human genetics.
- Genomic sequencing of cancer cells reveals thousands of mutations that arise during the process of transformation into cancer cells. By comparing large numbers of cancer genome sequences, researchers can distinguish driver mutations, which contribute to

cell transformation, from passenger mutations, which do not contribute to cell transformation.

6.5 Using Cloned DNA Fragments to Study Gene Expression

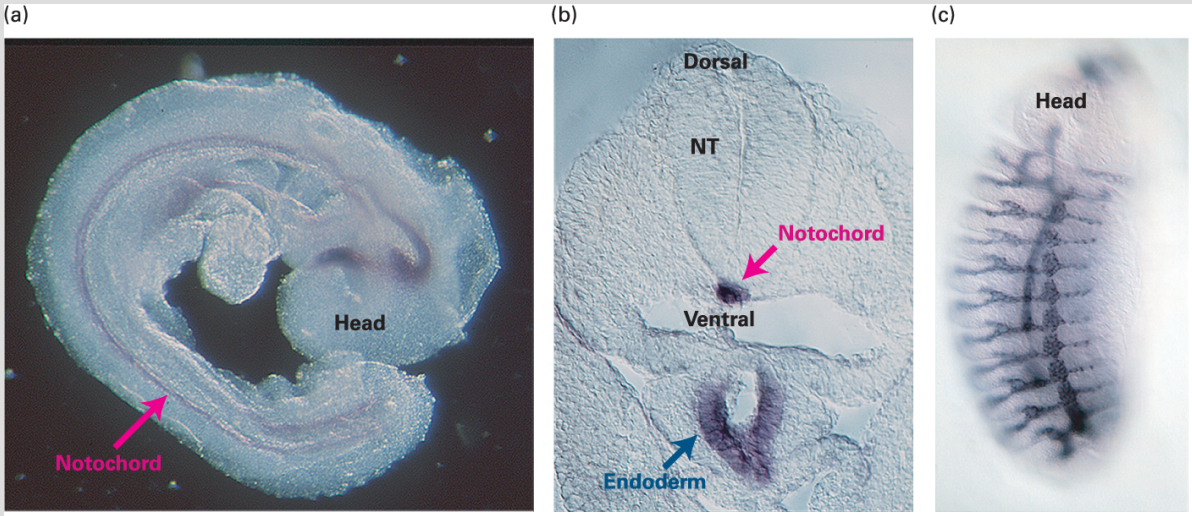
In the previous sections, we described how to identify and characterize genes by analysis of mutations and by bioinformatic analysis of genome sequences. We also described the basic techniques for using recombinant DNA technology to isolate specific DNA clones and ways in which those clones can be further characterized. Here we consider how an isolated DNA clone can be used to study gene expression. We discuss several widely used general techniques that rely on nucleic acid hybridization to elucidate when and where genes are expressed, as well as methods for generating large quantities of protein and otherwise manipulating amino acid sequences to determine their expression patterns, structure, and function.

In Situ Hybridization Techniques Permit Detection of Specific mRNAs

Hybridization depends on the ability of complementary single-stranded DNA or RNA molecules to associate (hybridize) specifically with each other via base pairing. As discussed in Chapter 5, double-stranded (duplex) DNA can be denatured (melted) into single strands by heating in a dilute salt solution. If the temperature is then lowered and the ion concentration raised, complementary single

strands will reassociate (hybridize) into duplexes. In a mixture of nucleic acids, only complementary single strands (or strands containing complementary regions) will reassociate; moreover, the extent of their reassociation is virtually unaffected by the presence of noncomplementary strands. The power of the hybridization method comes from the ability to adjust the hybridization conditions so that only perfectly complementary sequences will hybridize.

One of the most basic ways to characterize a cloned gene is to determine when and where in an organism the gene is expressed. The expression of a particular gene can be followed by identifying in which cells or tissues its mRNA is found. One way to locate a gene's corresponding mRNA is to probe for it by hybridization to a specific complementary oligonucleotide that is fluorescently labeled. Thus an oligonucleotide of only about 20 nucleotides can be used to label a single unique complementary mRNA. To retain precise information about the location of the mRNA, a chemically fixed whole or sectioned tissue, or even a fixed whole permeabilized embryo, may be subjected to **in situ hybridization** to detect the mRNA encoded by a particular gene. By examining the fixed tissue in a fluorescence microscope, the amount of gene transcription can be monitored in both time and space (Figure 6-30).



EXPERIMENTAL FIGURE 6-30 In situ hybridization can detect activity of specific genes in whole and sectioned embryos. The specimen is made permeable to nucleic acids by treatment with detergent and a protease to expose the mRNA to the probe. A DNA or RNA probe, specific for the mRNA of interest, is made with nucleotide analogs containing chemical groups that can be recognized by antibodies. After the permeabilized specimen has been incubated with the probe under conditions that promote hybridization, the excess probe is removed with a series of washes. The specimen is then incubated in a solution containing an antibody that binds to the probe. This antibody is covalently joined to a reporter enzyme (e.g., horseradish peroxidase or alkaline phosphatase) that produces a colored reaction product. After excess antibody has been removed, substrate for the reporter enzyme is added. A colored precipitate forms where the probe has hybridized to the mRNA being detected. (a) A whole mouse embryo at about 10 days of development probed for *Sonic hedgehog* mRNA. The stain marks the notochord (red arrow), a rod of mesoderm running along the future spinal cord. (b) A cross section of a mouse embryo similar to that in part (a). The dorsal/ventral axis of the neural tube (NT) can be seen, with the notochord expressing *Sonic hedgehog* (red arrow) below it and the endoderm (blue arrow) still farther ventral. (c) A whole *Drosophila* embryo probed for an mRNA produced during trachea development. The repeating pattern of body segments is visible. Anterior (head) is up; ventral is to the left.

DNA Microarrays Can Be Used to Evaluate the Expression of Many Genes at Once

The expression of thousands of genes can be monitored simultaneously using DNA microarray analysis, another technique based on the concept of nucleic acid hybridization. A **DNA microarray** consists of an organized array of thousands of individual, closely packed sequences, each specific to a single gene, attached to the surface of a glass microscope slide. By coupling microarray analysis with the results from genomic sequencing projects, researchers can analyze the global patterns of gene expression in an organism during specific physiological responses or developmental processes.

Preparation of DNA Microarrays

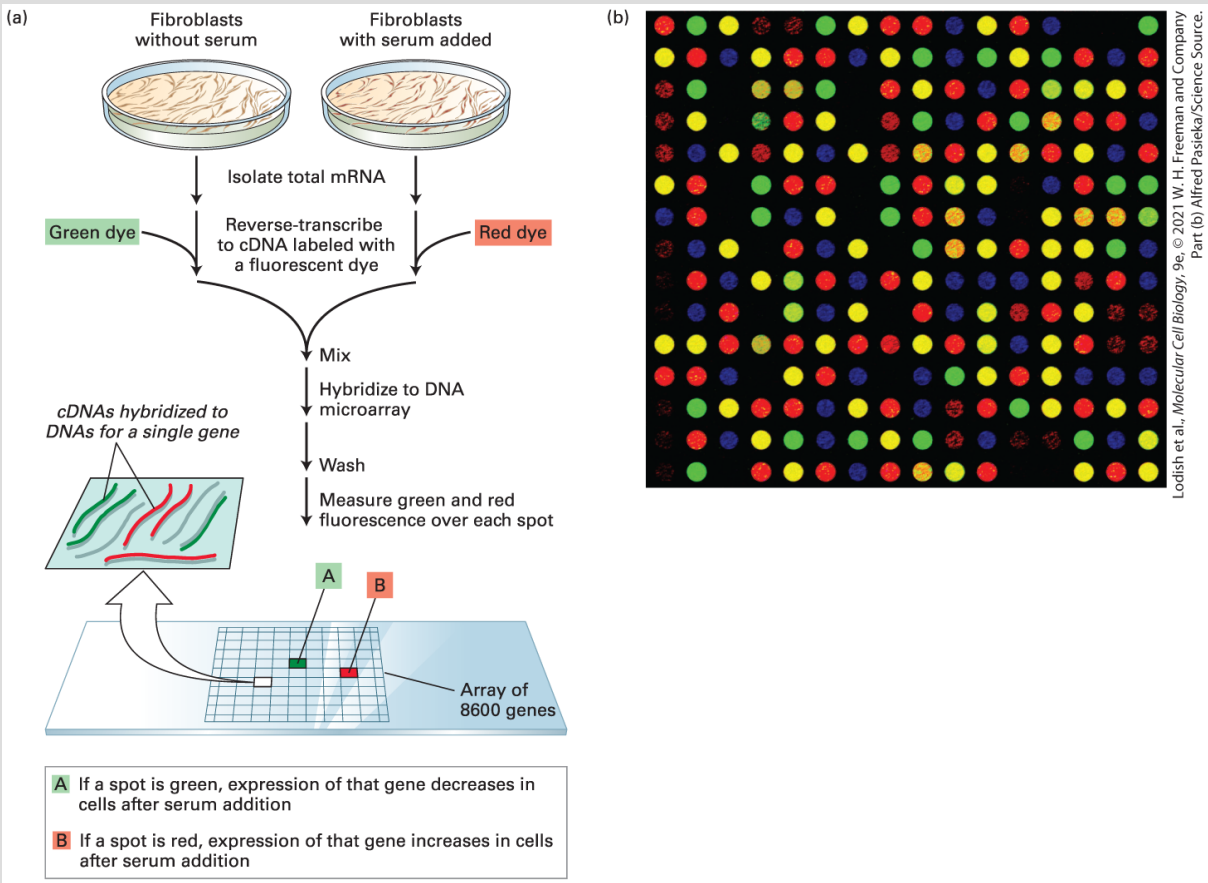
In one method of preparing microarrays, a DNA segment of about 1 kb corresponding to part of the coding region of each of many genes to be analyzed is individually amplified by PCR. A robotic device is used to apply each amplified DNA sample to the surface of a glass microscope slide, which is then chemically processed to permanently attach the DNA sequences to the glass surface and to denature them. A typical array might contain some 6000 spots of DNA in a 2×2–cm grid.

In an alternative method, multiple DNA oligonucleotides, usually at least 20 nucleotides in length, are synthesized, each starting from an initial nucleotide that is covalently bound to the surface of a glass slide. The synthesis of an oligonucleotide of specific sequence can be programmed in a small region on the surface of the slide. Several oligonucleotide sequences corresponding to different regions of a single gene are thus synthesized in neighboring regions of the slide. Using this method, oligonucleotides representing thousands of genes can be produced on a single glass slide. Because the methods for constructing these arrays of synthetic oligonucleotides were adapted from methods for manufacturing microscopic integrated circuits used in computers, these types of oligonucleotide microarrays are often called *DNA chips*.

Using Microarrays to Compare Gene Expression Under Different Conditions

The initial step in a microarray expression study is to prepare cDNAs from the mRNAs expressed by the cells under study and attach fluorescent labels to them (see Figure 6-17). When the cDNA preparation is applied to a microarray under appropriate conditions, DNA spots representing genes that are expressed will hybridize to their complementary cDNAs in the labeled probe mix and can subsequently be detected in a scanning laser microscope.

Figure 6-31 depicts how this method can be applied to examine the changes in gene expression observed after starved human fibroblasts are transferred to a rich, serum-containing growth medium. In this type of experiment, the separate cDNA preparations from starved and from serum-grown fibroblasts are labeled with differently colored fluorescent dyes. A DNA microarray comprising 8600 mammalian genes is then incubated with a mixture containing equal amounts of the two cDNA preparations under hybridization conditions. After unhybridized cDNA is washed away, the intensity of green and red fluorescence at each DNA spot is measured using a fluorescence microscope and is recorded under the name of each gene according to its known position on the slide. The relative intensities of red and green fluorescence signals at each spot provide a measure of the relative level of expression of that gene in response to serum. Genes that are not transcribed under these growth conditions give no detectable signal. Genes that are transcribed at the same level under both conditions hybridize equally to both red- and green-labeled cDNA preparations. Microarray analysis of gene expression in fibroblasts showed that transcription of about 500 of the 8600 genes examined changed substantially after addition of serum.



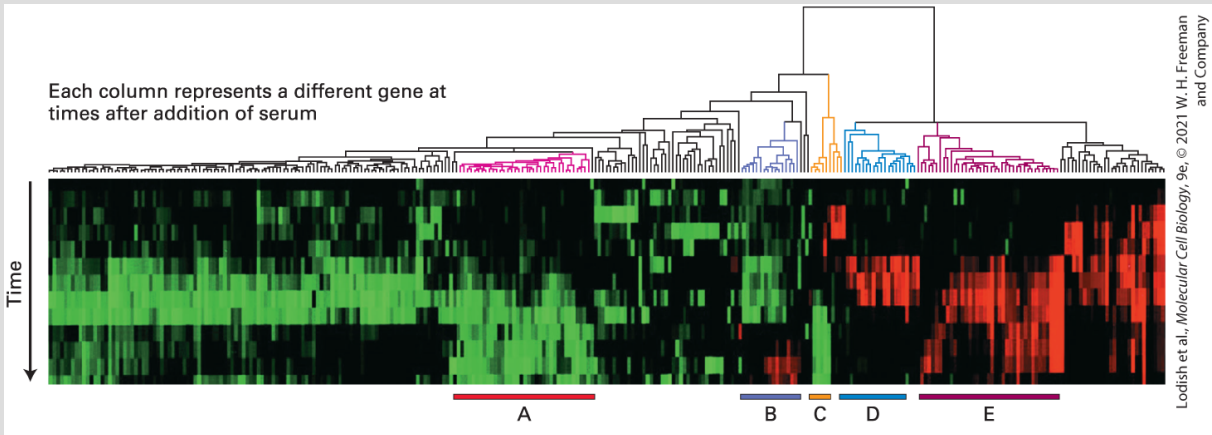
EXPERIMENTAL FIGURE 6-31 DNA microarray analysis can reveal differences in gene expression in fibroblasts under different experimental conditions. (a) In this example, cDNA prepared from mRNA isolated from fibroblasts either starved for serum or after addition of serum is labeled with different fluorescent dyes. A microarray composed of DNA spots representing 8600 mammalian genes is exposed to an equal mixture of the two cDNA preparations under hybridization conditions. The ratio of the intensities of red and green fluorescence over each spot, detected with a scanning confocal laser microscope, indicates the relative expression of each gene in response to serum. (b) A micrograph of a small segment of an actual DNA microarray. Each spot in this 16×16 array contains DNA from a different gene hybridized to control and experimental cDNA samples labeled with red and green fluorescent dyes. (A yellow spot indicates equal hybridization of green and red fluorescence, indicating no change in gene expression.)

Cluster Analysis of Multiple Expression Experiments Identifies Co-Regulated Genes

Microarray studies reveal sets of genes that show similar patterns of gene expression; these genes are expressed at the same times in the same tissues. Genes that are expressed under the same conditions and in the same cell types are hence likely to be closely related functionally. However, a firm conclusion can rarely be drawn from a single microarray experiment. For example, many of the observed differences in gene expression just described in fibroblasts could be indirect consequences of the many different changes in cell physiology that occur when cells are transferred from one medium to another. In other words, genes that appear to be co-regulated in a single microarray expression experiment may undergo changes in expression for very different reasons and may actually have very different biological functions. A solution to this problem is to combine the information from a set of microarray expression experiments to find genes that are similarly regulated under a variety of conditions or over a period of time.

This more informative use of multiple microarray expression experiments is illustrated by examining the relative expression of the 8600 genes mentioned above at different times after serum addition to fibroblasts, which generated more than 10^4 individual pieces of data. A computer program, related to the one used to

determine the relatedness of different protein sequences, can organize these data and cluster genes that show similar expression over the time course after serum addition. This *cluster analysis* based on gene transcription efficiently groups sets of genes whose encoded proteins participate in a common cellular process, such as cholesterol biosynthesis or the cell cycle (Figure 6-32).



EXPERIMENTAL FIGURE 6-32 Cluster analysis of data from multiple microarray expression experiments can identify co-regulated genes. The expression of 8600 mammalian genes was detected by microarray analysis at time intervals over a 24-hour period after serum-starved fibroblasts were provided with serum. The cluster diagram shown here is based on a computer algorithm that groups genes showing similar changes in expression compared with a serum-starved control sample over time. Each column of colored boxes represents a single gene, and each row represents a time point. A red box indicates an increase in expression relative to the control; a green box, a decrease in expression; and a black box, no significant change in expression. The “tree” diagram at the top shows how the expression patterns for individual genes can be organized in a hierarchical fashion to group together the genes with the greatest similarity in their patterns of expression over time. Five clusters of coordinately regulated genes were identified in this experiment, as indicated by the bars at the bottom. Each cluster contains multiple genes whose encoded proteins function in a particular cellular process: cholesterol biosynthesis (A), the cell cycle (B), the immediate-early response (C), signaling and angiogenesis (D), and wound healing and tissue remodeling (E).

Microarray analysis is a powerful diagnostic tool in medicine. Previously indistinguishable disease variations are now detectable. For instance, particular sets of mRNAs have been found to distinguish tumors with a poor prognosis from those with a good prognosis. Analysis of these tumor biopsies will help physicians to select the most appropriate treatment. As more patterns of gene expression characteristic of various diseased tissues are recognized, the diagnostic use of DNA microarrays will be extended to other conditions. ■

Sequencing of cDNAs Allows Analysis of Gene Expression in Individual Cells

A limitation of DNA microarray analysis is that while hybridization is highly selective, it is relatively insensitive. Many mRNAs that are present but are of low abundance fail to produce a detectable signal. A powerful method is available that enables detection of mRNAs that are present at only one copy per cell. Known as **whole transcriptome shotgun sequencing** (or **RNA-Seq**), the method is to directly subject a large number ($>10^6$) cDNAs from a tissue to shotgun sequence analysis. A sequencing read of about 100 bases is obtained for each cDNA, which allows each read sequence to be unambiguously assigned to a particular mRNA. The method makes it possible to quantify relative mRNA abundance by simply counting the relative number of sequencing reads corresponding to each

mRNA. For example, when cDNA prepared from hepatocytes is subjected to RNA-Seq analysis, sequence reads corresponding to abundant mRNAs such as that for serum albumin may be detected thousands of times out of a total of a million reads. Since one read per million is used as the threshold for detection of very rare mRNAs, we can see that RNA-Seq can measure mRNA abundance over a range of 10^4-fold, which is far greater than any experiment based on hybridization could achieve.

Refinements of RNA-Seq, such as introducing a few rounds of amplification of cDNAs by PCR, have increased the sensitivity of the method so that RNA abundance even in single cells can be measured. This method, called **single cell RNA-Seq (scRNA-Seq)**, allows detection of differences in gene expression among individual cells that could be obscured when different cell types in a tissue are combined for preparation of a bulk mRNA sample.

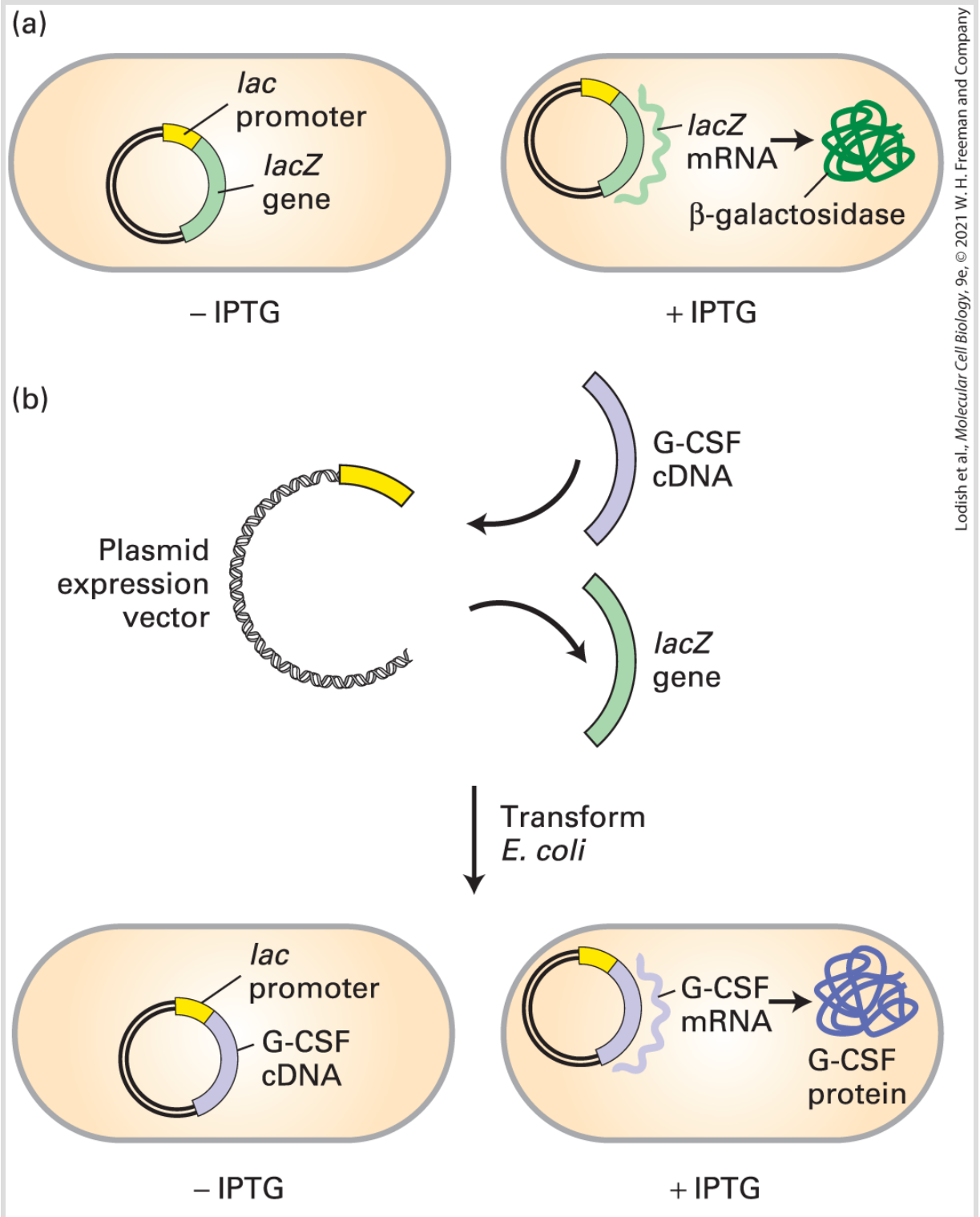
E. coli Expression Systems Can Produce Large Quantities of Proteins from Cloned Genes

An important application of our knowledge of how to engineer genes for high levels of expression is the production of large quantities of useful proteins. Many protein hormones and other signaling or regulatory proteins are normally expressed at very low

concentrations and so cannot be isolated and purified in large quantities by standard biochemical techniques. To enable the widespread therapeutic use of such proteins, as well as basic research on their structure and functions, efficient procedures are needed for producing them in large amounts at reasonable cost. Recombinant DNA techniques that turn *E. coli* cells into factories for synthesizing low-abundance proteins are now used to produce commercially a variety of human proteins with therapeutic uses, including granulocyte colony-stimulating factor (G-CSF), insulin, growth hormone, and erythropoietin. For example, G-CSF stimulates the production of granulocytes, the phagocytic white blood cells that are critical to defense against bacterial infections. Administration of G-CSF to cancer patients helps offset the reduction in granulocyte production caused by chemotherapeutic agents, thereby protecting patients against serious infection while they are receiving chemotherapy. ■

The first step in producing large amounts of a low-abundance protein is to obtain a cDNA clone encoding the full-length protein by the methods (see Figure 6-17). The second step is to engineer plasmid vectors that will express large amounts of the encoded protein when they are inserted into *E. coli* cells. The key to designing such expression vectors is to include a promoter, a DNA sequence from which transcription of the cDNA can begin. Consider, for example, the relatively simple system for expressing G-CSF shown in Figure 6-33. In this case, G-CSF is expressed in *E. coli* transformed

with plasmid vectors that contain the *lac* promoter adjacent to the cloned cDNA encoding G-CSF. Transcription from the *lac* promoter occurs at high rates only when lactose, or a lactose analog, such as isopropylthiogalactoside (IPTG), is added to the culture medium. Even larger quantities of a desired protein can be produced in more complicated *E. coli* expression systems.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 6-33 Some eukaryotic proteins can be produced in *E. coli* cells from plasmid vectors containing the *lac* promoter. (a) This plasmid expression vector contains a fragment of the *E. coli* chromosome containing the *lac* promoter and the

neighboring *lacZ* gene. In the presence of the lactose analog IPTG, RNA polymerase normally transcribes the *lacZ* gene, producing *lacZ* mRNA, which is translated into the encoded protein, β-galactosidase. (b) The *lacZ* gene can be cut out of the expression vector with restriction enzymes and replaced by a cloned cDNA, in this case, one encoding granulocyte colony-stimulating factor (G-CSF). When the resulting plasmid is inserted into *E. coli* cells, addition of IPTG and subsequent transcription from the *lac* promoter produce G-CSF mRNA, which is translated into G-CSF protein.

To aid in purification of a eukaryotic protein produced in an *E. coli* expression system, researchers often modify the cDNA encoding the recombinant protein to make it easier to separate from endogenous *E. coli* proteins. For example, a short nucleotide sequence is commonly added to the end of the cDNA, so that the expressed protein will have six histidine residues at the C-terminus. Proteins modified in this way bind tightly to an affinity matrix that contains chelated nickel atoms, whereas *E. coli* proteins do not bind to such a matrix. The bound proteins can be released from the nickel atoms by decreasing the pH of the surrounding medium. In most cases, the modified protein is still functional, since addition of short amino acid sequences to either the C-terminus or the N-terminus of a protein usually does not interfere with the protein's biochemical activity.

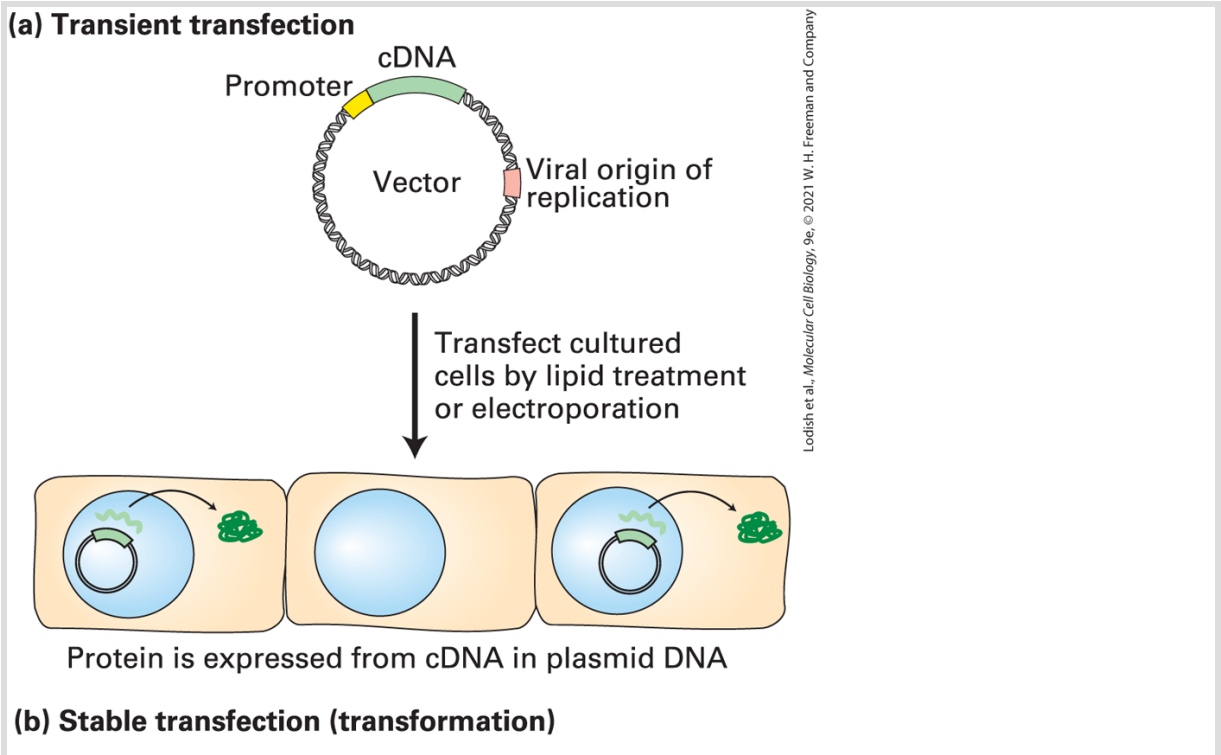
Plasmid Expression Vectors Can Be Designed for Use in Animal Cells

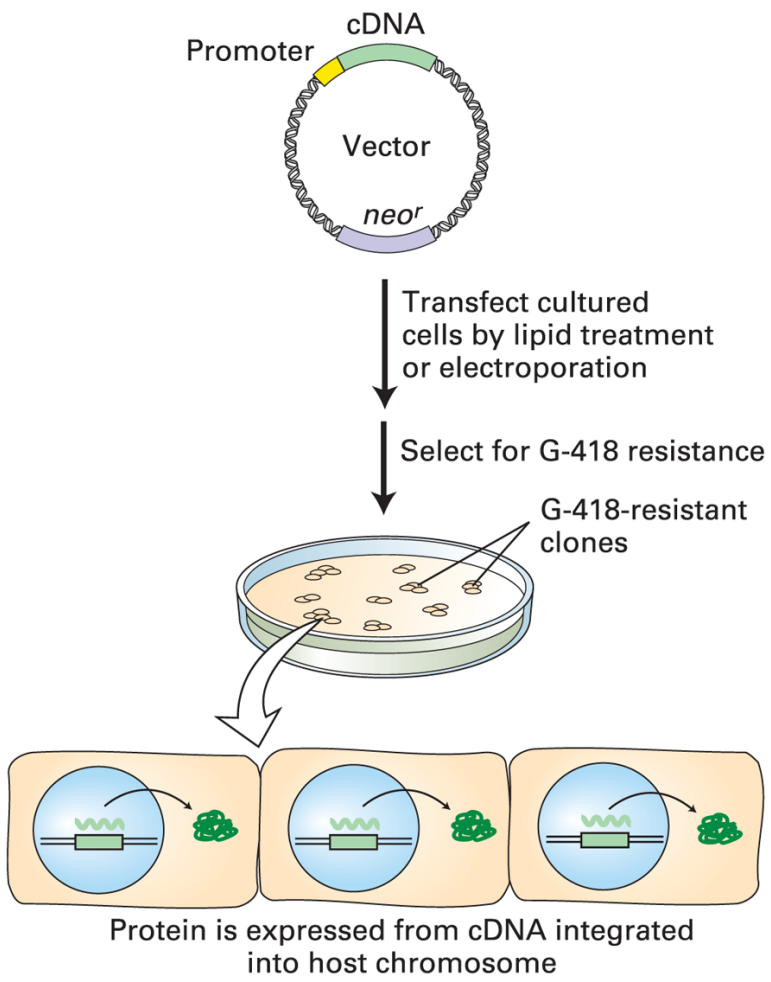
Although bacterial expression systems can be used successfully to create large quantities of some proteins, bacteria are not suitable expression hosts in many cases. Many experiments are designed to examine the function of a protein in the context of an appropriate cell. In such cases, a genetically modified protein must be expressed in cultured animal cells. To accomplish this, genes are cloned into specialized eukaryotic expression vectors and are introduced into cultured animal cells by a process called **transfection**. Two common methods for transfecting animal cells differ in whether the recombinant vector DNA is or is not integrated into the host-cell genomic DNA.

In both methods, cultured animal cells must be treated to facilitate their initial uptake of the recombinant plasmid vector. This can be done by exposing cells to a preparation of lipids that penetrates the plasma membrane, increasing its permeability to DNA. Alternatively, cells are subjected to a brief electric shock of several thousand volts, a technique known as *electroporation* that makes the cells transiently permeable to DNA. Usually the plasmid DNA is added in sufficient concentration to ensure that a large proportion of the cultured cells will receive at least one copy. Researchers have also harnessed viruses for transfection; viruses can be modified to contain DNA of interest, which is then introduced into host cells by simply infecting them with the recombinant virus.

Transient Transfection

The simpler of the two transfection methods, called *transient transfection*, employs a plasmid vector similar to the yeast shuttle vectors described previously. For use in mammalian cells, plasmid vectors are engineered to carry a replication origin derived from a virus that infects mammalian cells, a strong promoter recognized by mammalian RNA polymerase, and the cloned cDNA encoding the protein to be expressed adjacent to the promoter (Figure 6-34a). Once such a plasmid vector enters a mammalian cell, the viral replication origin allows it to replicate efficiently, generating numerous plasmids from which the protein is expressed. However, during cell division, such plasmids are not faithfully segregated into both daughter cells, and in time, a substantial fraction of the cells in a culture will not contain a plasmid, hence the name *transient transfection*.





EXPERIMENTAL FIGURE 6-34 Transient and stable transfection with specially designed plasmid vectors permits expression of cloned genes in cultured animal cells. Both methods employ plasmid vectors that contain the usual elements — replication origin, selectable marker (e.g., amp^r), and polylinker — that permit propagation in *E. coli* as well as a cloned cDNA with an adjacent animal promoter. For simplicity, these elements are not depicted. (a) In transient transfection, the plasmid vector contains a replication origin from a virus that can replicate in the cultured animal cells. Since the vector is not incorporated into the genome of the cultured cells, production of the cDNA-encoded protein continues for only a limited time. (b) In stable transfection, the vector carries a selectable marker such as neo^r, which confers resistance to G-418. The relatively few transfected animal cells that integrate the exogenous DNA into their genomes are selected on medium containing G-418. Because the vector is integrated into the genome, these stably transfected, or transformed, cells will continue to produce the cDNA-encoded protein as long as the culture is maintained.

Stable Transfection (Transformation)

The second of the two transfection methods integrates the introduced vector into the genome of the host cell. In that case, the genome is permanently altered, and the cell is said to be *transformed*. Integration is most likely accomplished by endogenous enzymes that normally function in DNA repair and recombination. The basic procedure for expressing a cloned cDNA by *stable transfection* is outlined in Figure 6-34b. A commonly used selectable marker is the gene for neomycin phosphotransferase (designated neo^r), which confers resistance to a toxic compound chemically related to neomycin, known as G-418. In the presence of a high concentration of G-418, only those cells that have integrated the expression vector into the host chromosome will survive and give rise to a clone. Because the expression vector is integrated at random sites in the genome, individual transformed clones will differ in their rates of transcribing the inserted cDNA. Therefore, the stable transfectants are usually screened to identify those that produce the protein of interest at the highest levels.

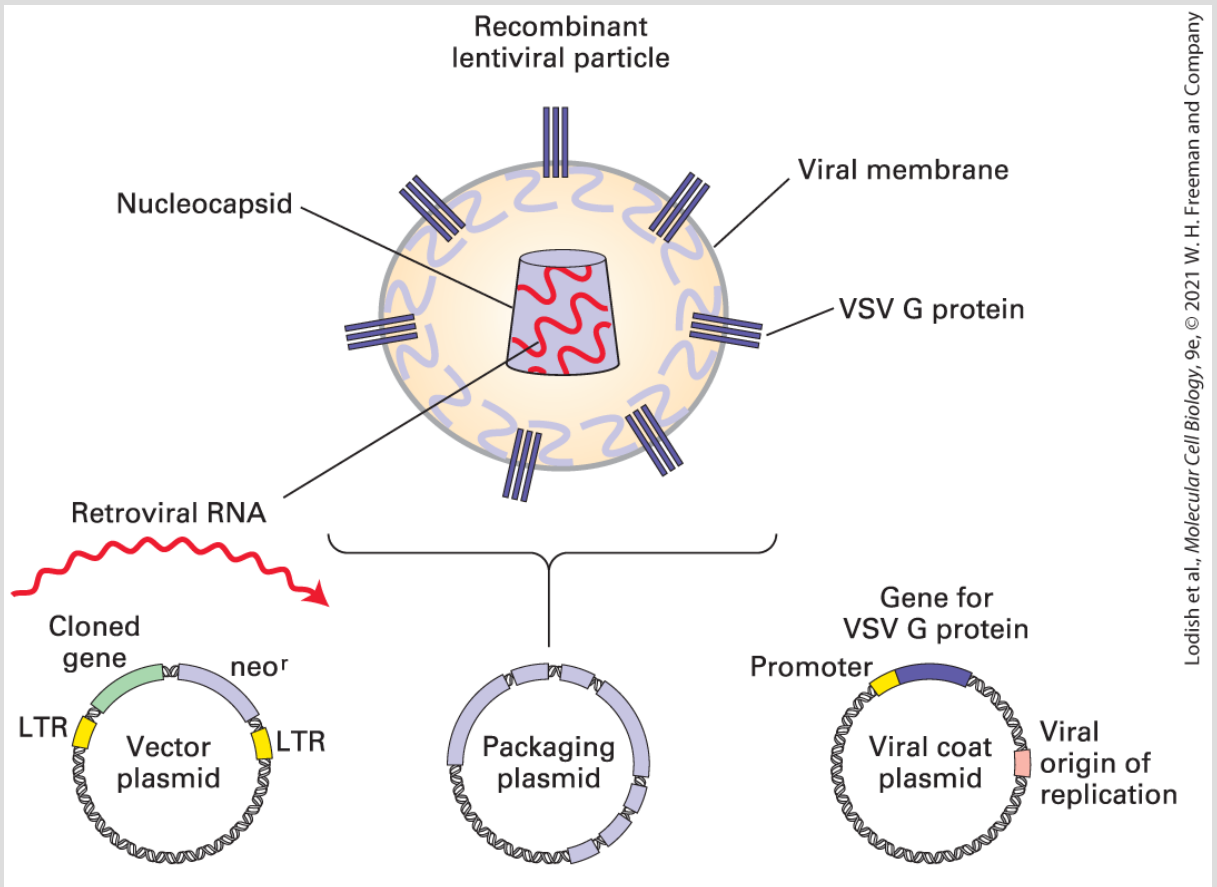
Retroviral Expression Systems

Researchers have exploited the basic mechanisms used by viruses for inserting their genetic material into the host's chromosomal DNA to greatly increase the efficiency by which a modified gene can be stably expressed in animal cells. One such viral expression system is derived from a class of retroviruses known as *lentiviruses*.

The system utilizes lentivirus particles that carry an RNA version of the cloned gene incorporated into the virus's RNA genome. The cloned gene is flanked by lentivirus LTR sequences. In the target cell, the LTR sequences direct the copying of the RNA into double-stranded DNA by reverse transcription, resulting in the integration of that DNA into chromosomal DNA following a sequence of events depicted in Figure 7-14.

The key to the procedure is the production of lentivirus particles suitable for efficient introduction of a cloned gene into target animal cells. As shown in Figure 6-35, three different plasmids, introduced into cells by transient transfection, are used to produce recombinant lentivirus particles. The first plasmid, known as the *vector plasmid*, contains a cloned gene of interest next to a selectable marker such as neo^r flanked by lentivirus LTR sequences. The left LTR sequence directs synthesis of an RNA molecule that carries lentiviral LTR sequences at either end and thus has many of the properties of native retroviral RNA. In an appropriate host, this LTR-bearing RNA can be packaged into viral particles with the assistance of the other two plasmids. The second plasmid, known as the packaging plasmid, carries all of the viral genes except for the major viral envelope protein, necessary for packaging of LTR containing viral RNA into the nucleocapsid of a functional lentiviral particle. The final plasmid allows expression of a viral envelope protein that, when incorporated into a recombinant lentivirus, allows the resulting hybrid virus particles to infect a desired target cell type. A

common envelope protein used in this context is the glycoprotein of the vesicular stomatitis virus (VSV-G protein), which can readily replace the normal lentivirus envelope protein on the surface of completed virus particles. Virus particles coated with the VSV-G protein are able to infect a wide variety of mammalian cell types, including hematopoietic stem cells, neurons, and muscle and liver cells.



EXPERIMENTAL FIGURE 6-35 Retroviral vectors can be used for efficient integration of cloned genes into the mammalian genome. Recombinant lentiviral particles are produced in cells expressing viral proteins from a packaging plasmid and that include a coat protein that confers a broad host range expressed from the viral coat plasmid. These viral particles will package viral RNA synthesized from a vector plasmid that contains a gene sequence of interest. When a recombinant lentiviral particle formed in this way infects a recipient cell the

cloned gene of interest is introduced into the cell in a form that can integrate into chromosomal DNA and thus be stably expressed in the recipient cell.

After cell infection, the cloned gene flanked by the viral LTR sequences is reverse-transcribed into DNA, which is transported into the nucleus and then integrated into the host genome. If necessary, as in the case of stable transfection, cells with a stably integrated cloned gene and neo^r marker can be selected by resistance to G-418. Many of the techniques for inactivating the function of specific genes (see Section 6.6) require that an entire population of cultured cells be genetically modified simultaneously. Engineered lentiviruses are particularly useful for such experiments because they infect cells with such high efficiency that every cell in a population will receive at least one copy of the lentivirus-borne plasmid.

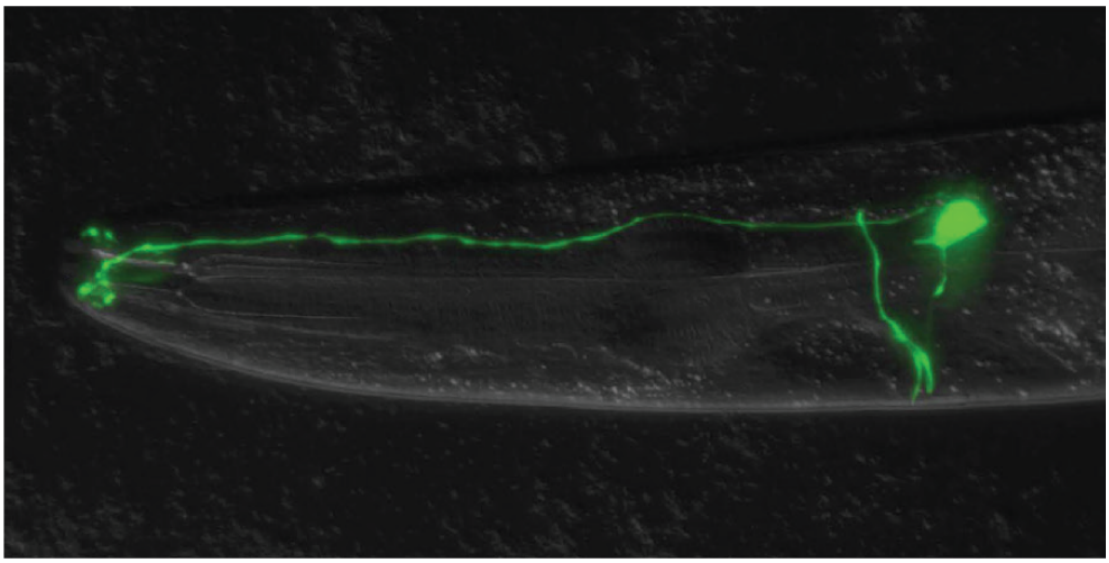
Gene and Protein Tagging

Expression vectors can provide a way to study the expression of eukaryotic proteins and their location with the cell. Such studies often rely on the use of a reporter protein, such as *green fluorescent protein (GFP)*, which can be conveniently detected in cells (see Figure 4-15). Here we describe two ways to create a hybrid gene that connects expression of the reporter protein to expression of the protein of interest. When the hybrid gene is reintroduced into cells, either by transfection with a plasmid expression vector containing

the modified gene or by creation of a transgenic animal as described in Section 6.6, the expression of the reporter protein can be used to determine where and when a gene is expressed. This method provides data similar to that from the in situ hybridization experiments described previously, but often with greater resolution and sensitivity.

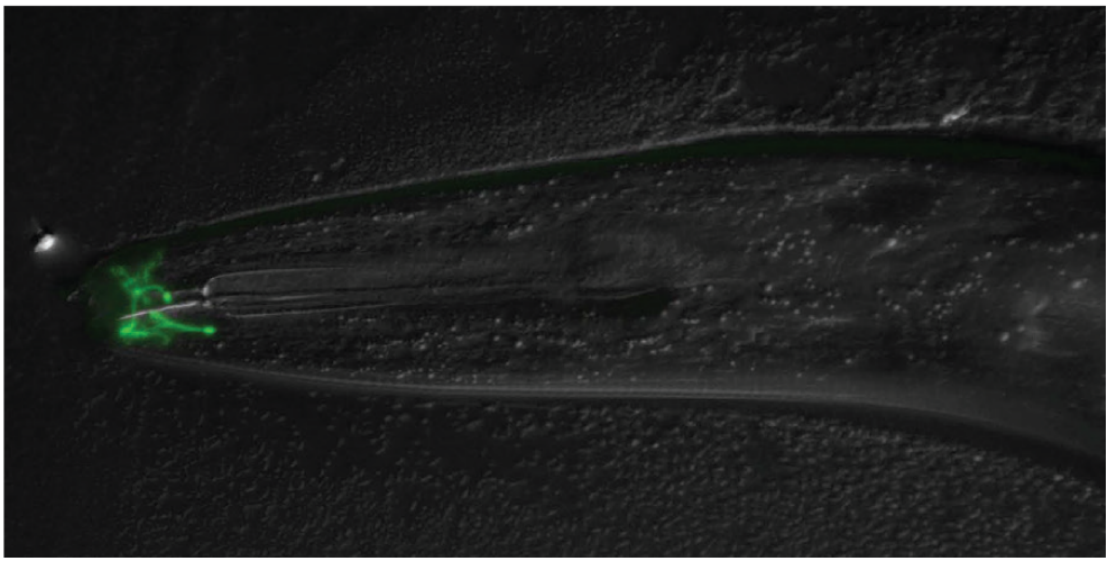
Figure 6-36 illustrates the use of two different types of GFP-tagging experiments to study the expression of an odorant receptor protein in *C. elegans*. In the first type of experiment, the coding sequence of GFP is placed under control of the promoter for the odorant receptor in a configuration usually known as a *promoter-fusion*. In the second type of experiment, the coding sequence of GFP is linked directly to the coding sequence of the odorant receptor, in a configuration known as *protein-fusion.* When GFP is linked to the promoter, the GFP is expressed throughout specific neurons; thus this promoter fusion experiment provides readout of gene expression. But to look at the protein in the cell, protein-fusion methods are used. When GFP is linked to the odorant receptor coding sequence, it is localized at the distal cilia in sensory neurons, the site at which the receptor protein is normally located.

(a) Promoter-fusion; ODR10 promoter fused to GFP



Ashish Maurya.

(b) Protein-fusion; ODR10-GFP fusion protein



EXPERIMENTAL FIGURE 6-36 Gene and protein tagging facilitate cellular localization of proteins expressed from cloned genes. In this experiment, the gene encoding a chemical odorant receptor, *Odr10*, of *C. elegans* was fused to the gene encoding green fluorescent protein (GFP). (a) A promoter-fusion was generated by linking just the promoter of *Odr10* to the coding sequence for GFP. The result is that GFP is expressed in the cytoplasm of the same specific sensory neurons in the head of *C. elegans* where Odr10 is expressed. (b) A

protein-fusion was constructed by linking GFP to the end of the full-length Odr10 coding sequence. In this case, the Odr10-GFP fusion protein is targeted to the membrane at the tip of the sensory neurons and is apparent only at the distal end of the sensory cilia. The observed distribution can be inferred to reflect the normal location of Odr10 protein in specific neurons. Because the promoter-fusion shown in (a) lacks the Odr10 localization sequences, the expressed GFP fills the entire cell cytoplasm rather than being localized just to the distal tip of the sensory cilia.

An alternative to GFP tagging for detecting the location of a protein inside the cell is to modify the gene of interest by appending to it a short DNA sequence that encodes a short stretch of amino acids recognized by a known monoclonal antibody. Such a short peptide that can be bound by an antibody is called an *epitope*; hence this method is known as *epitope tagging*. After transfection of cells with a plasmid expression vector containing the modified gene, the expressed epitope-tagged form of the protein can be detected by immunofluorescence microscopy utilizing a monoclonal antibody specific for the epitope. The choice of whether to use a short epitope or GFP to tag a given protein often depends on what types of modification a cloned gene can tolerate and still remain functional.

GFP expressed in *C. elegans* can be imaged by fluorescent microscopy because the body of a living worm is transparent. To image gene expression in a larger animal such as the mouse a more sensitive bioluminescent reporter can be used. This method uses an enzyme derived from insects known as luciferase, which is responsible for the flashes of light emitted by fireflies. When provided with the substrate luciferin, luciferase will catalyze a

chemical reaction that produces visible light. Visible light photography can detect the light from a luciferase reporter emitted from within the body of a living mouse. The chapter opening figure shows a transgenic mouse expressing a promoter fusion of the T-cell specific promoter CD2 expressing luciferase.

KEY CONCEPTS OF SECTION 6.5

Using Cloned DNA Fragments to Study Gene Expression

- Hybridization of a labeled DNA to the complementary RNA or DNA sequence allows specific detection of a single mRNA species in the context of a whole cell.
- The presence and distribution of specific mRNAs within tissues can be determined by in situ hybridization and fluorescence microscopy.
- DNA microarray analysis simultaneously detects the relative levels of expression of thousands of genes in different types of cells or in the same cells under different conditions (see Figure 6-30).
- Cluster analysis of the data from multiple microarray expression experiments can identify genes that are similarly regulated under various conditions. Such co-regulated genes commonly encode proteins that have biologically related functions.
- RNA-Seq analysis, based on direct sequencing of millions of cDNAs, gives a picture of all of the mRNAs that are expressed in a single cell.
- Expression vectors derived from plasmids allow the production of abundant amounts of a protein from a cloned gene.
- Eukaryotic expression vectors can be used to express cloned genes in the context of a living organism. An important application of these methods is to the tagging of proteins with GFP or some other identifiable marker to determine the pattern of expression and intracellular location.

6.6 Altering the Function of Specific Genes by Design

The vast amount of genomic DNA sequence information obtained in recent years has led to the identification of many new genes. The function of a new gene can often be deduced from the similarity of the encoded proteins with proteins of known function, but its roles in the context of a living organism can't be known with certainty unless a mutant forms of the gene is available. In this section, we describe several ways of disrupting the normal function of a specific gene in the genome of an organism. Analysis of the resulting mutant phenotype often helps reveal the in vivo function of the normal gene and its encoded protein.

We will begin this section with a description of how gene deletions are constructed in yeast. Yeast is the most easily genetically manipulatable organism and is the first organism for which deletions for every gene have been constructed. This comprehensive collection of null alleles has allowed researchers to observe the effect of completely inactivating every gene in an organism. We will then turn to the recently developed, powerful genome editing method known as CRISPR, which is based on a bacterial system for making cuts in specific DNA sequences. This system has been adapted to edit specific genes in a wide variety of organisms, including mammals. The basic CRISPR protocol will cause a small deletion in the targeted gene, which usually generates a null allele.

Alternatively, gene researchers often desire to study the effect of a particular allele of a gene, such as a dominant gain of function allele, and CRISPR has been modified to carry out this more specific type of genome editing as well.

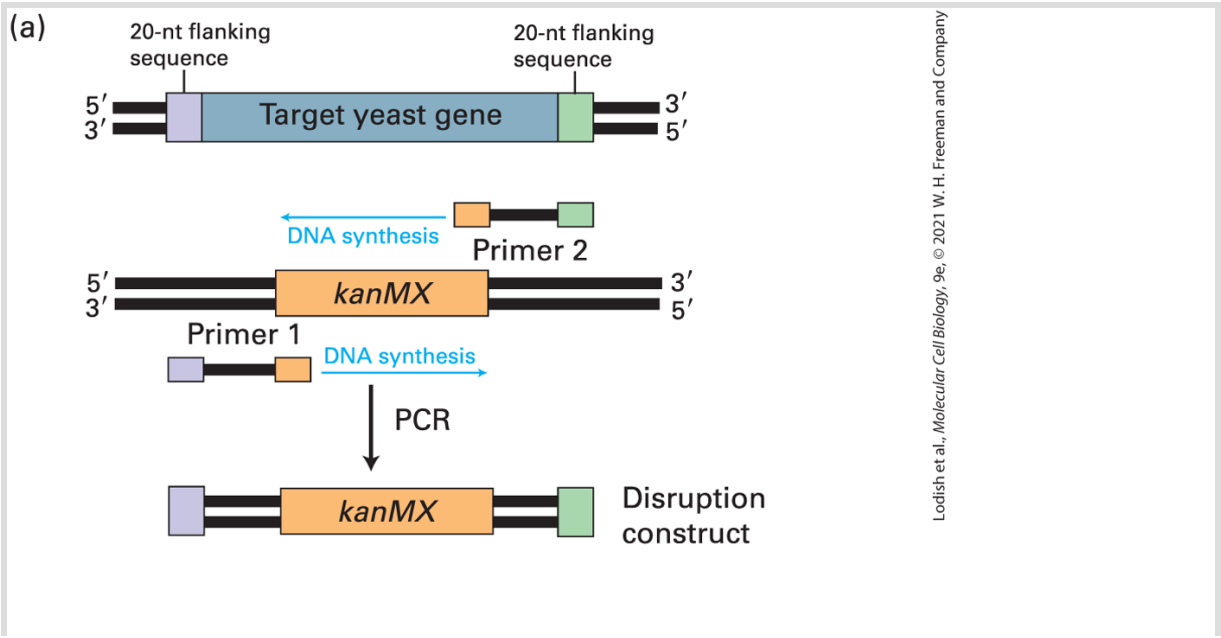
We will also describe other approaches that are used widely in model organisms to study the effects of gene inactivation: (1) use of transgenic mice to construct gene alleles that can be selectively inactivated in particular tissues and (2) promoting destruction of the mRNA transcribed from a gene using RNA interference.

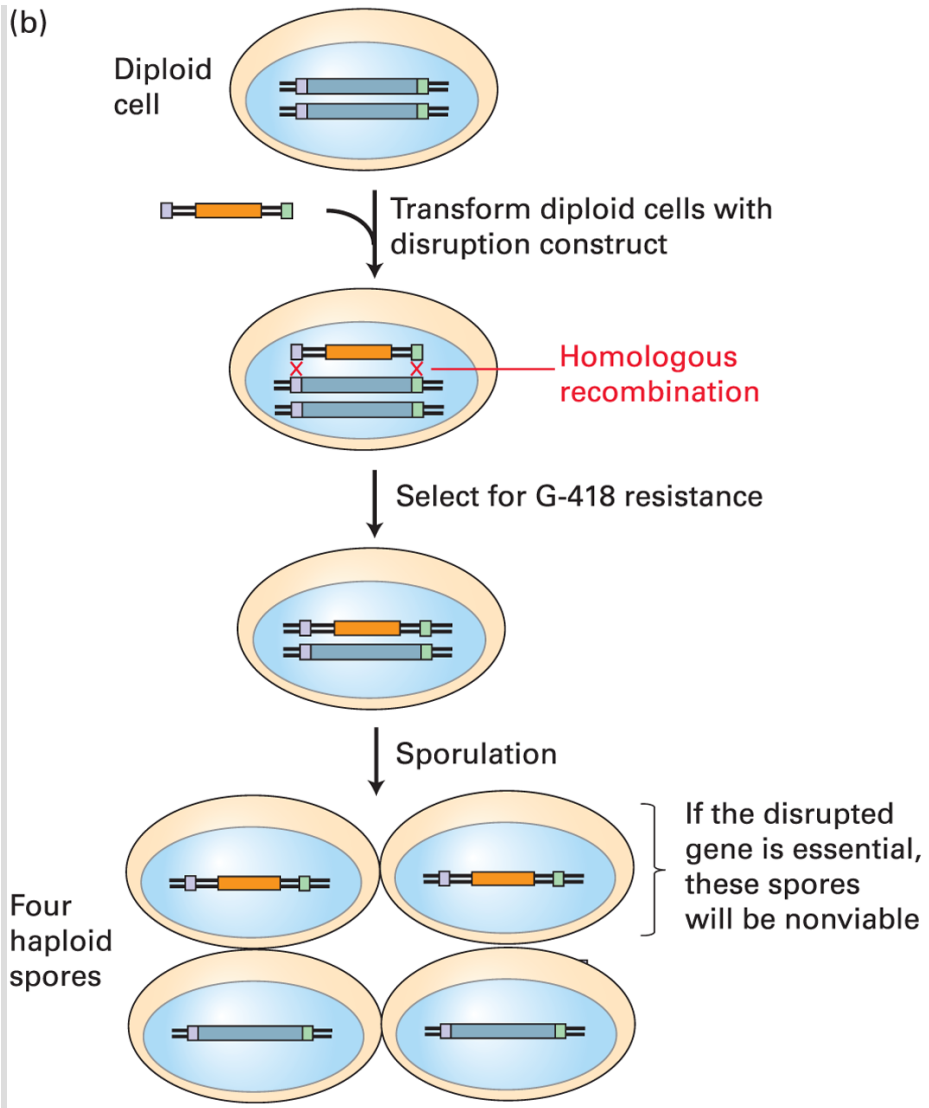
Normal Yeast Genes Can Be Replaced with Mutant Alleles by Homologous Recombination

Modifying the genome of the yeast *S. cerevisiae* is particularly easy for two reasons: yeast cells readily take up exogenous DNA under certain conditions, and the introduced DNA is efficiently exchanged for the homologous chromosomal site in the recipient cell. This specific, targeted recombination of identical stretches of DNA allows any gene in yeast chromosomes to be replaced with a mutant allele.

In one popular method for disrupting yeast genes in this fashion, a gene is replaced by a selectable marker. PCR is used to generate a *disruption construct* containing the selectable marker, which is

subsequently transfected into yeast cells. As shown in Figure 6-37a, the two primers for PCR amplification of the selectable marker are each designed to include about 20 nucleotides identical to sequences flanking the yeast gene to be replaced. The resulting amplified construct comprises the selectable marker (e.g., the *KanMX* gene, which, like neo^r, confers resistance to G-418) flanked by about 20 bp at each end that match the ends of the target yeast gene. Once transfected into a diploid yeast cell, the disruption construct will typically replace one of the two copies of the target endogenous gene. These transformed diploid yeast cells can be identified by their resistance to G-418 or other selectable phenotype. Heterozygous diploid yeast cells generally grow normally regardless of the function of the target gene, but half the haploid spores derived from these cells will carry only the disrupted allele (Figure 6-37b). If a gene is essential for viability, then spores carrying a disrupted allele will not survive.





EXPERIMENTAL FIGURE 6-37 Homologous recombination with transfected disruption constructs can inactivate specific target genes in yeast. (a) A suitable construct for disrupting a target gene can be prepared using PCR. The two primers designed for this purpose each contain a sequence of about 20 nucleotides (nt) that is homologous to one end of the target yeast gene as well as sequences needed to amplify a segment of DNA carrying a selectable marker gene such as *kanMX*, which confers resistance to G-418. (b) When recipient diploid *Saccharomyces* cells are transformed with the disruption construct, homologous recombination between the ends of the construct and the corresponding chromosomal sequences integrates the marker gene into the chromosome, replacing the target-gene sequence. The recombinant diploid cells will grow on a medium containing G-418, whereas untransformed cells will not. If the target gene is essential for viability, half the haploid spores that form after sporulation of recombinant diploid cells will be nonviable.

Disruption of genes by this method is proving particularly useful in assessing the roles of proteins identified by analysis of the entire genomic sequence of *S. cerevisiae*. Each of the approximately 6000 genes has been disrupted with the *KanMX* construct in diploids, and gene disruptions in haploid spores have also been produced. These analyses have shown that about 4500 of the 6000 yeast gene disruptions can reside in viable haploid spores, revealing an unexpectedly large number of apparently nonessential genes. In some cases, disruption of a particular gene may give rise to subtle defects that do not compromise the viability of yeast cells growing under laboratory conditions. Alternatively, cells carrying a disrupted gene may be viable because of the operation of backup or compensatory pathways. The existence of compensatory pathways and processes is revealed through the analysis of double mutants as shown in Figure 6-10.

The construction of *KanMX* deletion constructs for *S. cerevisiae* included the insertion of two sequences of 20 nucleotides (nt) that are unique to each deleted gene. These sequences can be used as molecular bar codes that can be read out either by DNA sequencing or by hybridization to a microarray. By counting the number of copies of a given bar code the growth properties of the corresponding deletion mutant can be followed in a pooled mixture of all 4500 nonessential gene deletion mutants. In an early application of the bar code technology for parallel screening of

deletion mutants, the pool of all 4500 deletion mutants was tested for sensitivity to various stress conditions. For example, the pool was evaluated by hybridization to a microarray to determine the relative numbers of each deletion strain. After this initial calibration the pool was placed in growth medium containing 1 M NaCl, a condition that causes stress and will slow the growth of wild-type yeast. Deletion mutants that are less able to withstand salt stress will exhibit decreased fitness relative to the average strain when grown in 1 M NaCl. From among 4500 deletion strains 62 gene deletions were identified that exhibited a significant decrease in cell number as determined by counting the number of the corresponding bar codes after growth in 1 M NaCl. The typical approach to conducting an equivalent screen would be to test the growth of each of the 4500 deletion mutants individually, whereas this method of following each strain by its molecular bar code enables growth characteristics of the entire collection of mutants to be screened in a single culture flask. This efficient method for screening mutants can be applied whenever individual mutant strains can be identified by a unique sequence tag, as we will see in the next section on mutations generated by the CRISPR system.

Engineered CRISPR Systems Allow Precise Genome Editing

Precise alterations can be made in genomic DNA sequences, a process known as *genome editing,* of virtually all organisms through the use of CRISPR. This genome-editing technique is an adaptation of a natural mechanism that evolved to protect bacterial cells against foreign DNA, such as phage DNA. The name **CRISPR** (*c*lustered *r*egularly *i*nterspaced *s*hort *p*alindromic *r*epeats) refers to the curious arrays of tandem repeated sequences found in about half of the bacterial genomes that have been sequenced. The arrays of repeated sequences are flanked by a conserved set of genes, known as *Cas* (CRISPR-*as*sociated) genes, that show similarity to genes encoding nucleases.

The breakthrough in understanding the function of CRISPR sequences came from the observation that one set of repeated sequence elements in the array often matched short segments in phage genomes. It was subsequently shown that the CRISPR element and associated Cas genes confer on bacterial cells the ability to cleave phage DNA at precisely the site that corresponds to repeated sequences in the CRISPR array. By this means the bacterial cells acquire immunity to a phage. The acquisition of immunity takes place in two stages. In the first stage, a phage-infected bacterium cleaves the phage DNA into short segments and adds those segments to the CRISPR array so that they are interspersed with highly conserved repeats. In the second stage, transcription of the CRISPR array and processing of the resulting RNA yields mature bipartite RNA molecules, known as guide RNAs, that carry both the

conserved repeat sequences and a phage-derived spacer sequence. The guide RNA assembles with Cas proteins to form an interference complex, which can target infecting phage DNA (that has been released from the viral particle) through base pairing of the guide RNA and with its complementary sequence in the phage DNA. Once targeted to a specific DNA sequence, nucleases in the Cas proteins cleave both strands of the target DNA molecule at a site adjacent to the region base-paired with the guide sequence.

Using CRISPR to Modify Eukaryotic Genomes

Although CRISPR elements have been found only in prokaryotic organisms, researchers proposed that the guide RNA and Cas proteins should be able to perform their functions if expressed in eukaryotic cells. To adapt CRISPR to function in virtually any cell type, researchers have developed minimal systems consisting of the nuclease Cas9 and an engineered guide RNA (Figure 6-38a). Cas9 contains the enzymatic activities necessary for precise DNA cleavage, including two separate endonuclease activities — one for each strand of DNA. The guide RNA is composed of two regions: the first is a scaffold made up of two complementary sequences that form a double-stranded hairpin structure that binds to Cas9, and the second region of about 20 nt identifies the target sequence and is designed to perfectly match a specific target site in genomic DNA. When both Cas9 and the guide RNA are expressed in a recipient cell, Cas9 will cleave both DNA strands at the chromosomal site specified by the guide RNA sequence (Figure 6-38b, c). Transfection with

expression plasmids for Cas9 and a specific guide RNA has been shown to give rise to specific DNA cleavages in cells derived from a variety of organisms, including *Drosophila, C. elegans*, zebrafish, mouse, rat, and human.

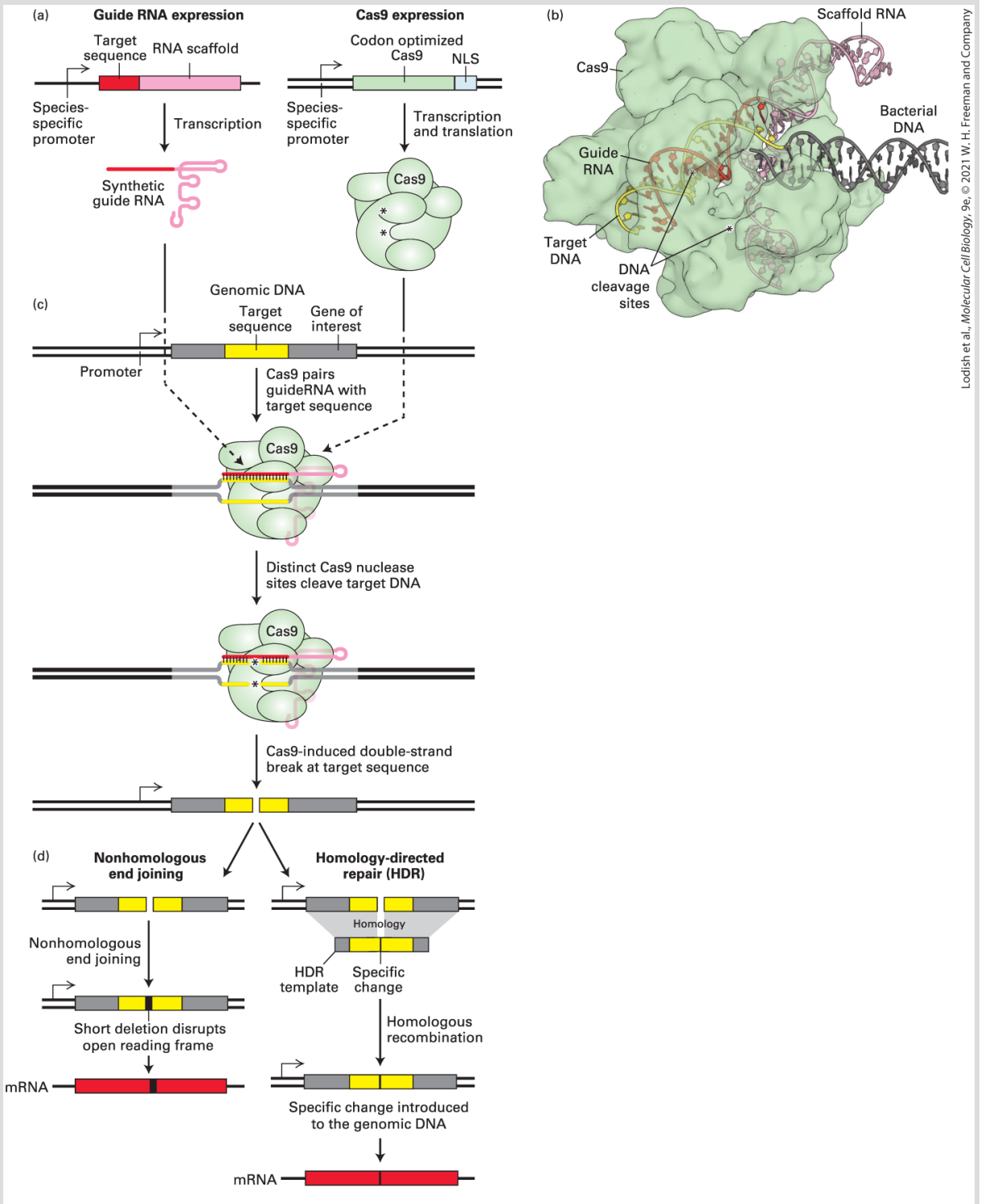


FIGURE 6-38 Single-nucleotide mutations can be introduced into the genome using an engineered CRISPR-Cas9 system. (a) The genome of a target cell can be modified by expression of the double-stranded DNA endonuclease Cas9 and a guide RNA. Expression of these components can be achieved by transfection with plasmids carrying genes for Cas9

and the guide RNA or by direct injection of Cas9 mRNA and guide RNA. The guide RNA is composed of two parts: a sequence that folds into a hairpin scaffold structure that binds to Cas9, and a sequence of approximately 20 nt corresponding to the targeted site in the genome. (b) A complex of guide RNA bound to Cas9 is targeted to the genome by base pairing of the guide RNA with the complementary genomic DNA sequence. This structure allows the two distinct nuclease active sites of Cas9 to cleave both strands of the target DNA adjacent to the heteroduplex formed with the guide RNA. (c) By this mechanism, the expression of both Cas9 and a bipartite guide RNA designed to target a specific gene sequence leads to a double-strand cleavage of the target gene. (d) Cleaved DNA can be repaired via a nonhomologous end joining (NHEJ) process, which usually removes a small number of bases at the cleavage site. If the cleavage occurs in a coding sequence, NHEJ will usually inactivate gene function by producing a frameshift mutation. If an ~100-nt single-stranded DNA segment that spans the sequences flanking the cleavage site is injected along with Cas9 mRNA and the guide RNA, the cleaved DNA can be repaired by homologous recombination (homology-directed repair, HDR). By this mechanism, single base changes can be introduced into the repaired genomic DNA.

[Part (b) data from C. Anders et al., 2014, *Nature* **513**:569–573, PDB ID 4un3.]

A particularly efficient way to modify the germ line of a mouse begins with the microinjection of Cas9 mRNA and a guide RNA into a mouse zygote. The resulting double-strand break in the specific target sequence is typically repaired by a set of enzymes that ligate the free DNA ends back together in a process known as nonhomologous end joining (Figure 6-40d). Typically, a few base pairs are deleted at the site of cleavage because nucleases remove bases from the free DNA ends before the end joining reaction is completed. If the cleavage site is within a gene-coding sequence, the deletion of several bases will usually cause a frameshift mutation, thus producing a loss of function mutation that inactivates the targeted gene.

An important refinement of this method of genome editing allows exact single-base changes in the target DNA to be produced. This method first requires targeting Cas9 cleavage close to the site of the desired single-base change. With the addition of a segment of DNA, typically about 100-nt long, that matches the sequences flanking the cleavage site and contains the desired single base mutation, the free ends can be repaired by homologous recombination using the added homologous DNA segment, leaving the single-base mutation in the repaired DNA (see Figure 6-40d). In a dramatic demonstration of this genome editing method, a single-base mutation was corrected in *Crygc*, a gene for gamma-crystallin that leads to the formation of cataracts. The mutated allele was corrected back to the wild-type form by injection of a mutant mouse zygote with Cas9 mRNA, a guide RNA that directs cleavage at the site of the mutation, and a 90-nt DNA segment spanning the cleavage site and containing the wild-type *Crygc* sequence. This experiment also illustrates the remarkable selectivity of CRISPR-Cas9 targeting, since the guide RNA was able to discriminate between the mutant and wild-type alleles even though they differed by only one base pair.

Genome-Wide CRISPR Screens

The ability of CRISPR genome editing to produce genetically modified mammals, including humans, has sparked both public interest and controversy. But from the perspective of cell biologists hoping to learn more about cell function, genome editing by CRISPR

of somatic cells grown in culture is a powerful (and not at all controversial) research tool for exploring the function of mammalian genes in the context of a living cell. One of the most significant extensions of genome editing to somatic cells is the ability to test in a single experiment the effects of loss of function mutations in each of approximately 20,000 mammalian genes. One way to do this would be to transfect cultured cells with the gene for Cas9 and a gene for a guide RNA specifically targeting an exon of a selected gene. To test for the effects of loss of function of each of 20,000 mammalian genes would require 20,000 different transfections and 20,000 different phenotypic assays of the CRISPR edited cell lines. As was the case for screening the collection of yeast gene deletions, a much more efficient way to conduct such a screen would be to pool all 20,000 transfected cell lines and then to use the sequence of the guide RNA as a unique molecular tag to follow the growth of each individual cell line.

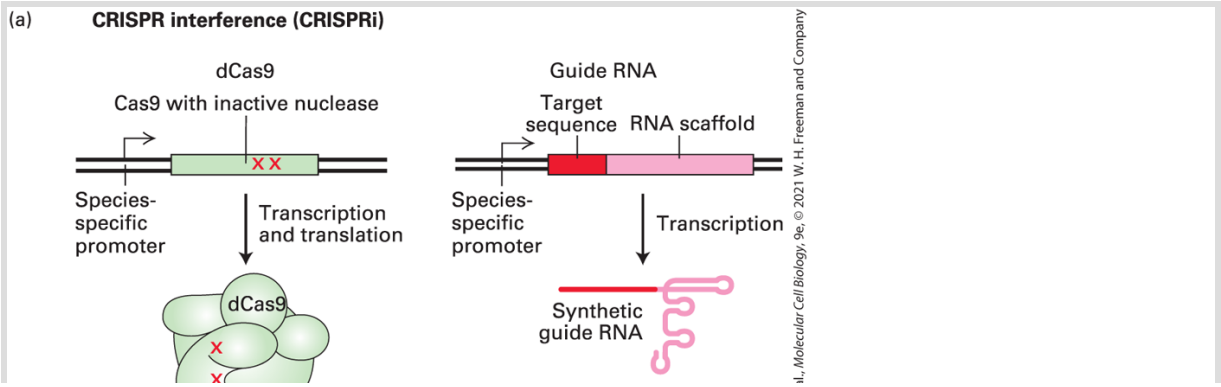
To see how such a pooled screen would work, suppose we wanted to identify genes that, when mutated, conferred resistance to Taxol, an anticancer drug which interferes with the function of microtubules during cell division. First, a collection of cells transfected with genes for Cas9 and with a guide RNA for each of 20,000 different genes is generated as a large pool. Next DNA isolated from the pool is subjected to large-scale sequencing of the transfected guide RNA genes such that $>10^6$ separate guide RNA gene sequences are determined from the pool. The relative number of reads corresponding to each specific guide RNA will give a count of the

relative numbers of cells carrying each guide RNA and provide a baseline for the experiment. After pooled cells have been exposed to Taxol for a period of time long enough for the number of Taxol-sensitive cells to decline substantially (e.g., 36 hours), the guide RNA genes in the pool are sequenced again. The typical cell line would decrease in relative number compared to a baseline determined at the start of the experiment because sensitivity to Taxol would have slowed cell growth. However, the relative number for some cell lines would increase after Taxol treatment. These would contain guide RNA genes that promote mutations that increase resistance to Taxol. Some cell lines would show an even greater decrease in relative number relative to the average cell line; this would indicate that the corresponding guide RNA genes promoted mutations that cause hypersensitivity to Taxol. Thus in a single large experiment, the effect of CRISPR inactivation of every gene in the mammalian genome can be tested for increased resistance or sensitivity to Taxol — an important advance in identifying all of the genes that may play a role in modulating the efficacy of Taxol as an anticancer agent.

Inducible Activation of CRISPR

A significant limitation of genome editing using CRISPR, which generally produces loss of function mutations, is that mutations in essential genes are difficult to study. Moreover, for some genes, it is often desirable to study the effects of gain-of-function alleles. Examples are the genes for proteins involved in signal transduction or gene regulation. In such cases, it is desirable to transiently

increase or decrease the expression of a given gene rather than permanently inactivate it by genome editing. The CRISPR system has been adapted for this purpose through the development of two different systems, one for transiently decreasing gene expression and one for increasing it. One system, known as CRISPRi, targets Cas9 to the promoter so as to interfere with transcription initiation, whereas a complementary gene-activating system, known as CRISPRa, activates transcription of a gene of interest by delivering a transcriptional activation domain to its promoter (Figure 6-39). Both CRISPRi and CRISPRa employ a modified version of Cas9 with the endonuclease activity inactivated (this nuclease dead form is called *dCas9*). When dCas9 is targeted to the promoter region of a gene by transiently expressing an appropriate guide RNA, dCas9 will bind to the promoter without cleaving the DNA. For CRISPRi, dCas9 is targeted to a region of the promoter that will directly compete with the ability of RNA polymerase to initiate transcription and thus will transiently interfere with gene expression. For CRISPRa, dCas9 is fused to a powerful transcription activation domain that will cause a transient activation of gene expression when this hybrid protein is targeted to a promoter.



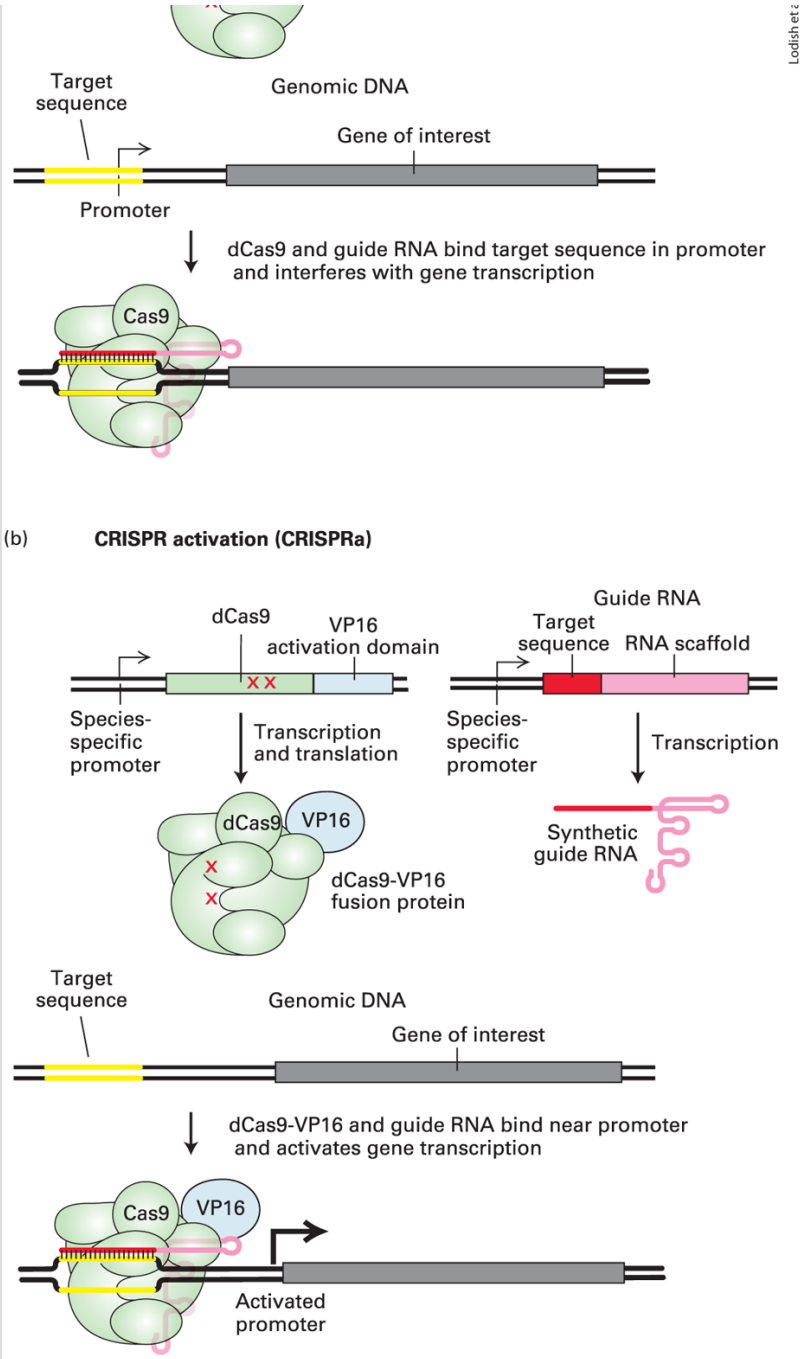


FIGURE 6-39 CRISPR-Cas9 engineered to produce transient inactivation or hyperactivation of a gene. (a) CRISPR interference (CRISPRi) can be used to inactivate transcription of a specific gene. For CRISPRi, a synthetic guide RNA is designed to target the RNA polymerase binding region within the promoter for a gene of interest. When this guide RNA is expressed together with a mutated form of Cas9 that lacks endonuclease activity (dCas9), the guide RNA will direct dCas9 to bind to the promoter, preventing RNA

polymerase from binding and thus blocking transcription of the gene. (b) CRISPRa employs a similar strategy except that the region targeted by the guide RNA is upstream of the promoter of interest and dCas9 is fused to a powerful transcription activation domain (dCas9-VP16). When dCas9-VP16 is targeted to the promoter region of the gene of interest the activation domain will activate transcription of that gene.

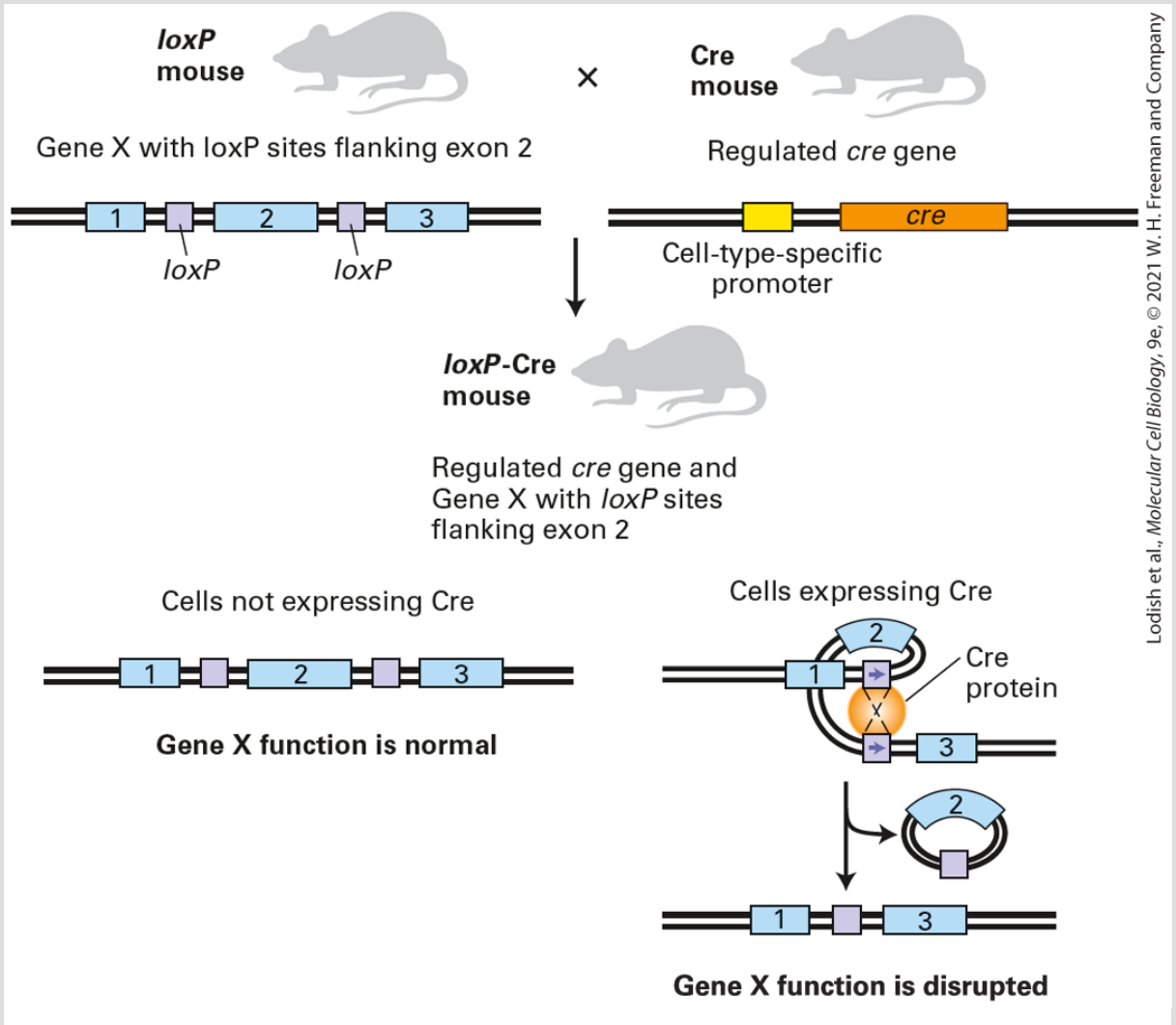
Somatic Cell Recombination Can Inactivate Genes in Specific Tissues of Mice

The same principles applied to genetically modify yeast by homologous recombination can be applied to mammalian genes. The use of homologous recombination to replace a normal mouse gene by a gene knockout has been a historically important way to construct null alleles of mouse genes, but these methods have been largely replaced by the powerful CRISPR-Cas9 genome editing methods. However, investigators are often interested in examining the effects of knockout mutations in a particular tissue of the mouse, at a specific stage in development. The problem is that mice carrying a germ-line mutation may have defects in numerous tissues or die before the developmental stage of interest. To address this problem, mouse geneticists have devised a clever technique to inactivate target genes in specific types of somatic cells or at particular times during development.

In the first stage of this procedure, a DNA construct containing a specifically modified allele of a particular target gene is introduced into embryonic stem (ES) cells. These cells, which are derived from the blastocyst, can be grown in culture through many generations (see Figure 22-6). In a small fraction of transfected cells, the introduced DNA undergoes homologous recombination with the target gene, although recombination at nonhomologous chromosomal sites occurs much more frequently. To enable selection for cells in which the construct allele has successfully replaced the target gene, the recombinant DNA construct introduced into ES cells includes the selectable marker (neo^r), which confers resistance to G-418. To produce mice carrying this construct in their genome, modified ES cells are injected into a recipient wild-type mouse blastocyst, which subsequently is transferred into a pseudopregnant female mouse. The resulting progeny will be chimeras, containing tissues derived from both the transplanted ES cells and the host cells. Chimeric mice are then mated with mice that are homozygous for another allele of the marker trait to determine if the knockout mutation has been incorporated into the germ line. Finally, mating of mice, each heterozygous for the knockout allele, will produce progeny homozygous for the knockout mutation.

Homologous recombination can also be used to produce a mouse in which the gene of interest has been inactivated in only a single tissue. This technique employs short DNA recombination sites

(called loxP *sites*) and the enzyme Cre that catalyzes recombination between them. The loxP-*Cre recombination system* is derived from bacteriophage P1, but this site-specific recombination system also functions when placed in mouse cells. An essential feature of this technique is that expression of Cre is controlled by a promoter specific to a single cell type. In *loxP*-Cre mice generated by the procedure depicted in Figure 6-40, inactivation of the gene of interest (X) occurs only in cells in which the promoter controlling the *cre* gene is active. By using different promoters, researchers can study the effects of knocking out gene X in various types of cells.



EXPERIMENTAL FIGURE 6-40 The *loxP*-Cre recombination system can knock out genes in specific cell types. The outlined protocol is designed to produce a knockout of a target gene (GeneX) targeted to a specific cell type. A a *loxP* mouse contains the conditional gene X knockout generated by inserting *loxP* sites (purple) on each side of an essential exon (exon 2) of the target gene (blue). Since the *loxP* sites are in introns, they do not disrupt the function of gene X. The Cre mouse carries the *cre* gene (orange) from bacteriophage P1 expressed from a mouse promoter specific for the desired cell-type (yellow). The *cre* gene is incorporated into the mouse genome by nonhomologous recombination and does not affect the function of endogenous mouse genes. The Cre mouse is also heterozygous for a gene X knockout allele. Crossing these two strains generates among the progeny *loxP*-Cre mice that contain as the only functional copy of gene X the conditional knockout construct as well as having regulated expression of cre. Cre protein produced only in those cells in which the

promoter is active will catalyze recombination between the *loxP* sites, leading to deletion of exon 2 and a complete loss of function of gene X in all cells expressing Cre.

An early application of this technique provided strong evidence that a particular neurotransmitter receptor is important for learning and memory. Previous pharmacological and physiological studies had indicated that normal learning requires the NMDA class of glutamate receptors in the hippocampus, a region of the brain. But mice in which the gene encoding an NMDA receptor subunit was knocked out died shortly after birth, precluding analysis of the receptor's role in learning. Following the protocol in Figure 6-40, researchers generated mice in which the receptor subunit gene was inactivated in the hippocampus but expressed in other tissues. These mice survived to adulthood and showed learning and memory defects, confirming a role for these receptors in the ability of mice to encode their experiences into memory.

RNA Interference Causes Gene Inactivation by Destroying the Corresponding mRNA

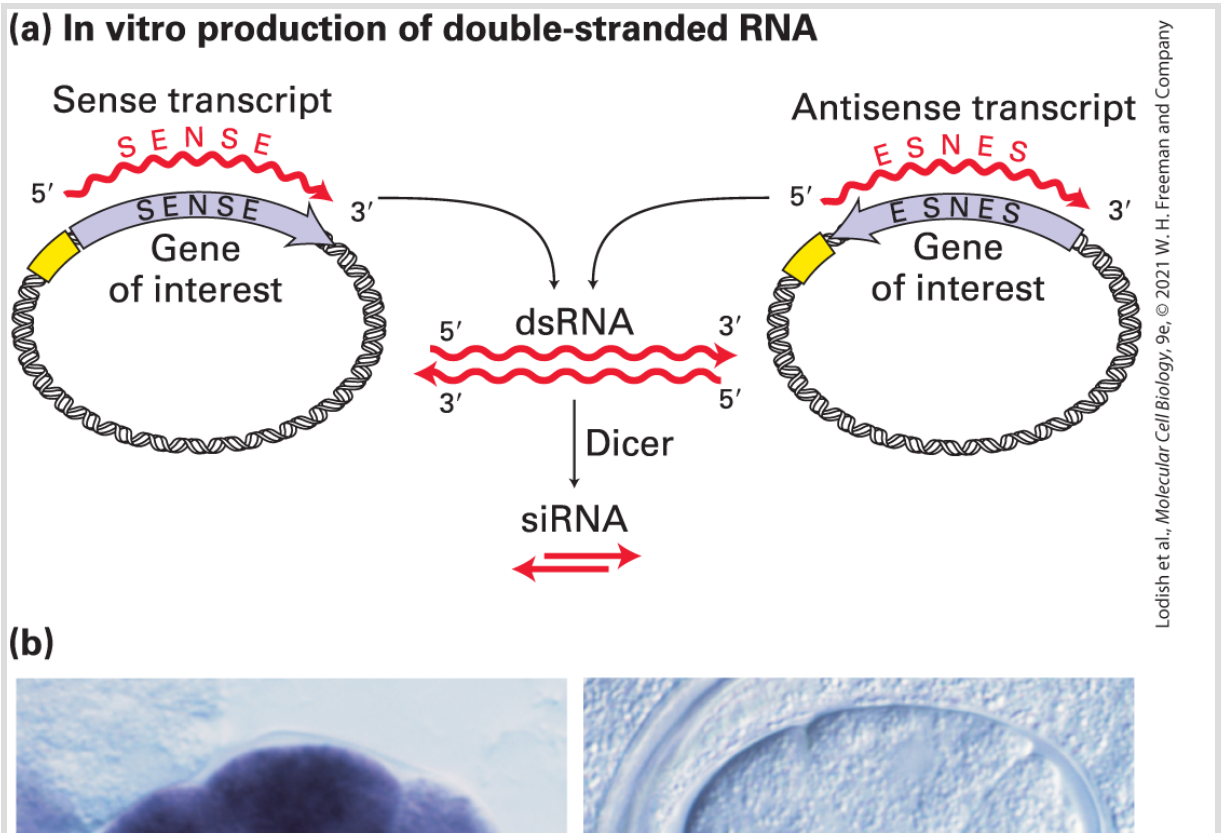
First observed in the roundworm *C. elegans*, **RNA interference (RNAi)** refers to the ability of double-stranded RNA to block expression of its corresponding single-stranded mRNA, but not that of mRNAs with a different sequence. The phenomenon provides a

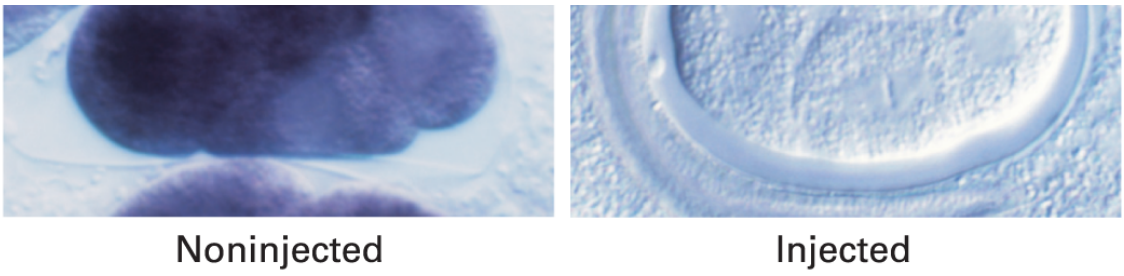
straightforward method to inhibit transiently the function of specific genes.

As described in Chapter 9, the phenomenon of RNAi rests on the ability of eukaryotic cells to cleave double-stranded RNA into short (23-nt) double-stranded segments known as small inhibitory RNAs (siRNAs). The RNA endonuclease that catalyzes this reaction, known as Dicer, is found in all metazoans, but not in simpler eukaryotes such as yeast. The siRNA molecules, in turn, can cause cleavage of mRNA molecules of matching sequence, in a reaction catalyzed by a protein complex known as *RISC*. Attached to RISC, one strand of the siRNA recognizes and hybridizes with its complementary sequence on the target mRNA; subsequently, specific nucleases in the RISC complex cleave the mRNA-siRNA hybrid. This model accounts for the specificity of RNAi, since it depends on base pairing, and for its potency in silencing gene function, since the complementary mRNA is permanently destroyed by nucleolytic degradation. The normal function of both Dicer and RISC is to allow for gene regulation by small endogenous RNA molecules known as *microRNAs*, or *miRNAs*.

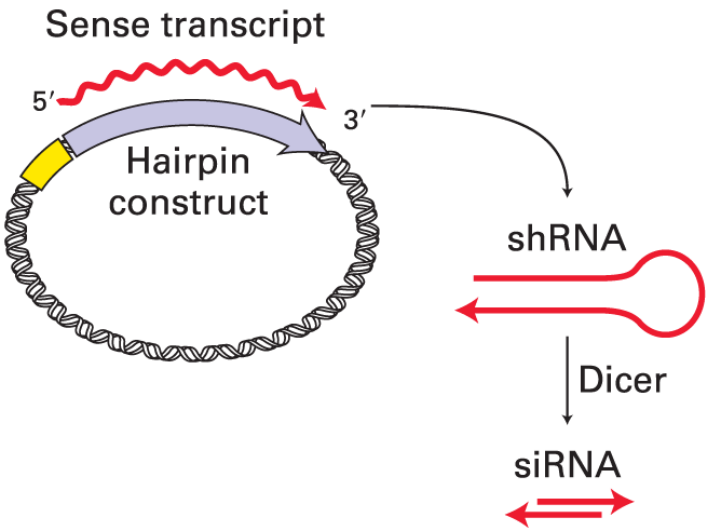
Researchers can exploit the microRNA pathway to intentionally silence a *C. elegans* gene of interest by using either of two general methods for generating siRNAs of defined sequence. In the first method, a double-stranded RNA corresponding to the target-gene sequence is produced by in vitro transcription of both sense and antisense copies of that sequence (Figure 6-41a). This dsRNA is then

injected into the gonad of an adult worm, where it is converted to siRNA by Dicer in the developing embryos. In conjunction with the RISC complex, the siRNA molecules cause the corresponding mRNA molecules to be destroyed rapidly. The resulting worms display a phenotype similar to the one that would result from disruption of the target gene itself. In some cases, insertion of just a few molecules of a particular dsRNA into a cell is sufficient to inactivate many copies of the corresponding mRNA. Figure 6-41b illustrates the ability of an injected dsRNA to interfere with the production of the corresponding endogenous mRNA in *C. elegans* embryos. In this experiment, the mRNA levels in embryos were determined by in situ hybridization, as described earlier, using a fluorescently labeled probe.





(c) In vivo production of double-stranded RNA



EXPERIMENTAL FIGURE 6-41 RNA interference (RNAi) can inhibit gene function in *C. elegans* and other organisms. (a) In vitro production of double-stranded RNA (dsRNA) for interference with a specific target gene. The coding sequence of the gene, derived from either a cDNA clone or a segment of genomic DNA, is placed in two orientations in a plasmid vector adjacent to a strong promoter. Transcription of both constructs in vitro using RNA polymerase and ribonucleoside triphosphates yields many RNA copies in both the sense orientation (identical to the mRNA sequence) and the complementary antisense orientation. Under suitable conditions, these complementary RNA molecules hybridize to form dsRNA. When the dsRNA is injected into cells, it is cleaved by Dicer into siRNAs. (b) Inhibition of *mex3* RNA expression in *C. elegans* embryos by RNAi (see the text for the mechanism). Expression of *mex3* RNA in embryos was assayed by in situ hybridization to a probe specific for this mRNA, linked to an enzyme that produces a colored (purple) product. (*Left*) Wild-type embryo. (*Right*) Embryo derived from a worm injected with double-stranded *mex3* RNA. Each four-cell-stage embryo is ~50 mm in length. (c) In vivo production of double-stranded RNA via an engineered plasmid introduced directly into cells. The synthetic gene construct is a tandem arrangement of both sense and antisense sequences of the target gene. When it is

transcribed, double-stranded small hairpin RNA (shRNA) forms. The shRNA is cleaved by Dicer to form siRNA.

[Reprinted by permission from Macmillan Publishers Ltd, from A. Fire, 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**(6669):806–811; permission conveyed through Copyright Clearance Center, Inc.]

The second method is to produce a specific double-stranded RNA in vivo. An efficient way to do this is to express a synthetic gene that is designed to contain tandem segments of both sense and antisense sequences corresponding to the target gene (Figure 6-41c). When this gene is transcribed, a double-stranded RNA hairpin structure forms, known as *small hairpin RNA*, or *shRNA*. The shRNA will then be cleaved by Dicer to form siRNA molecules. The lentiviral expression vectors are particularly useful for introducing synthetic genes for the expression of shRNA constructs into animal cells.

Both RNAi methods lend themselves to systematic studies in which researchers inactivate each of the known genes in an organism and observe what goes wrong. For example, in initial studies with *C. elegans*, RNA interference with 16,700 genes (about 86 percent of the genome) yielded 1722 visibly abnormal phenotypes. The genes whose functional inactivation causes particular abnormal phenotypes can be grouped into sets; each member of a set presumably controls the same signals or events. The regulatory relations between the genes in a set — for example, the genes that control muscle development — can then be worked out.

RNAi-mediated gene inactivation has been successful in many organisms, including *Drosophila*, many kinds of plants, zebrafish, the frog *Xenopus*, and mice, and these organisms are now the subjects of large-scale RNAi screens. For example, lentiviral vectors have been designed to inactivate by RNAi more than 10,000 different genes expressed in cultured mammalian cells. The functions of the inactivated genes can be inferred from defects in the growth or morphology of cell clones transfected with lentiviral vectors.

KEY CONCEPTS OF SECTION 6.6

Altering the Function of Specific Genes by Design

- Once a gene has been cloned, important clues about its normal function in vivo can be deduced from the observed phenotypic effects of mutating the gene.
- Genes can be disrupted in yeast by inserting a selectable marker gene into one copy of a wild-type gene via homologous recombination, producing a heterozygous mutant. When such a heterozygote forms spores, two nonviable haploid spores will be produced (see Figure 6-37).
- Each yeast gene deletion has been tagged with a unique DNA sequence that can serve as a molecular bar code that allows the growth of individual deletion strains to be traced in a pooled population containing all of the deletion strains.
- A bacterial system that evolved to precisely target and cleave foreign DNA, known as CRISPR, has been adapted for introducing specific changes in the genomic DNA of many organisms. Cleavage of chromosomal DNA at a specific site by CRISPR-Cas9 usually results in a short deletion at the cleavage site. If an appropriately designed DNA segment is provided by transfection, specific DNA changes, such as point mutations, can be introduced at the cleavage site (see Figure 6-39).
- All genes in the mammalian genome can be disrupted by CRISPR-Cas9 genome editing, and the growth properties of specific gene disruptions in a pooled population can be traced.
- Cas9 protein with an inactivated nuclease can be used for the transient interference (CRISPRi) or activation (CRISPRa) of a targeted gene.

- In mice, modified genes can be incorporated into the germ line at their original genomic location by homologous recombination. The *loxP*-Cre recombination system permits production of mice in which a gene is knocked out in a specific tissue.
- In many organisms, including the roundworm *C. elegans*, double-stranded RNA triggers destruction of the all the mRNA molecules with the same sequence. This phenomenon, known as RNAi (RNA interference), provides a specific and potent means of inactivating genes without altering their structure.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

alleles
clone
complementary DNA (cDNA)
conservation
CRISPR
DNA cloning
DNA library
DNA microarray
dominant
functional complementation
genome-wide association study (GWAS)
genotype

heterozygous
homozygous
hybridization
in situ hybridization
linkage
mutagen
mutation
open reading frames (ORFs)
phenotype
plasmids
point mutation
polymerase chain reaction (PCR)
recessive
recombinant DNA
recombination
restriction enzymes
RNA interference (RNAi)
RNA-Seq
segregation
single cell RNA-Seq (scRNA-Seq)
structural motifs
transfection
transformation
vector
whole transcriptome shotgun sequencing (RNA-Seq)
wild type

Review the Concepts

1. Genetic mutations can provide insights into the mechanisms of complex cellular or developmental processes. How might your analysis of a genetic mutation be different depending on whether a particular mutation is recessive or dominant?
2. What is a temperature-sensitive mutation? Why are temperature-sensitive mutations useful for uncovering the function of a gene?
3. Describe how complementation analysis can be used to reveal whether two mutations are in the same or in different genes. Explain why complementation analysis will not work with dominant mutations.
4. Restriction enzymes and DNA ligase play essential roles in DNA cloning. How is it that a bacterium that produces a restriction enzyme does not cut its own DNA? Describe some general features of restriction sites. What are the three types of DNA ends that can be generated after cutting DNA with restriction enzymes? What reaction is catalyzed by DNA ligase?
5. Bacterial plasmids often serve as cloning vectors. Describe the essential features of a plasmid vector. What are the advantages and applications of plasmids as cloning vectors?
6. A DNA library is a collection of clones, each containing a different fragment of DNA, inserted into a cloning vector. What is the difference between a cDNA library and a genomic DNA library? Suppose you would like to clone

gene X, a gene expressed only in neurons, into a vector using a library as the source of the insert. You have the following libraries at your disposal: a genomic library from skin cells, a cDNA library from skin cells, a genomic library from neurons, and a cDNA library from neurons. Which could you use, and why?

7. In 1993, Kary Mullis won the Nobel Prize in Chemistry for his invention of the PCR process. Describe the three steps in each cycle of a PCR reaction. Why was the discovery of a thermostable DNA polymerase so important for the development of PCR?
8. A number of foreign proteins have been expressed in bacterial and mammalian cells. Describe the essential features of a recombinant plasmid that are required for expression of a foreign gene. How can the foreign protein be modified to facilitate its purification? What is the advantage of expressing a protein in mammalian cells versus expressing the same protein in bacteria?
9. RT-PCR and microarrays can be used to analyze gene expression. A lab studies yeast cells, comparing their growth in two different sugars, glucose and galactose. One student is comparing expression of the gene *HMG2* under these two conditions. Which technique(s) could he use and why? Another student wants to compare expression of all the genes on chromosome 4, of which there are approximately 800. What technique(s) could she use and why?

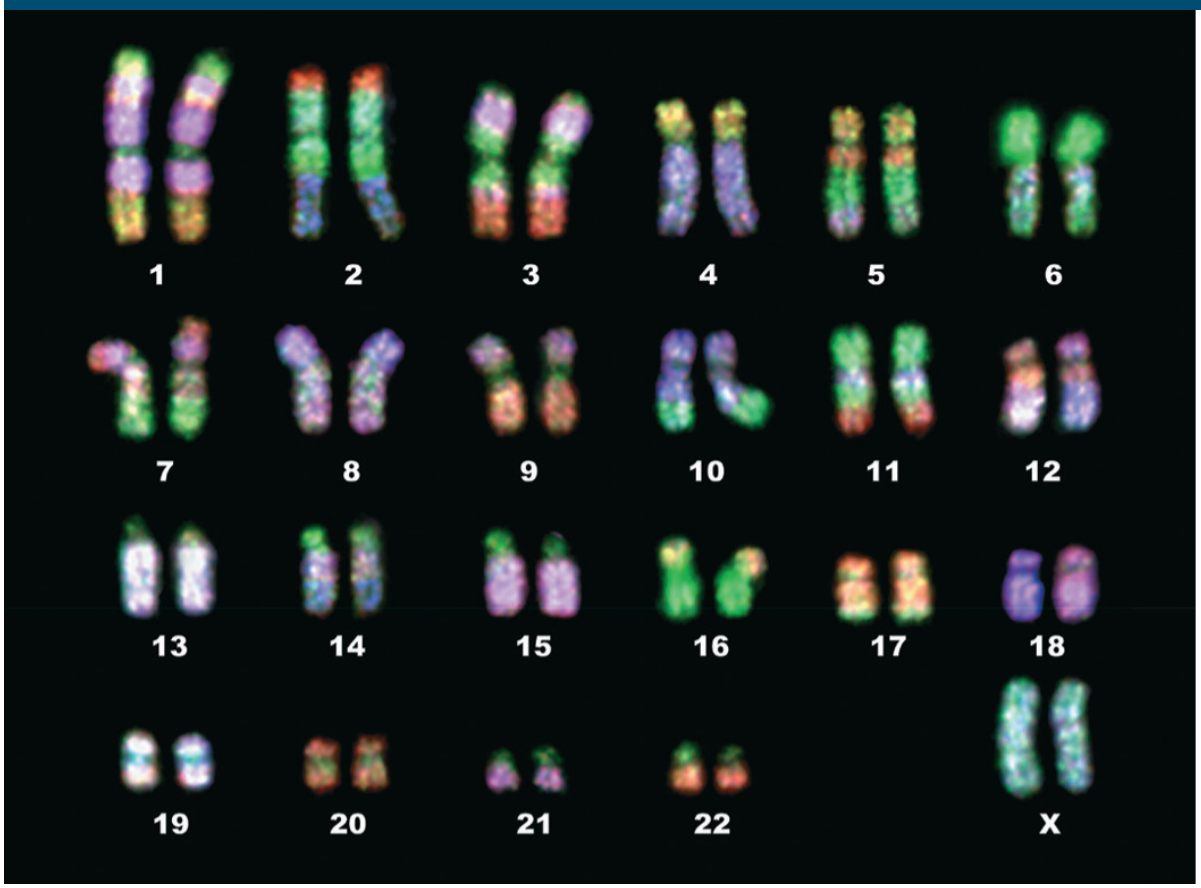
10. What are paralogous and orthologous genes? What are some of the explanations for the finding that humans are a much more complex organism than the roundworm *C. elegans*, yet have only 5 percent more protein-coding genes (21,000 versus 20,000)?
11. In determining the identity of the protein that corresponds to a newly discovered gene, it often helps to know the pattern of tissue expression for that gene. For example, researchers have found that a gene called *SERPINA6* is expressed in the liver, kidney, and pancreas but not in other tissues. What techniques might researchers use to find out which tissues express a particular gene?
12. DNA polymorphisms can be used as DNA markers. Describe the differences between SNPs and STR polymorphisms. How can these markers be used for DNA-mapping studies?
13. How can linkage-disequilibrium mapping sometimes provide a much higher resolution of gene location than classical linkage mapping?
14. Genetic linkage studies can usually locate the chromosomal position of a disease gene only roughly. How can expression analysis and DNA sequence analysis help locate a disease gene within the region identified by linkage mapping?
15. The ability to selectively modify the genome in the mouse has revolutionized mouse genetics. Outline the procedure for generating a knockout mouse at a specific genetic locus. How can the *loxP*-Cre system be used to conditionally

knock out a gene? What is an important medical application of knockout mice?

16. Two methods for functionally inactivating a gene without altering the gene sequence involve dominant-negative alleles and RNA interference (RNAi). Describe how each method can inhibit expression of a gene.

CHAPTER 7

Genes, Chromatin, and Chromosomes



L. Willatt/Science Source.

These brightly colored FISH-painted chromosomes are not only beautiful, but also useful in revealing chromosome anomalies and in comparing karyotypes of different species.

OUTLINE

7.1 Eukaryotic Gene Structure and Organization

7.2 Chromosomal Organization of Genes and Noncoding DNA
7.3 Transposable (Mobile) DNA Elements
7.4 Structural Organization of Eukaryotic Chromatin and Chromosomes
7.5 Morphology and Functional Elements of Eukaryotic Chromosomes

In previous chapters, we learned how the structure and composition of proteins allow them to perform a wide variety of cellular functions. We also examined another vital component of cells, nucleic acids, and the process by which information encoded in the sequence of DNA is translated into protein. In this chapter, our focus again is on DNA and proteins as we consider the characteristics of eukaryotic nuclear genomes. The chapter is organized around two major topics: (1) the organization and evolution of genetic information and (2) the structure and function of chromatin and chromosomes; the former being the complex of nuclear DNA and abundant nuclear proteins, and the latter being those structures that condense and become visible by light microscopy during mitosis.

By the beginning of the twenty-first century, biologists had complete genome sequences of hundreds of viruses, scores of bacteria, and one unicellular eukaryote, the budding yeast *Saccharomyces cerevisiae*. Today the vast majority of the genomic sequence is also known for the fission yeast *Schizosaccharomyces pombe*; many other fungi; the model plant organism *Arabadopsis thaliana*; multiple other plants; multicellular animals (metazoans), including the roundworm *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, mice,

humans, and hundreds of other vertebrates and invertebrates, with at least one representative from each of the 35 or so metazoan phyla. Detailed analysis of these sequencing data has led to insights into evolution, genome organization, and gene function. It has also allowed researchers to identify previously unknown genes and to estimate the total number of protein-coding genes in each of the sequenced genomes. Comparisons among gene sequences often provide insight into possible functions of newly identified genes.

It was a surprise to many researchers when genomic sequencing revealed that large portions of the genomes of metazoans and plants do not encode mRNAs or any other RNAs required by the organism. Remarkably, about 98.5 percent of human chromosomal DNA is noncoding DNA! The noncoding DNA includes transcription regulatory sequences recognized and bound by proteins that regulate transcription of genes within tens to hundreds of kb away in the linear DNA sequence. However, the vast majority of noncoding DNA in multicellular organisms includes many regions that do not seem to be involved directly in gene control or DNA replication. A large fraction of noncoding DNA in the genomes of individual organisms (~50 percent for humans) includes many regions that are similar but not identical in sequence to one another. There is enough variation within this *repetitive DNA* among individuals that every person can be distinguished by a unique DNA *fingerprint* based on these sequence variations. Moreover, the location in the genome of some repetitive DNA sequences varies

among different individuals of the same species. At one time, all noncoding DNA was collectively termed *junk DNA* and was considered to serve no purpose. We now understand the evolutionary basis of this noncoding DNA and its variation in location among individuals. Cellular genomes harbor transposable (mobile) DNA elements that can copy themselves and move to new locations throughout the genome. Although most transposable DNA elements seem to have little function in the life cycle of an individual organism, over evolutionary time they have helped to shape our genomes and contributed to the rapid evolution of multicellular organisms.

A eukaryotic genome is made up of very long molecules of DNA. For example, the 46 chromosomes in a human cell contain about 2 m of DNA (which is the height of someone six and a half feet tall). All of this DNA must be contained within a nucleus whose diameter is less than 20 μm; a compaction ratio of greater than 100,000 to 1. Specialized proteins associated with nuclear DNA fold and organize the DNA so that it fits into a nucleus. And yet, at the same time, any given portion of this highly compacted DNA can be accessed readily for transcription, replication, and repair of damage without the long DNA molecules becoming tangled or broken. Furthermore, the integrity of DNA must be maintained during the process of cell division when the entire genome is replicated and partitioned into daughter cells. In eukaryotes, the complex of DNA and the proteins that bind and organize the DNA is called *chromatin.* During mitosis,

chromatin can be visualized by light microscopy as individual chromosomes. As we will see in this and the following chapter, the organization of DNA into chromatin enables mechanisms of gene regulation that are not available in bacteria.

In the first three sections of this chapter, we walk through the landscape of eukaryotic genes and genomes. First we discuss the structure of eukaryotic genes and the complexities that arise in multicellular organisms from processing mRNA precursors into alternatively spliced mRNAs. Next we discuss the main sequence classes of eukaryotic DNA, including the special properties of transposable DNA elements and how they have shaped contemporary genomes. The final two sections of the chapter address how DNA is physically organized in eukaryotic cells. We consider how DNA and histone proteins are packaged into the compact complexes called *nucleosomes*, we describe the large-scale structure of chromosomes, and we look at the functional elements required for chromosome duplication and segregation. Figure 7-1 provides an overview of these interrelated subjects. The understanding of genes, chromatin, and chromosomes gained in this chapter will prepare you to explore in the following two chapters how the synthesis and amount of each protein and functional RNA in a cell is regulated.

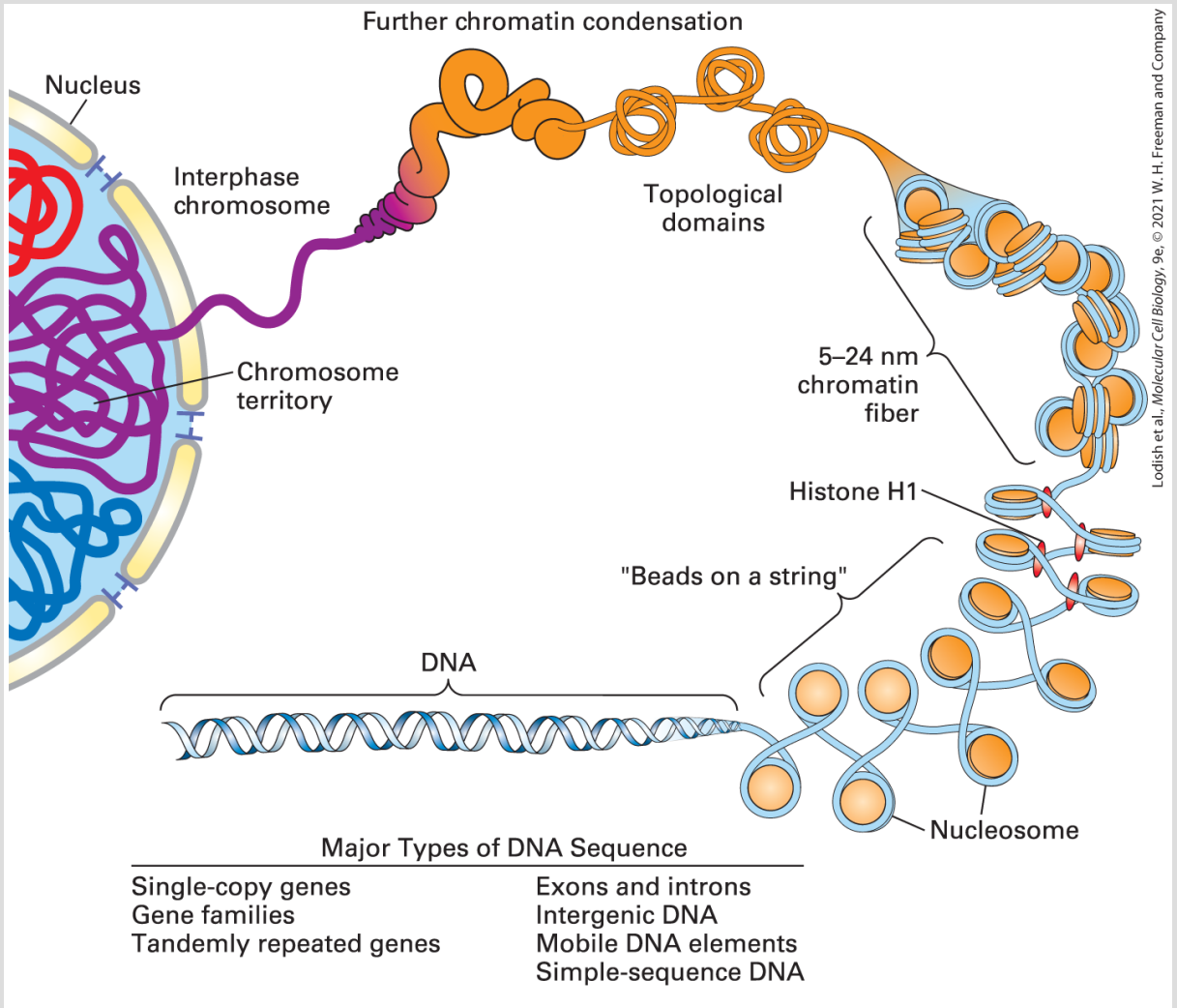


FIGURE 7-1 Overview of the structure of genes and chromosomes. DNA of multicellular eukaryotes consists of both unique and repeated sequences. Each chromosome consists of a single, long molecule of DNA (as long as 280 Mb in humans), organized into nucleosomes that fold and interact to generate various levels of chromatin condensation. This compaction process is carried out by the histone and nonhistone proteins that interact with the DNA molecule. The combination of a DNA molecule, histone proteins, and nonhistone proteins is called chromatin. Each chromosome occupies its own *territory* in the nucleus (note the blue, purple, and red chromosomes visible in the nucleus).

7.1 Eukaryotic Gene Structure and Organization

In molecular terms, a *gene* is commonly defined as *the entire nucleic acid sequence that is necessary for the synthesis of a functional gene product (polypeptide or RNA)*. According to this definition, a gene includes more than the nucleotides encoding an amino acid sequence or a functional RNA, which is referred to as the *coding region*. A gene also includes all of the DNA sequences required for synthesis of a particular RNA transcript, no matter where those sequences are located in relation to the coding region. For example, in genes of multicellular animals, transcription-control regions known as **enhancers** can lie 50 kb or more from the coding region. As we learned in Chapter 5, other critical noncoding regions in eukaryotic genes include **promoters**, which determine where transcription initiates on a DNA template; *poly(A) sites* that specify 3′ cleavage and polyadenylation, generating the 3′-end of an mRNA; and *splice sites* where exons in pre-mRNA molecules are spliced together (see Figure 5-27). Since these noncoding regions control transcription initiation and RNA processing, mutations in these sequences affect the normal expression and function of RNAs, producing distinct phenotypes in mutant organisms, even though the coding regions of the associated genes are normal. We examine these various transcription-control elements and mechanisms

regulating post-transcriptional RNA processing in greater detail in Chapters 8 and 9.

Although most well-defined genes are transcribed into mRNAs that encode proteins, some DNA sequences are transcribed into important functional RNAs that do not encode proteins [e.g., tRNAs and rRNAs, described in Chapters 5 and 9; miRNAs and siRNAs that regulate mRNA translation and stability, discussed in Chapter 9; and long noncoding RNAs (lncRNAs) that regulate transcription, discussed in Chapter 8]. Because the DNA sequences that encode these functional RNAs can cause specific phenotypes when they are mutated, these DNA regions also are generally referred to as *genes*, even though their final products are RNA molecules and not proteins.

In this section, we will examine the structure of genes in eukaryotes and how that structure influences gene expression and evolution. The nucleotide sequences within the chromosomal DNA of multicellular organisms can be classified on the basis of their functions and the number of times closely related sequences are found in the genomes of individuals of the same species, as shown in Table 7-1 and described in subsequent sections of the chapter.

TABLE 7-1 • Major Classes of Nuclear Eukaryotic DNA and Their Representation in the Human Genome

Class	Length	Copy Number in Human Genome	Fraction of Human

			Genome (%)[i]
Protein-coding genes	0.5–2200 kb	~19,000	~40[ii] (2.0[iii])
Long noncoding RNA genes	0.2–50 kb	~10,000	~15 (0.9[iii])
Tandemly repeated genes			
U2 snRNA	6.1 kb[iv]	~20	<0.0001
rRNAs	43 kb[iv]	~300	0.4
Repetitive DNA			
Simple-sequence DNA	1–500 bp	Variable	~6
Interspersed repeats (mobile DNA elements)			
DNA transposons	2–3 kb	300,000	3
LTR retrotransposons	6–11 kb	440,000	8
Non-LTR retrotransposons			
LINEs	6–8 kb	860,000	21
SINEs	100–400 bp	1,600,000	13
Processed pseudogenes	Variable	~12,500	~0.4
Intergenic regions	Variable	n.a.	~25

[i]The sum of "Fraction of the Human Genome (%)" totals more than 100% because mobile DNA elements are counted twice: once to show the different classes of human mobile DNA elements; and second as part of the intergenic regions and protein-coding genes where they are located in introns and 3′ untranslated regions of terminal exons.

[ii]Complete transcription units including exons and introns.

[iii]Total length of all exons. Protein-coding regions total 1.2 percent of the genome.

[iv]Length of each repeat in a tandemly repeated sequence.

SOURCE: Data from International Human Genome Sequencing Consortium, 2001, *Nature* **409**:860 and 2004, *Nature* **431**:931.

Most Genes of Multicellular Eukaryotes Contain Introns and Produce mRNAs Encoding Single Proteins

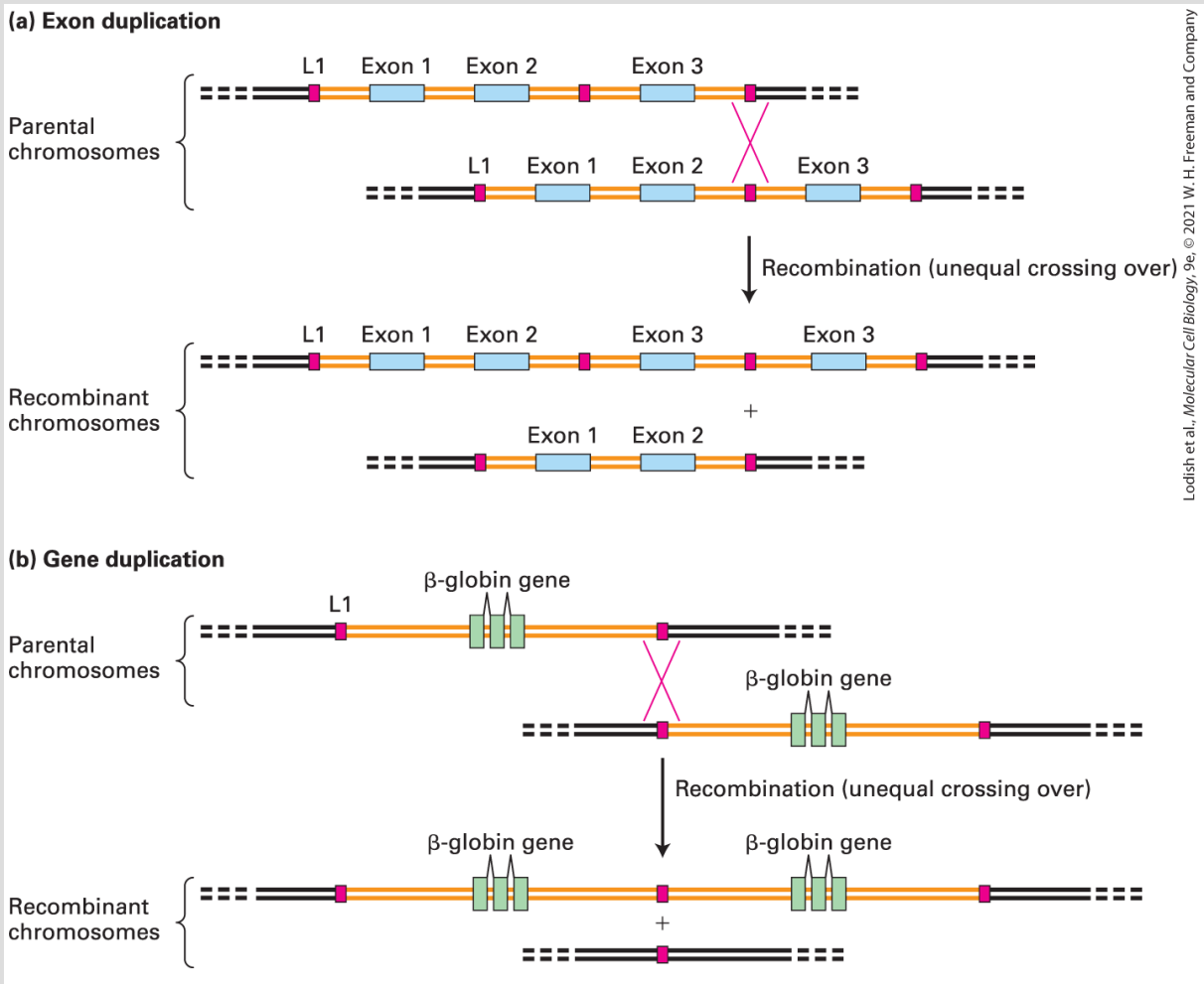
Many bacterial mRNAs (e.g., the mRNA encoded by the *trp* operon) include the coding region for several proteins that function together in a biological process. Such mRNAs are said to be *polycistronic*. (A *cistron* is a genetic unit encoding a single polypeptide.) In contrast, most eukaryotic mRNAs are *monocistronic*; that is, each mRNA molecule encodes a single protein. This difference between polycistronic mRNAs in bacteria and monocistronic mRNAs in eukaryotes correlates with a fundamental difference in their translation mechanisms. Within a bacterial polycistronic mRNA, a

ribosome-binding site is located near the start site for each of the protein-coding regions, or cistrons, in the mRNA. Translation initiation can begin at any of these multiple internal sites, producing alternative proteins from one polycistronic mRNA molecule. In most eukaryotic mRNAs, however, the 5′ cap directs ribosome binding, and translation begins at the closest AUG start codon to the 5′ cap. As a result, translation of most eukaryotic mRNAs begins only at this site, even if, as for many viral mRNAs, the mRNA molecule contains additional open reading frames beginning with an AUG initiation codon downstream from the first open reading frame that is used in that mRNA.

The term *primary transcript* refers to the initial RNA transcript from a gene before it is modified by RNA splicing or addition of a poly(A) tail. In many cases, the primary transcripts of eukaryotic protein-coding genes are processed into a single type of mRNA, which is translated to give a single type of polypeptide (see Figure 5-27). However, unlike bacterial and yeast genes, which generally lack introns, most genes in multicellular animals and plants contain introns, which are removed during RNA processing in the nucleus before the fully processed mRNA is exported to the cytosoplasm for translation. In many cases, the introns in a gene are considerably longer than the exons. The median intron length in human genes is 3.3 kb. However, some are much longer: the longest known human intron is 17,106 bp and lies within the *titin* gene, which encodes a very large structural protein in muscle cells (see Chapter 17). In

comparison, most exons in human genes contain only 50–200 bp. The typical human gene encoding an average-sized protein is about 50,000 bp long, but more than 95 percent of that sequence consists of introns and flanking noncoding 5′ and 3′ regions.

Many large proteins with required functions in multicellular organisms have repeated domains encoded by repeats of similar exons separated by introns of variable length. An example is fibronectin, a component of the extracellular matrix. The fibronectin gene contains multiple copies of five types of exons (see Figure 5-28). Such genes evolved by tandem duplication of the DNA encoding the repeated exon, probably by unequal crossing over during meiosis, as shown in Figure 7-2a.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 7-2 Exon and gene duplication. (a) Exon duplication often results from unequal crossing over during meiosis. Each parental chromosome (*top*) contains one ancestral gene containing three exons (blue) and two introns (orange). Homologous noncoding sequences called L1 long interspersed elements (red) lie 5′ and 3′ of the gene as well as in the intron between exons 2 and 3. As we will see later in the chapter, L1 elements have been repeatedly transposed to new sites in the genome over the course of human evolution, so that all chromosomes are peppered with them. The parental chromosomes are shown displaced relative to each other, so that a pair of the L1 elements are aligned. Homologous recombination between these L1 elements as shown would generate one recombinant chromosome in which the gene now has four exons (two copies of exon 3) and one chromosome in which the gene is missing exon 3. (b) The same process can generate duplications of entire genes. Each parental chromosome (*top*) contains one ancestral β-globin gene. After unequal recombination between L1 elements, subsequent independent mutations in the resulting duplicated genes could lead to slight changes in sequence that

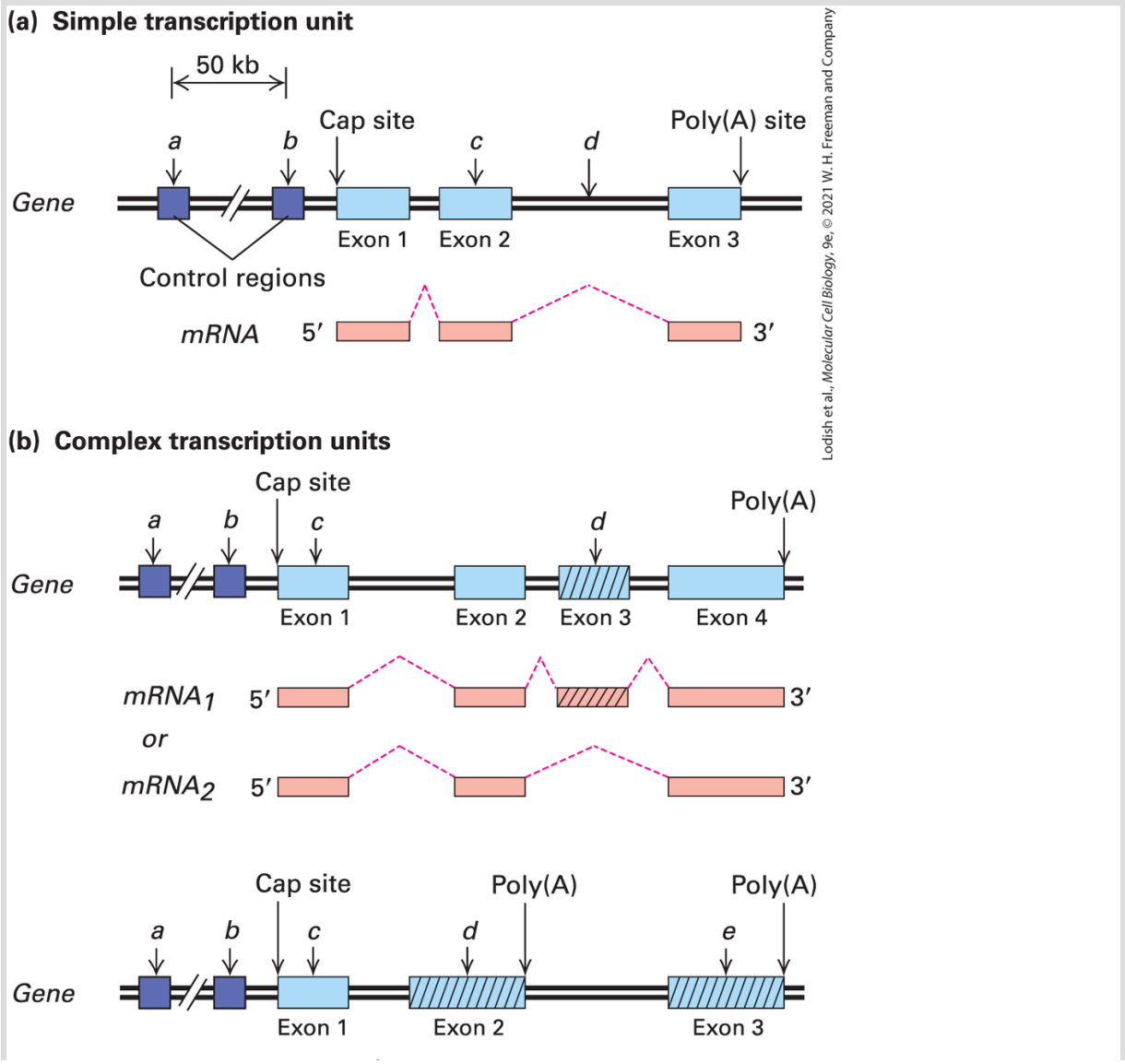
might result in slightly different functional properties of the encoded proteins. Unequal crossing over can also result from rare recombinations between unrelated sequences. See D. H. A. Fitch et al., 1991, *Proc. Nat'l. Acad. Sci. USA* **88**:7396.

Simple and Complex Transcription Units Are Found in Eukaryotic Genomes

The cluster of genes that form a bacterial operon constitutes a single **transcription unit**, which is transcribed from a specific promoter in the DNA sequence to a termination site, producing a single primary transcript. As a consequence, genes and transcription units are often distinguishable in prokaryotes, since a single transcription unit contains several genes when they are part of an operon. In contrast, most eukaryotic genes are expressed from separate transcription units, so that each mRNA is translated into a single protein.

Eukaryotic transcription units are classified into two types — simple and complex — depending on the fate of the primary transcript, that is, the newly transcribed RNA before it is subjected to any modifications. A simple transcription unit contains the protein-encoding exons separated by introns, and upstream control regions (Figure 7-3a). The primary transcript produced from a *simple* transcription unit is processed to yield a single type of mRNA,

encoding a single protein (Figure 7-3a). In humans, simple transcription units, such as the one encoding β-globin, are rare. Approximately 95 percent of human transcription units are *complex*. Complex transcription units produce primary RNA transcripts that can be processed in more than one way, leading to formation of mRNAs containing different combinations of exons (see Figure 7-3b). Each alternative mRNA is translated into a single polypeptide, with translation usually initiating at the first AUG in the mRNA.



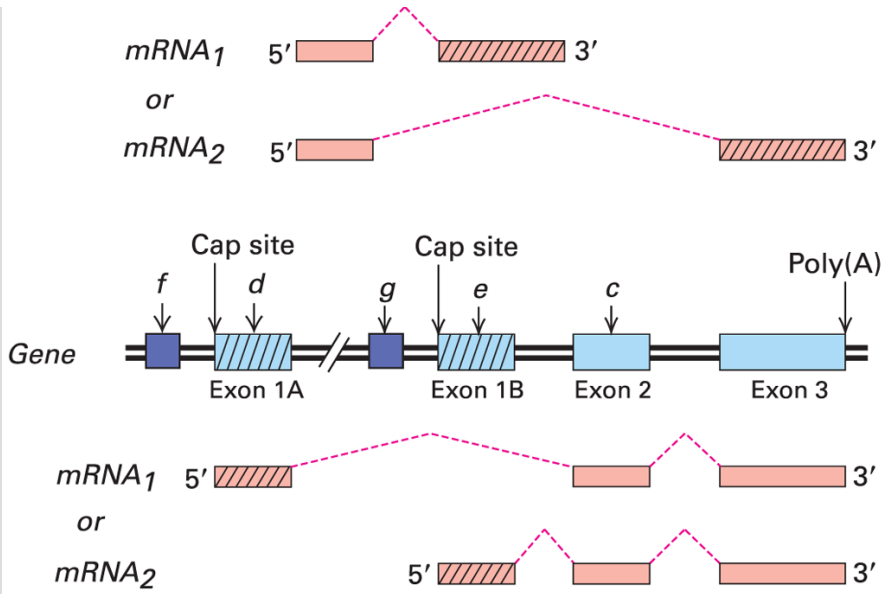


FIGURE 7-3 Simple and complex eukaryotic transcription units. (a) A simple transcription unit includes a region that encodes one protein, which extends from the $5'$ cap site to the $3'$ poly(A) site, plus the associated control regions. Exons (light blue rectangles) are separated by introns. Introns are removed during processing of the primary transcripts (dashed red lines); thus they do not occur in the functional monocistronic mRNA. (b) Complex transcription units produce primary transcripts that can be processed in alternative ways. (*Top*) If a primary transcript contains alternative splice sites, it can be processed into mRNAs with the same $5'$ and $3'$ exons but different internal exons. (*Middle*) If a primary transcript has two poly(A) sites, it can be processed into mRNAs with alternative $3'$ exons. (*Bottom*) If alternative promoters (*f* or *g*) are active in different cell types, mRNA_1, produced in a cell type in which *f* is activated, has a different first exon (1A) than mRNA_2, which is produced in a cell type in which *g* is activated (and in which exon 1B is used).

Multiple mRNAs can arise from one primary transcript in three ways. First, when a primary transcript contains **alternative splice sites**, the transcript can be processed into mRNAs with the same $5'$ and $3'$ exons but different internal exons (Figure 7-3b, *top*). Second, when a primary transcript has two poly(A) sites, the transcript can be processed into mRNAs with alternative $3'$ exons (Figure 7-3b,

middle). Third, alternative promoters may be present but active in different cell types (Figure 7-3b, *bottom*), For example, when promoter *f* is activated in cell type 1, then the first exon in $mRNA_1$ is exon 1A. In cell type 2, where promoter *f* is inactive but promoter *g* is activated, $mRNA_2$ begins with exon 1B.

It is common for one mRNA to be produced from a complex transcription unit in some cell types, while a different mRNA is made in other cell types. For example, the fibronectin gene (see Figure 5-28) encoding long, linear fibrous proteins is transcribed at high level in both fibroblasts, the major cell type in connective tissue, and hepatocytes, the principal liver cell type that secretes most of the extracellular proteins in blood. When expressed in fibroblasts, the fibronectin protein includes domains encoded by related exons EIIIA and EIIIB (green) that adhere to specific proteins anchored to the fibroblast cell surface. In contrast, when fibronectin is expressed in hepatocytes, the EIIIA and EIIIB exons are left out of the mRNA so that the hepatocyte isoform of fibronectin does not bind to cell surfaces and is free to circulate in the blood, where it can be incorporated into blood clots when they form through interactions with other fibronectin domains. This type of alternative mRNA splicing occurs for ~95 percent of human genes, greatly expanding the number of proteins encoded in the genomes of multicellular organisms.

Mutations in exons, introns, and transcription-control regions may influence expression of the protein encoded by a simple transcription unit. For example, in Figure 7-3a, mutation *a* or *b* in the transcription control region might reduce or prevent transcription from occurring, thus little or none of the encoded protein is synthesized. Mutation *c* within an exon could alter the amino acid sequence so that a stop codon is introduced or an abnormal protein is made. Finally, mutation *d* is located within an intron. If it introduces a new splice site, that will affect the sequence of the mRNA molecule, resulting in a nonfunctional protein.

In complex transcription units, the relationship between a mutation and a gene is not always straightforward. A mutation in the control region or in an exon shared by alternatively spliced mRNAs will affect all of the alternative proteins encoded by a given complex transcription unit. On the other hand, a mutation in an exon present in only one of the alternative mRNAs will affect only the protein encoded by that mRNA. Figure 7-3b shows examples of mutations in various components of a complex transcription unit.

As explained in Chapter 6, **genetic complementation** tests are commonly used to determine if two mutations are in the same or different genes (see Figure 6-7). However, in the complex transcription unit shown in Figure 7-3b (*middle*), mutations *d* and *e* would complement each other in a genetic complementation test even though they occur in the same gene, because a chromosome

with mutation *d* can express a normal protein encoded by $mRNA_2$ and a chromosome with mutation *e* can express a normal protein encoded by $mRNA_1$. Both mRNAs produced from this gene would be present in a diploid cell carrying both mutations, generating both protein products and hence a wild-type phenotype. However, a chromosome with mutation *c* in an exon common to both mRNAs would not complement either mutation *d* or mutation *e*. In other words, mutation *c* would be in the same complementation groups as mutations *d* and *e*, even though *d* and *e* themselves would not be in the same complementation group!

Given these complications with the genetic definition of a gene, the molecular definition outlined at the beginning of this section is commonly used. In the case of protein-coding genes, a gene is the DNA sequence transcribed into a pre-mRNA precursor, equivalent to a transcription unit, plus any regulatory elements required for synthesis of the primary transcript.

Protein-Coding Genes May Be Solitary or Belong to a Gene Family

The set of all human proteins is encoded by transcription units that comprise approximately 40 percent of the human genome. However, most of this sequence is in introns that average about 3.3 kb in length but vary greatly in size and can be up to ~17 kb. Compare this

with most human exons that are only 50–200 bp long. In multicellular organisms, roughly 25–50 percent of the protein-coding genes are represented only once in the haploid genome and thus are termed *solitary* genes. A well-studied example of a solitary protein-coding gene is the gene for lysozyme. Lysozyme is an enzyme that cleaves the polysaccharides in bacterial cell walls, causing the bacteria to lyse and die. It is abundant in chicken egg whites and is also found in human tears, where its activity helps to keep the egg and the surface of the eye sterile. The 15-kb DNA sequence encoding chicken lysozyme constitutes a simple transcription unit containing four exons and three introns. The flanking regions, extending about 20 kb upstream and downstream from the transcription unit, do not encode any detectable mRNAs and are thus examples of intergenic regions (i.e., DNA sequences between genes).

Duplicated genes constitute the second group of protein-coding genes. These genes have close but nonidentical sequences and are often located within 5–50 kb of one another. A set of duplicated genes that encodes proteins with similar but nonidentical amino acid sequences is called a **gene family**, and the encoded, closely related, homologous proteins constitute a **protein family**. A few protein families, such as protein kinases, vertebrate immunoglobulins, and olfactory receptors, include hundreds of members. Most protein families, however, include from a few to 30

or so members; common examples are cytoskeletal proteins, the myosin heavy chain, and the α-like and β-like globins in vertebrates.

The genes encoding the β-like globins are a good example of a gene family. As shown in Figure 7-4a, the β-like globin gene family contains five functional genes, designated *HBB* (encoding the most abundant adult β-globin), *HBD* (a minor adult β-globin), *HBG1* and *HBG2* (fetal β-globins), and *HBE1* (embryonic β-globin). Two identical β-like globin polypeptides combine with two identical α-like globin polypeptides (encoded by another gene family expressed during embryonic, fetal, and adult stages of development) and four heme prosthetic groups (see Figure 12-17) to form a hemoglobin molecule. All the hemoglobins formed from the different α-like and β-like globins carry oxygen in the blood, but they exhibit somewhat different properties that are suited to their specific functions in human physiology. For example, hemoglobins containing either the *HBG1*- or *HBG2*-encoded polypeptides are expressed only during fetal life. Because these fetal hemoglobins have a higher affinity for oxygen than adult hemoglobins, they can effectively extract oxygen from the maternal circulation in the placenta. The lower oxygen affinity of adult hemoglobins, which are expressed after birth, permits better release of oxygen to the tissues, especially muscles, which have a high demand for oxygen during exercise. The embryonic hemoglobin assembled from polypeptides encoded by the *HBE1* gene and the embryonic α-like globin gene *HBZ* has an

even higher affinity for oxygen than the fetal and adult hemoglobins.

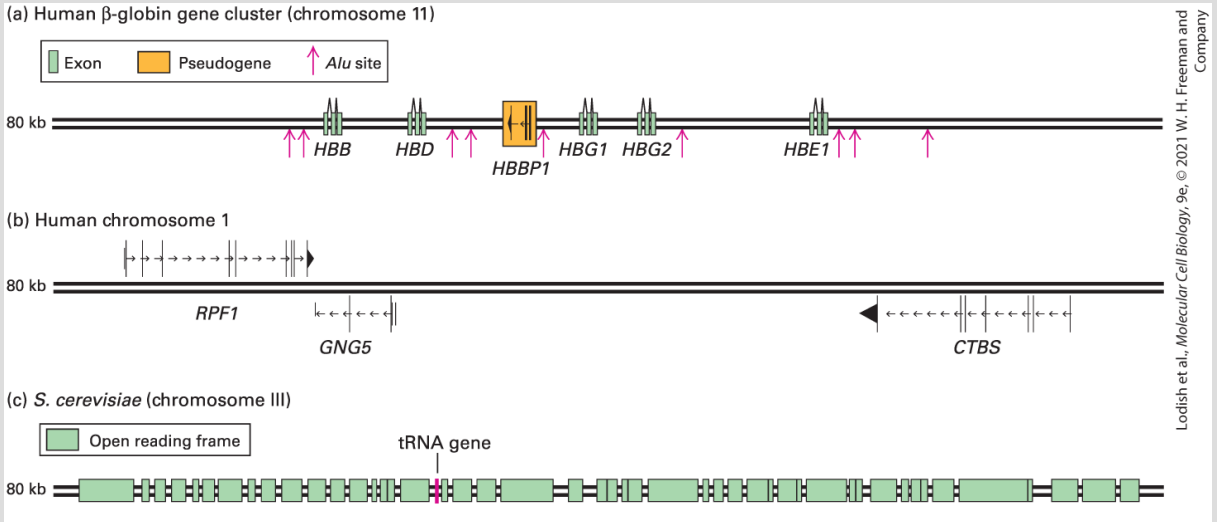


FIGURE 7-4 Comparison of gene density in unicellular and multicellular eukaryotes. (Double black lines represent genomic DNA. Genes diagrammed above these lines are transcribed to the right; genes below are transcribed to the left.) (a) In this diagram of the β-globin gene cluster on human chromosome 11, the green vertical lines represent exons of β-globin–related genes. All of the human β-globin genes are transcribed to the left. The human β-globin gene cluster also contains a pseudogene (orange highlight) that is related to the functional β-globin genes but is not transcribed. Each red arrow indicates the location of an *Alu* sequence, a roughly 300-bp noncoding repeated sequence discussed in Section 7.3 that is abundant in the human genome. See F. S. Collins and S. M. Weissman, 1984, *Prog. Nucl. Acid Res. Mol. Biol.* **31**:315. (b) Diagram of genes in a randomly chosen 80 kb region of human chromosome 1. *RPF1* is transcribed to the right; *GNG5* and *CTBS* are transcribed to the left. Where the $3'$-exon is long enough, its $3'$ end is indicated by an arrowhead. (c) Diagram of yeast genes in an 80 kb region of chromosome III. The green boxes indicate open reading frames. Most of these are functional genes without introns. Note the much higher proportion of noncoding to coding sequences in the human DNA than in the yeast DNA. See S. G. Oliver et al., 1992, *Nature* **357**:28.

The different β-like globin genes arose by duplication of an ancestral gene, most likely as the result of unequal crossing over during meiotic recombination in a developing germ cell (egg or sperm) (see Figure 7-2b). Over evolutionary time, the two copies of the gene that resulted accumulated random mutations, resulting in *sequence drift*. Beneficial mutations that conferred some refinement in the basic oxygen-carrying function of hemoglobin were retained by natural selection. Repeated gene duplications and subsequent sequence drift and selection are thought to have generated the contemporary β-like globin genes observed in humans and other mammals today.

One region in the human β-like globin gene cluster contains a nonfunctional sequence, called a **pseudogene**, that is similar to the functional β-like globin genes (Figure 7-4a). Sequence analysis shows that this pseudogene has the same apparent exon–intron structure as the functional β-like globin genes, suggesting it arose by duplication of the same ancestral gene. However, there was little selective pressure to maintain the function of this gene. Consequently, sequence drift during evolution generated sequences that terminate translation and block mRNA processing, rendering this region nonfunctional. Because such pseudogenes are not deleterious, they remain in the genome and mark the location of a gene duplication that occurred in one of our ancestors, followed by sequence drift eliminating the gene function.

Duplications of segments of a chromosome (called *segmental duplication*) occurred fairly often during the evolution of multicellular plants and animals. As a result, a large fraction of the genes in these organisms today have been duplicated, allowing the process of sequence drift to generate gene families and pseudogenes. The extent of sequence divergence between duplicated copies of the genome and characterization of the homologous genomic sequences in related organisms allows us to estimate the time in evolutionary history when the duplication occurred. For example, the human fetal globin genes (*HBG1* and *HBG2*) evolved following the duplication of a 5.5-kb region in the β-globin locus that included the single *HBG*-globin gene in the common ancestor of catarrhine primates (Old World monkeys, apes, and humans) and platyrrhine primates (New World monkeys) about 50 million years ago.

Although members of gene families that arose relatively recently in evolution, such as the genes of the human β-globin family, are often found near one another on the same chromosome, members of gene families may also be found on different chromosomes in the same organism. This is the case for the human α-like globin genes, which were separated from the β-globin genes by an ancient chromosomal translocation. Both the α- and β-globin genes evolved from a single ancestral globin gene that was duplicated (see Figure 7-2b) to generate the predecessors of the contemporary α- and β-globin genes in mammals. Both the primordial α- and β-globin genes

then underwent further duplications to generate the different genes of the α- and β-globin gene clusters found in mammals today.

Several different gene families encode the various proteins that make up the cytoskeleton. These proteins are present in varying amounts in almost all cells. In vertebrates, the major cytoskeletal proteins are the actins, tubulins, and intermediate filament proteins such as the keratins, discussed in Chapters 17, 18, and 20. We examined the origin of one such family, the tubulin family, in Section 6.3. The keratins, important for determining the shape and physical properties of epithelial cells that cover tissue surfaces, have expanded into at least 54 functional genes in humans. Although the physiological rationale for the evolution of so many keratins is not as obvious as it is for the globins, the different members of the family probably have similar but subtly different functions suited to the particular type of cell in which they are expressed.

Heavily Used Gene Products Are Encoded by Multiple Copies of Genes

In vertebrates and invertebrates, the genes encoding ribosomal RNAs and some other nonprotein-coding RNAs, such as those involved in RNA splicing, occur as *tandemly repeated arrays*. Most often, copies of these sequences appear one after the other, in a head-to-tail fashion, over a long stretch of DNA. Within a tandem

array of rRNA genes, each copy is nearly exactly like all the others. Although the transcribed portions of the genes are the same, the nontranscribed regions between the transcribed regions can vary in length and sequence.

These tandemly repeated rRNA genes have evolved to meet the great cellular demand for their transcripts during cell replication. To understand why, consider that a fixed maximal number of rRNA molecules can be produced from a single gene during one cell generation when the gene is fully loaded with RNA polymerase molecules. If more RNA is required than can be transcribed from one gene, multiple copies of the gene are necessary. For example, during early embryonic development in humans, many embryonic cells have a doubling time of about 24 hours and contain 5–10 million ribosomes. To produce enough rRNA to form this many ribosomes, an embryonic human cell needs at least 100 copies of genes encoding the large and small rRNAs, and most of these genes must be close to maximally active for the cell to divide every 24 hours; that is, multiple RNA polymerases must be transcribing each rRNA gene at the same time. Indeed, all eukaryotes, including yeasts, contain 100 or more copies of the 5S rRNA gene as well as similar numbers of the genes encoding the other rRNAs.

The genes encoding tRNA and those encoding the histone proteins are also present in multiple copies in eukaryotic cells. As we will see later in this chapter, histones bind and organize nuclear DNA. Just as the cell requires multiple rRNA and tRNA genes to produce

sufficient numbers of ribosomes and tRNAs, multiple copies of the histone genes are required to produce sufficient histone protein to bind the large amount of nuclear DNA produced in each round of cell replication. While tRNA and histone genes often occur in clusters, they generally do not occur in tandem arrays in the human genome.

Nonprotein-Coding Genes Encode Functional RNAs

In addition to rRNA and tRNA genes, there are thousands of additional genes that are transcribed into nonprotein-coding RNAs, some with various known functions and many whose functions are not yet known. For example, **small nuclear RNAs (snRNAs)** function in RNA splicing, and **small nucleolar RNAs (snoRNAs)** function in rRNA processing and base modification in the nucleolus. The RNase P RNA is a ribozyme that functions in tRNA processing, and a large family (~19,000 in humans) of short **micro-RNAs (miRNAs)** regulates the translation and stability of specific mRNAs. The functions of these nonprotein-coding RNAs are discussed in Chapter 9. An RNA found in telomerase (see Figure 7-40) maintains the sequence at the ends of chromosomes, and the 7SL RNA is a component of the signal recognition particle that functions in the transport of secreted proteins and most membrane proteins into the endoplasmic reticulum (see Chapter 13). These and other nonprotein-coding

RNAs encoded in the human genome, and their functions when known, are listed in Table 7-2. Recent advances in DNA sequencing have led to the discovery of about 10,000 *long noncoding RNAs* (lncRNAs) in nuclei of mammalian cells. Some of these have been found to function in regulating the expression of specific protein-coding genes. Pursuing the functions of lncRNAs is currently a highly active area of research.

TABLE 7-2 • Known Nonprotein-Coding RNAs and Their Functions

RNA	Number of Genes in Human Genome	Function
rRNAs	~300	Protein synthesis
tRNAs	~500	Protein synthesis
snRNAs	~40	mRNA splicing
U7 snRNA	1	Histone mRNA $3'$ processing
snoRNAs	~85	Pre-rRNA processing and rRNA modification
miRNAs	~19,000	Regulation of gene expression
piRNAs	$>6 \times 10^7$	Suppress transposon transpositions in germ cells. Associated with PIWI proteins.
Xist	1	X-chromosome inactivation
7SK	1	Transcription control
RNase P RNA	1	tRNA $5'$ processing

7SL RNA	3	Protein secretion (component of signal recognition particle, SRP)
RNase MRP RNA	1	rRNA processing, mtDNA[i] replication
Telomerase RNA	1	Template for addition of telomeres
Vault RNAs	3	Components of Vault ribonucleoproteins (RNPs). Regulate autophagy.
hY1, hY3, hY4, hY5	~30	Components of ribonucleoproteins (RNPs), function unknown
H19	1	Unknown

[i]Mitochondrial DNA

SOURCE: Data from International Human Genome Sequencing Consortium, 2001, *Nature* **409**:860, and P. D. Zamore and B. Haley, 2005, *Science* **309**:1519.

KEY CONCEPTS OF SECTION 7.1

Eukaryotic Gene Structure and Organization

- In molecular terms, a gene is the entire DNA sequence required for synthesis of a functional protein or RNA molecule. In addition to the coding regions (exons), a gene includes control regions, and in multicellular animals and plants most genes include introns.
- A simple eukaryotic transcription unit produces a single monocistronic mRNA, which is translated into a single polypeptide.
- A complex eukaryotic transcription unit is transcribed into a primary transcript that can be processed into two or more different monocistronic mRNAs depending on the choice of splice sites or polyadenylation sites. A complex transcription unit with

alternative promoters generates two or more different mRNAs, often in different cell types (see Figure 7-3b).

- Many complex transcription units (e.g., the fibronectin gene) express one mRNA in one cell type and an alternative mRNA in a different cell type.
- About half the protein-coding genes in vertebrate genomic DNA are solitary genes, each occurring only once in the haploid genome. The remaining are duplicated genes, which arose by duplication of an ancestral gene and subsequent independent mutations (see Figure 7-2b). The proteins encoded by a gene family have homologous but nonidentical amino acid sequences and exhibit similar but slightly different properties that are optimal for the cell type in which they are expressed.
- In invertebrates and vertebrates, rRNAs are encoded by multiple copies of genes located in tandem arrays in genomic DNA. Multiple copies of tRNA, snRNA, and histone genes also occur, often in clusters, but not generally in tandem arrays in humans.
- Many genes also encode functional RNAs that are not translated into proteins but nonetheless perform significant functions, such as rRNA, tRNA, and snRNA. Among these are micro-RNAs, whose biological significance in regulating gene expression has most recently been appreciated. Understanding the function of many short non-coding RNAs and thousands of newly discovered nuclear long noncoding RNAs (lncRNAs) is currently an intense area of research.

7.2 Chromosomal Organization of Genes and Noncoding DNA

Having reviewed the relationship between transcription units and genes, we now consider the organization of genes on chromosomes and of noncoding DNA that is not expressed as coding sequences in exons of mRNAs or as stable functional RNAs such as tRNAs, miRNAs, and lncRNAs.

Genomes of Many Organisms Contain a Large Fraction of Noncoding DNA

Comparisons of the total chromosomal DNA per cell in various species first suggested that much of the DNA in certain organisms does not encode proteins or functional RNA or have any apparent regulatory function. For example, yeasts, fruit flies, chickens, and humans have successively more DNA in their haploid chromosome sets (11.9, 137, 1043, and 2968 Mb, respectively), in keeping with what we perceive to be the increasing complexity of these organisms. Yet the vertebrates with the greatest amount of DNA per cell are amphibians, which are surely less complex than humans in their structure and behavior. Even more surprising, the unicellular protozoan *Amoeba dubia* has 200 times more DNA per cell than humans. Many plant species also have considerably more DNA per cell than humans have. Tulips, for example, have 10 times as much

DNA per cell as humans. The DNA content per cell also varies considerably between closely related species. All insects or all amphibians would appear to be similarly complex, but the amount of haploid DNA in species within each of these phylogenetic classes varies by a factor of 100.

Sequencing and identification of exons in chromosomal DNA have provided direct evidence that the genomes of multicellular eukaryotes contain large amounts of noncoding DNA. For instance, only a small portion of the β-globin gene cluster of humans, which is about 80 kb long, encodes protein (see Figure 7-4a). Similarly, a randomly chosen 80-kb region from chromosome 1 includes only the nine short exons of the *RFP1* gene, the four exons of the *GNG5* gene, and the seven exons of the *CTBS* gene. The total length of sequence encoding mRNA in this 80-kb region is only 9325 base pairs, or ~12 percent. In contrast, a typical 80-kb stretch of DNA from the yeast *S. cerevisiae* contains many closely spaced protein-coding sequences, few introns, and little noncoding DNA (see Figure 7-4c).

The density of genes varies among regions of human chromosomal DNA, from *gene-rich* regions, where a few hundred base pairs separate transcription units, to large *gene deserts*, where intergenic regions are a few million base pairs long. Of the 96 percent of human genomic DNA that has been sequenced, only about 2.9 percent is exons, and only about 1.5 percent encodes proteins. (The

fraction of the genome that corresponds to exons is much larger than the fraction that encodes proteins because many protein-coding genes include exons for long 3′ untranslated regions and because there are many exons in nonprotein-coding lncRNAs; see Chapter 9.) We learned in the previous section that the intron sequences of most human genes are significantly longer than the exon sequences. Approximately 55 percent of human genomic DNA is thought to be transcribed into pre-mRNAs, pre-lncRNAs, or other nonprotein-coding RNAs in one cell type or another, but some 95 percent of this sequence is intronic and is thus removed by RNA splicing. The remaining 45 percent of human DNA constitutes noncoding DNA between genes (*intergenic regions*) as well as the regions of repeated DNA sequences that make up the centromeres and telomeres of the human chromosomes. Consequently, about 97 percent of human DNA does not encode proteins, functional noncoding RNAs, or potentially functional lncRNAs.

Different selective pressures may account, at least in part, for the remarkable difference in the amount of nonfunctional DNA in different organisms. For example, many microorganisms must compete with other species of microorganisms in the same environment for limited amounts of available nutrients, and metabolic economy is thus a critical characteristic for these organisms. Because synthesis of nonfunctional (i.e., noncoding) DNA requires time, nutrients, and energy, presumably there was selective pressure to lose nonfunctional DNA during the evolution of

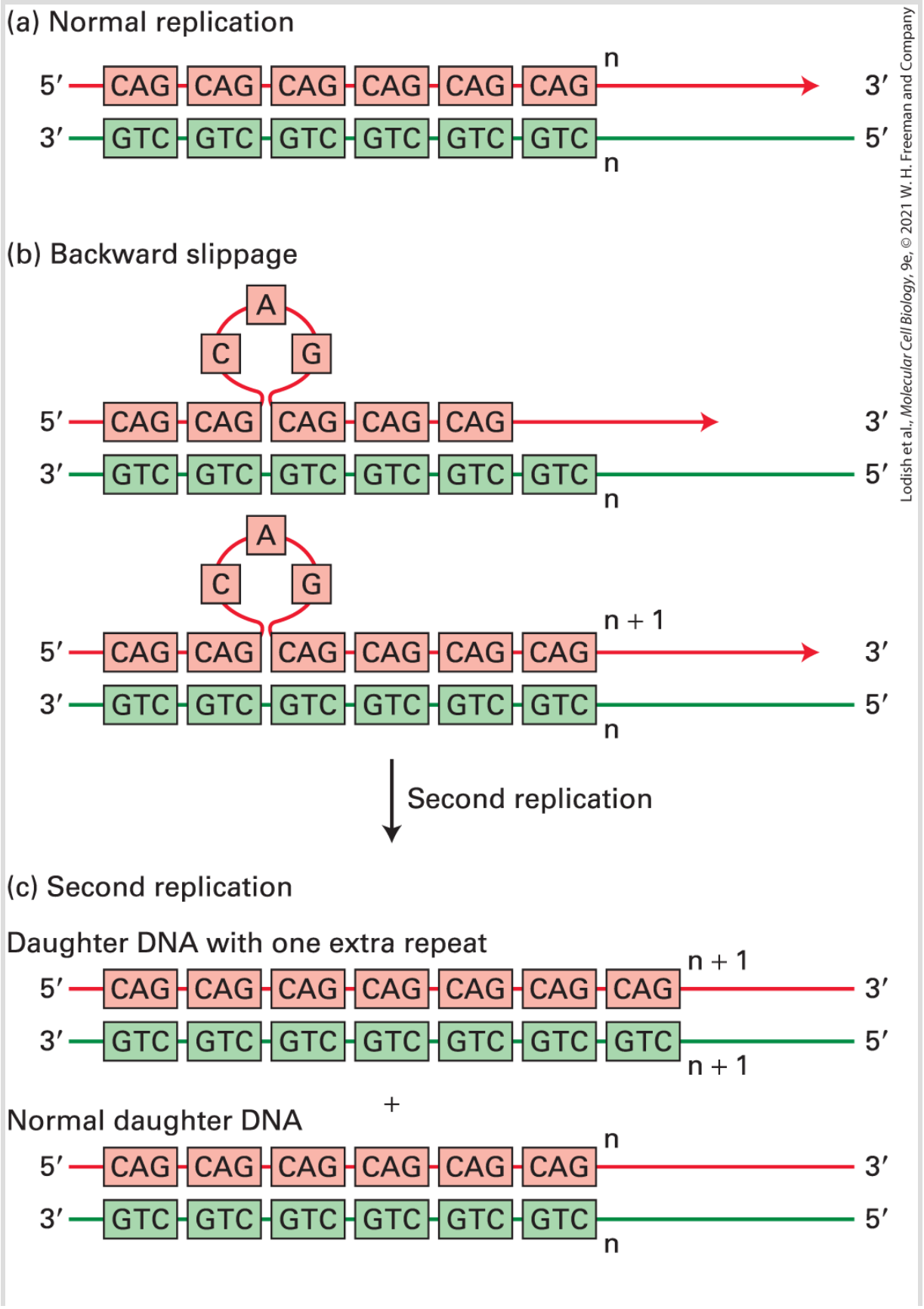
rapidly growing microorganisms such as the yeast *S. cerevisiae*. On the other hand, natural selection in vertebrates depends largely on their behavior. The energy invested in DNA synthesis is trivial compared with the metabolic energy required for the movement of muscles and the function of the nervous system; thus there may have been little selective pressure on vertebrates to eliminate nonfunctional DNA. Furthermore, the replication time of cells in most vertebrates and plants is much longer than in rapidly growing microorganisms, so there may have been little selective pressure to eliminate nonfunctional DNA in order to permit rapid cellular replication.

Most Simple-Sequence DNAs Are Concentrated in Specific Chromosomal Locations

The most abundant class of nuclear DNA found in human cells is *repetitive DNA* (see Table 7-1). There are two types of repetitive DNA: simple-sequence DNA and interspersed repeats (also called *mobile DNA elements*). Interspersed repeats are discussed in Section 7.3. Here we discuss simple-sequence DNA.

Simple-sequence DNA or **satellite DNA**, which constitutes about 6 percent of the human genome, is composed of perfect or nearly perfect repeats of relatively short sequences. The length of each

repeat in simple-sequence DNA can range from 1 to 500 base pairs. DNA sequences in which the repeats each contain 1–13 bp are often called **microsatellites**. Most microsatellite DNA has a repeat length of 1–4 bp, and the repeats usually occur in tandem sequences of 150 repeats or fewer. Microsatellites are thought to have originated by *backward slippage* of a daughter strand on its template strand during DNA replication so that the same short sequence was copied twice (Figure 7-5). Also, for triplet repeats that are self-complementary and include a G (e.g., CAG), excision repair of an oxidized G, the most frequent type of spontaneous DNA damage in humans, can lead to displacement of a strand with the trinucleotide repeat. This displaced region can fold back into a hairpin that is incorporated into the repair patch. Further repair of the single-stranded loop by additional DNA repair enzymes results in an expansion of the number of repeats.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 7-5 Generation of microsatellite repeats by backward slippage of the nascent daughter strand during DNA replication. (a) Normal replication. Daughter strand is red and template strand is green. (b) If the nascent daughter strand "slips" backward relative to the template strand by one repeat, one new copy of the repeat is added to the daughter strand when DNA replication continues. An extra copy of the repeat forms a single-stranded loop in the daughter strand of the daughter duplex DNA molecule. (c) If this single-stranded loop is not removed by DNA repair proteins before the next round of DNA replication, the extra copy of the repeat is added to one of the double-stranded daughter DNA molecules.

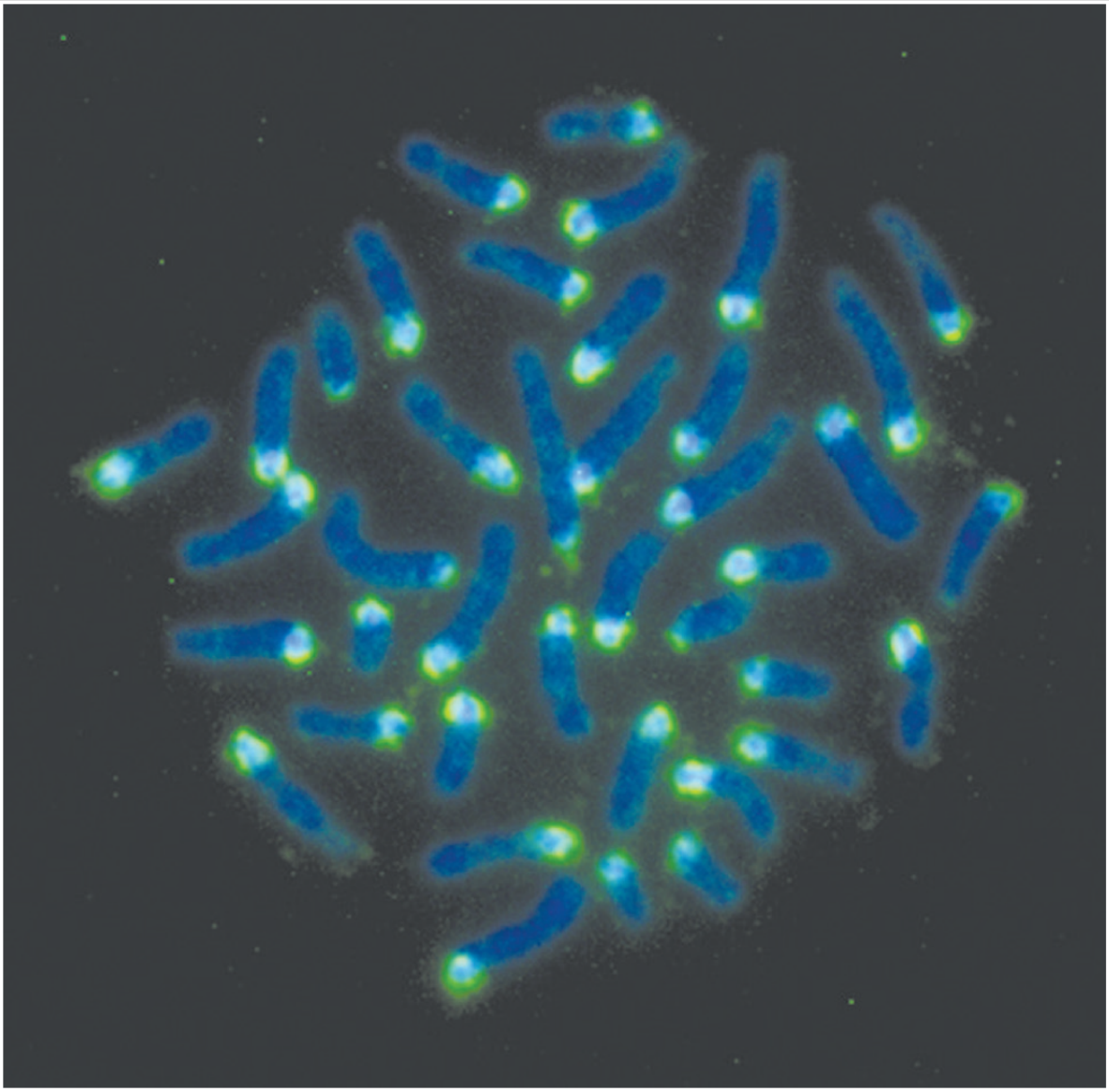
Microsatellites occasionally occur within transcription units. Some individuals are born with a larger number of repeats in specific genes than is observed in the general population, presumably because of daughter-strand slippage during DNA replication or repair of DNA damage in a region of triplet repeats in the germ cells from which they and their forebears developed. Such expanded microsatellites have been found to cause at least 14 different types of neuromuscular diseases. In some cases, expanded microsatellites behave like a recessive mutation because they interfere only with the function or expression of the gene in which they occur. But more commonly, expanded microsatellites behave like dominant mutations. In *Huntington's disease*, triplet repeats occur within a coding region, resulting in the formation of long polymers of a single amino acid that may aggregate over time in long-lived neuronal cells, eventually interfering with normal cellular function. For example, expansion of a CAG repeat in the first exon of the gene involved in Huntington's disease leads to synthesis of long stretches of polyglutamine, which over several decades form

toxic aggregates resulting in neuronal cell death in patients with the disease.

Pathogenic expanded repeats can also occur in the noncoding regions of some genes, where they are thought to function as dominant mutations because they interfere with the processing of a subset of mRNAs in the muscle cells and neurons where the affected genes are expressed. For example, in patients with *myotonic dystrophy type 1*, transcripts of the *DMPK* gene contain between 50 and 1500 repeats of the sequence CUG in the 3′ untranslated region, compared with 5–34 repeats in unaffected individuals. The extended stretch of CUG repeats in affected individuals is thought to form a long RNA hairpin. This RNA molecule binds and sequesters nuclear RNA-binding proteins that normally regulate alternative RNA splicing of certain pre-mRNAs essential for muscle and nerve cell function. ■

As opposed to microsatellite DNA composed of repeats of 1 to 13 base pairs, *simple-sequence* DNA is composed of repeats of 14–500 bp in tandem arrays that are 20–100 kb long. In situ hybridization studies with metaphase chromosomes have localized much of the simple-sequence DNA to the centromeres, the discrete chromosomal regions that attach to spindle microtubules during mitosis and meiosis (Figure 7-6). Experiments in the fission yeast *S. pombe* indicate that these sequences are required to form a specialized chromatin structure called *centromeric heterochromatin*,

which is necessary for the proper segregation of chromosomes to daughter cells during mitosis. Simple-sequence DNA is also found in long tandem repeats at the ends of chromosomes, known as the telomeres. Here the repeat sequences prevent chromosome shortening during DNA replication, a topic we discuss further in Section 7.5.



Courtesy of Sabine Mai, Ph.D., Manitoba Institute of Cell Biology, Canada.

EXPERIMENTAL FIGURE 7-6 Simple-sequence DNA is localized at the centromere in mouse chromosomes. Purified simple-sequence DNA from mouse cells was copied in vitro using *E. coli* DNA polymerase I and fluorescently labeled dNTPs to generate a fluorescently labeled DNA probe for mouse simple-sequence DNA. Chromosomes from cultured mouse cells were fixed and denatured on a microscope slide, and the chromosomal DNA was then hybridized in situ to the labeled probe (light blue). The slide was also stained with DAPI, a DNA-binding dye, to visualize the full length of the chromosomes (dark blue). Fluorescence microscopy shows that the simple-sequence probe hybridizes primarily to one end of the telocentric mouse chromosomes (i.e., chromosomes in which the centromeres are located at one end).

DNA Fingerprinting Depends on Differences in Length of Simple-Sequence DNAs

The nucleotide sequences of the repeat units in a simple-sequence DNA tandem array are highly conserved among individuals within a species. In contrast, the *number* of repeats, and thus the length of simple-sequence tandem arrays containing the same repeat unit, is quite variable among individuals. These differences in length are thought to result from unequal crossing over within regions of simple-sequence DNA during meiosis. As a consequence of this unequal crossing over, the lengths of some tandem arrays are unique in each individual.

In humans and other mammals, some simple-sequence DNA exists in relatively short 1–5-kb regions made up of 20–50 repeat units,

each containing 14–100 bp. These regions are called *minisatellites*, in contrast to microsatellites made up of 1–13-bp tandem repeats. Even slight differences in the total length of various minisatellites from different individuals can be detected by Southern blotting. This technique was exploited in the first application of *DNA fingerprinting*, which was developed to detect DNA *polymorphisms* (i.e., differences in sequence between individuals of the same species) (Figure 7-7).

FIGURE 7-7 Distinguishing individuals by DNA fingerprinting. (a) In this analysis of paternity, several minisatellite repeat lengths were determined by Southern blot analysis of restriction enzyme–digested genomic DNA and hybridization with a probe for a minisatellite sequence repeated at several different positions in the genome. This method generated hypervariable multiband patterns for each individual called "DNA fingerprints." Lane M shows the pattern of restriction fragment bands using the mother's DNA; C, using the child's

DNA; and F1 and F2 using DNA from two potential fathers. The child has minisatellite repeat lengths inherited from either the mother or F1, demonstrating that F1 is the father. Arrows indicate restriction fragments from F1, but not F2, found in the child's DNA. (b) In these DNA fingerprints of a specimen isolated from a victim and three people suspected of the crime, it is clear that minisatellite repeat lengths in the specimen match those of suspect 1. The victim's DNA was included in the analysis to ensure that the specimen DNA was not contaminated with DNA from the victim.

Today the far more sensitive polymerase chain reaction (PCR) technique (see Figure 6-18) is generally used in forensic genetic testing. Microsatellites consisting of 4-bp tandem repeats in 30–50 copies are usually analyzed today. The exact number of repeats at a specific location in the genome generally varies between the two homologous chromosomes of an individual (one inherited from the mother and one from the father) and between the Y chromosomes of different males. A mixture of pairs of PCR primers that hybridize to unique sequences flanking 13 of these short tandem repeats and a Y-chromosome short tandem repeat are used to amplify DNA in a sample from one individual. The resulting mixture of PCR product lengths is unique in the human population, except for identical twins. The use of PCR allows analysis of minute amounts of DNA, and individuals can be distinguished more precisely and reliably than by conventional fingerprinting.

Unclassified Intergenic DNA Occupies a Significant Portion of the Genome

About 45 percent of human DNA lies between transcription units. Much of this sequence is not repeated anywhere else in the genome. Enhancers on the order of 50–200 bp in length that help to regulate transcription at distant promoters occur in these long stretches of intergenic DNA, as well as in introns. These enhancers (discussed in Chapter 8) are often conserved during evolution, while the neighboring intergenic sequences are not conserved. Other conserved intergenic regions may perform significant functions that are not yet understood. For example, they may contribute to the structures of chromosomes discussed in Section 7.5.

KEY CONCEPTS OF SECTION 7.2

Chromosomal Organization of Genes and Noncoding DNA

- In the genomes of prokaryotes and most lower eukaryotes, which contain few nonfunctional sequences, coding regions are densely arrayed along the genomic DNA.
- Genomes of multicellular animals and plants contain many sequences that do not apparently code for functional RNAs or have any regulatory function. Much of this nonfunctional DNA is composed of repeated sequences.
- In humans, only about 2.9 percent of total DNA (the exons) actually encodes proteins or functional RNAs.
- Variation in the amount of nonfunctional DNA in the genomes of different species is largely responsible for the lack of a consistent relationship between the amount of DNA in the haploid chromosomes of an animal or plant and its phylogenetic complexity.
- Simple-sequence DNA, consisting of short sequences repeated in long tandem arrays, is preferentially located in centromeres and telomeres.
- The length of a particular simple-sequence tandem array is quite variable between individuals in a species, probably because of unequal crossing over during meiosis. Differences in the lengths of some simple-sequence tandem arrays form the basis for DNA fingerprinting (see Figure 7-7).

7.3 Transposable (Mobile) DNA Elements

The second type of repetitive DNA in eukaryotic genomes, **interspersed repeats**, is composed of a very large number of copies of relatively few sequence families (see Table 7-1). Also known as *moderately repeated DNA*, or *intermediate-repeat DNA*, these sequences are found throughout mammalian genomes and make up 25–50 percent of mammalian DNA (~45 percent of human DNA).

Because interspersed repeats have the unique ability to move within the genome, they are collectively referred to as **transposable DNA elements** or **mobile DNA elements** (we use these terms interchangeably). Barbara McClintock discovered the first mobile DNA elements while doing classical genetic experiments in maize (corn) during the 1940s. She characterized genetic entities that could move into and back out of genes, changing the phenotypes of corn kernels. Her theories were controversial until similar mobile elements were discovered in bacteria, where they were characterized as specific DNA sequences, and the molecular basis of their transposition was deciphered. Although transposable DNA elements were originally discovered in eukaryotes, they are also found in prokaryotes, although less frequently. Transposable DNA elements are essentially molecular symbionts that in most cases appear to have no specific function in the biology of their host

organisms, but exist only to maintain themselves. For this reason, Francis Crick referred to them as "selfish DNA."

The process by which these sequences are copied and inserted into a new site in the genome is called **transposition**. When transposition occurs in germ cells, the transposed sequences at their new sites are passed on to succeeding generations. In this way, mobile elements have accumulated in eukaryotic genomes over evolutionary time. Since mobile elements are eliminated very slowly from eukaryotic genomes, they now constitute a significant portion of the genomes of many eukaryotes.

Not only are mobile elements the source for much of the DNA in our genomes, but they also provide a mechanism for greatly increasing genome rearrangements by dispersing homologous sequences throughout the genome where meiotic recombination may occur (see Figure 7-2). Any one transposon transposes very rarely. However, since there are about 3.2 million transposons in the human genome (see Table 7-1), even with this very low frequency of transposition for each transposon (once in multiple generations), in humans there is about one new germ-line transposition for every eight individuals. Over time, transposon transpositions in germ cells have played an essential part in the evolution of genes that have multiple exons and of transcription-control regions that regulate transcription of genes in specific cell types and developmental periods (discussed in Chapter 8). In other words, although transposable elements probably evolved as molecular symbionts,

they have had an important function in the evolution of complex multicellular organisms. We will return to this topic at the end of this section.

Transposition also occurs within somatic cells. In this case, the transposed sequence is transmitted only to the daughter cells derived from that cell. In rare cases, somatic-cell transposition may lead to a mutation with detrimental phenotypic effects, such as the inactivation of a tumor-suppressor gene (see Chapter 25). In this section, we first describe the structure and transposition mechanisms of the major types of transposable DNA elements and then consider their likely role in evolution.

Movement of Mobile Elements Involves a DNA or an RNA Intermediate

The transposition of mobile elements occurs in two different ways: (1) transposition directly as DNA or (2) via an RNA intermediate transcribed from the mobile element by an RNA polymerase and then converted back into double-stranded DNA by a **reverse transcriptase** (Figure 7-8). Mobile elements that transpose directly as DNA are generally referred to as **DNA transposons**, or simply **transposons**. Eukaryotic DNA transposons excise themselves from one place in the genome, leaving that site and moving to another. Mobile elements that transpose to new sites in the genome via an

RNA intermediate are called **retrotransposons**. Retrotransposons make an RNA copy of themselves, convert that to double-stranded DNA, and introduce this new copy into another site in the genome, while also remaining at their original location. The movement of retrotransposons is analogous to the infectious process of retroviruses (see Figure 5-43). Indeed, retroviruses can be thought of as retrotransposons with genes encoding viral coats that allow them to transpose between cells. Retrotransposons can be further classified on the basis of their specific mechanism of transposition. To summarize, DNA transposons can be thought of as transposing by a cut-and-paste mechanism, while retrotransposons move by a copy-and-paste mechanism in which the copy is an RNA intermediate.

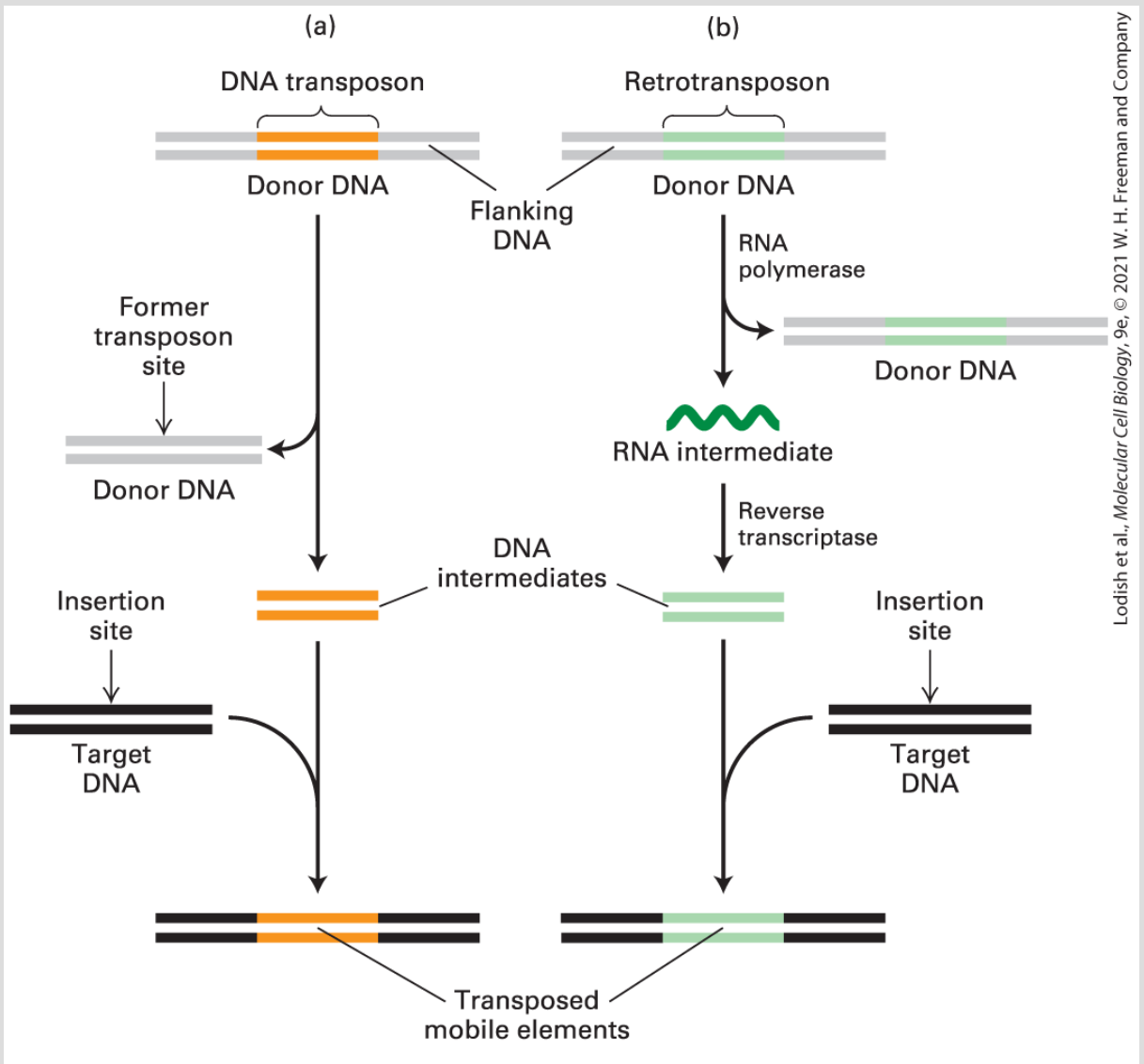


FIGURE 7-8 Two major classes of mobile elements. (a) Eukaryotic DNA transposons (orange) move via a DNA intermediate, which is excised from the donor site. (b) Retrotransposons (green) are first transcribed into an RNA molecule, which is then reverse-transcribed into double-stranded DNA. In both cases, the double-stranded DNA intermediate is integrated into the target-site DNA to complete movement. Thus DNA transposons move by a cut-and-paste mechanism, whereas retrotransposons move by a copy-and-paste mechanism.

Most Mobile Elements in Bacteria Are DNA Transposons Known as Insertion Sequences

Our first molecular understanding of mobile elements came from the study of *Escherichia coli* mutations caused by the insertion of a 1–2-kb DNA sequence into the middle of a gene. The inserted DNA is called an *insertion sequence*, or *IS element*. More than 1000 different IS elements have been found in *E. coli* and other bacteria.

Transposition of a bacterial IS element is a rare event, occurring in only one in 10^5–10^7 cells per generation. Since only a small fraction of bacterial DNA is noncoding or comprises promoters and operators involved in transcription control, when a transposition occurs in a bacterial cell, it may inactivate an essential gene, killing the host cell and the IS elements it carries. Consequently, the low rate of transposition of IS elements in bacteria may have evolved because higher rates of transposition would probably result in too great a mutation rate for the host species to survive. However, since IS elements transpose more or less randomly, some transposed sequences insert into nonessential regions of the host genome (e.g., regions between genes), allowing the host cell to survive. In this way, the number of transposons in a strain of a bacteria reaches an equilibrium at a relatively low number (usually less than 20). IS elements can also insert themselves into plasmids or lysogenic

viruses and can thus be transferred to other cells. In this way, IS elements can transpose into the chromosomes of new host cells.

The general structure of an IS element is diagrammed in Figure 7-9. An inverted repeat of 10–40 bp is invariably present at each end of an IS element. In an *inverted repeat*, the $5' \rightarrow 3'$ sequence on one strand is repeated on the other strand, such as

$$\begin{array}{l} \longrightarrow \\ 5'\ \mathbf{GAGC}\text{———}\mathbf{GCTC}\ 3' \\ 3'\ \mathbf{CTCG}\text{———}\mathbf{CGAG}\ 5' \\ \qquad\qquad\qquad \longleftarrow \end{array}$$

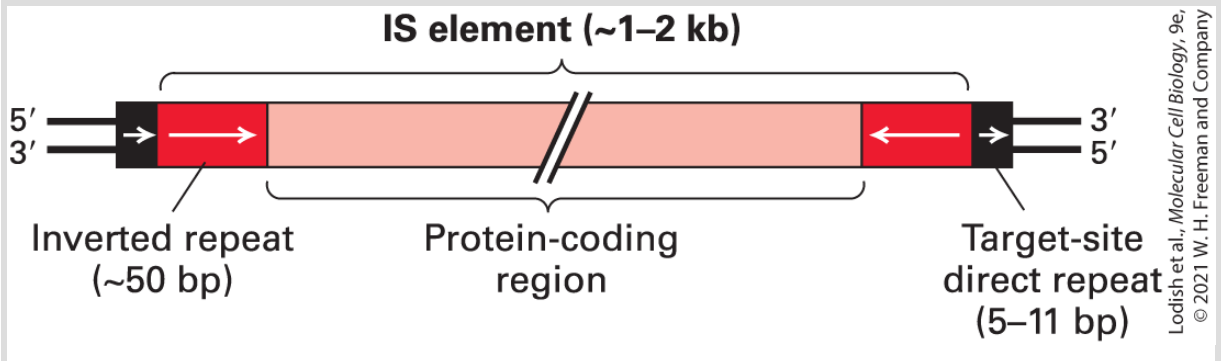


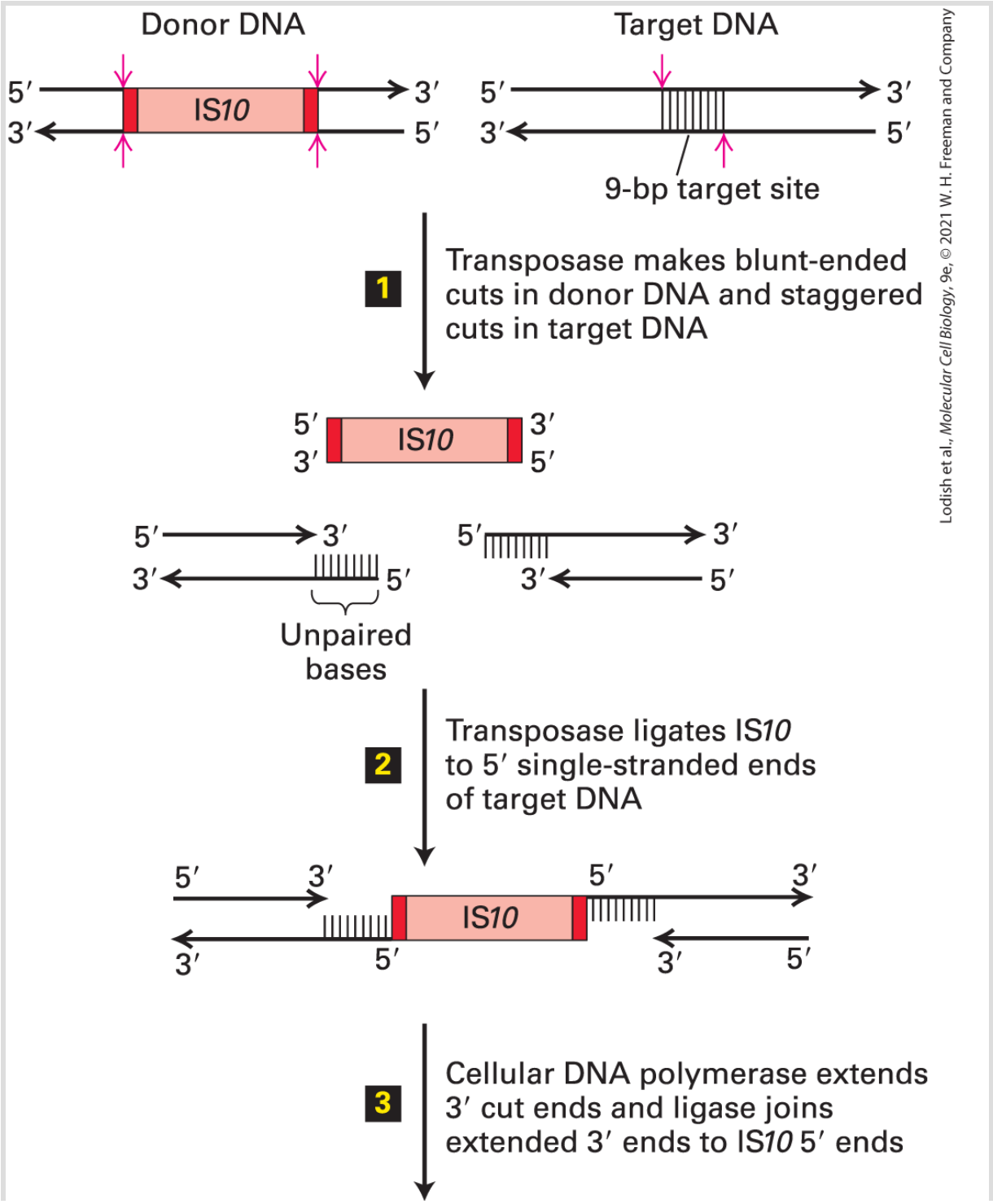
FIGURE 7-9 General structure of bacterial IS elements. The relatively large central region of an IS element, which encodes one or two enzymes required for transposition, is flanked by an inverted repeat at each end. The sequences of the inverted repeats are represented by the red regions, with white arrows indicating the direction of the sequence from $5'$ to $3'$. They are nearly identical, but are oriented in opposite directions. The sequence of the inverted repeat is characteristic of a particular type of IS element. The $5'$ and $3'$ short *direct repeats* (as opposed to the *inverted repeats*) are represented by black regions with a white arrow indicating the direction of the sequence $5'$ to $3'$). These short direct repeats are not transposed with the insertion element; rather, they are insertion-site sequences that become duplicated, with one copy at each end, during insertion of a mobile element. The length of the direct repeats is constant for a given IS element, but their sequence depends

on the site of insertion and therefore varies with each transposition of the IS element. The regions in this diagram are not to scale; the coding region makes up most of the length of an IS element.

Between the inverted repeats in an IS element is a region that encodes a *transposase*, an enzyme required for transposition of the IS element to a new site. The transposase is rarely expressed, accounting for the low frequency of transposition. A hallmark of IS elements is the presence of a 5–11-bp *direct repeat* sequence immediately adjacent to the ends of the inserted element. The *length* of the direct repeat is characteristic of each type of IS element, but its *sequence* depends on the target site where a particular copy of the IS element inserted. When the sequence of a mutated gene containing an IS element is compared with the wild-type gene sequence, only one copy of the direct repeat is found in the wild-type gene. Duplication of this target-site sequence to create the second direct repeat adjacent to an IS element occurs during the insertion process.

As depicted in Figure 7-10, transposition of an IS element occurs by a cut-and-paste mechanism. Transposase performs three functions in this process: it (1) precisely excises the IS element from the donor DNA (Figure 7-10, step **1**), (2) makes staggered cuts in a short sequence in the target DNA (Figure 7-10, step **1**), and (3) ligates the 3′ termini of the IS element to the 5′ ends of the cut donor DNA (Figure 7-10, step **2**). Finally, a host-cell DNA polymerase fills in the

single-stranded gaps, generating the short direct repeats that flank IS elements, and DNA ligase joins the free ends (Figure 7-10, step 3).



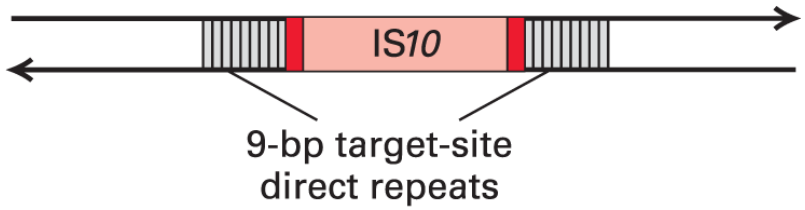


FIGURE 7-10 Model for transposition of bacterial insertion sequences. Step **1**: Transposase, which is encoded by the IS element (IS*10* in this example), cleaves both strands of the donor DNA containing the IS element (*top left*) next to the inverted repeats (dark red, see Figure 7-9 where the inverted repeats are also diagrammed in dark red), excising the IS*10* element. At a largely random target site, transposase makes staggered cuts in the target DNA (*top right*). In the case of IS*10*, the two cuts are 9 bp apart. Thin red vertical arrows indicate the phosphodiester bonds cut by the transposase. Step **2**: Ligation of the $3'$ ends of the excised IS element to the $5'$ ends in target site DNA left by the staggered cut in the target DNA is also catalyzed by transposase. Step **3**: The 9-bp gaps of single-stranded DNA left in the resulting intermediate are filled in by a cellular DNA polymerase. Finally, cellular DNA ligase forms the $3' \rightarrow 5'$ phosphodiester bonds between the $3'$ ends of the extended target DNA strands and the $5'$ ends of the IS*10* strands. This process results in duplication of the target-site sequence on each side of the inserted IS element. Note that the lengths of the target site and IS*10* are not to scale. See H. W. Benjamin and N. Kleckner, 1989, *Cell* **59**:373; and 1992, *Proc. Nat'l. Acad. Sci. USA* **89**:4648.

Eukaryotic DNA Transposons Move Using a Cut-and-Paste Process

McClintock's original discovery of mobile elements came from her observation of spontaneous mutations in maize that affect the production of enzymes required to make anthocyanin, a purple pigment in maize kernels. Mutant kernels are white and wild-type kernels are purple. One class of these mutations is revertible at high

frequency, whereas a second class does not revert unless the mutations occur in the presence of the first class of mutations. McClintock called the agents responsible for the first class of mutations *activator (Ac) elements* and those responsible for the second class *dissociation (Ds) elements* because they also tended to be associated with chromosome breaks.

Many years after McClintock's pioneering discoveries, cloning and sequencing revealed that Ac elements are equivalent to bacterial IS elements. Like IS elements, they contain inverted terminal repeat sequences that flank a region encoding a transposase, which recognizes the terminal repeats and catalyzes transposition to a new site in the host DNA. Ds elements are deleted forms of Ac elements in which a portion of the sequence encoding transposase is missing. Because it does not encode a functional transposase, a Ds element cannot move by itself. However, in plants that carry Ac elements and thus express a functional transposase, Ds elements can be transposed because they retain the inverted terminal repeats recognized by the transposase.

Since McClintock's early work on mobile elements in corn, DNA transposons have been identified in other eukaryotes. For instance, *Drosophila* possess a DNA transposon known as a *P element*, which moves by a mechanism similar to that used by bacterial insertion sequences. Current methods for constructing transgenic *Drosophila* depend on engineered, high-level expression of the P-element

transposase and use of the P-element inverted terminal repeats as targets for transposition.

DNA transposition by the cut-and-paste mechanism can result in an increase in the copy number of a transposon if it occurs during S phase of the cell cycle (see Figure 1-22), when DNA synthesis occurs. Transposon copy number increases when the donor DNA is in one of the two daughter DNA molecules in a region of a chromosome that has replicated but the target DNA is in a region that has not yet replicated. When DNA replication is complete at the end of S phase, the transposon DNA in its new location has also been replicated, resulting in a net increase in the total number of transposon copies in the cell (Figure 7-11). When such a transposition occurs during S phase preceding meiosis, one of the four germ cells produced contains the extra copy of the transposon. Repetition of this process over evolutionary time has resulted in the accumulation of large numbers of DNA transposons in the genomes of some organisms. Human DNA contains about 300,000 copies of full-length and deleted DNA transposons, amounting to about 3 percent of human DNA. As we will see shortly, this mechanism can lead to the transposition of genomic DNA as well as the transposon itself.

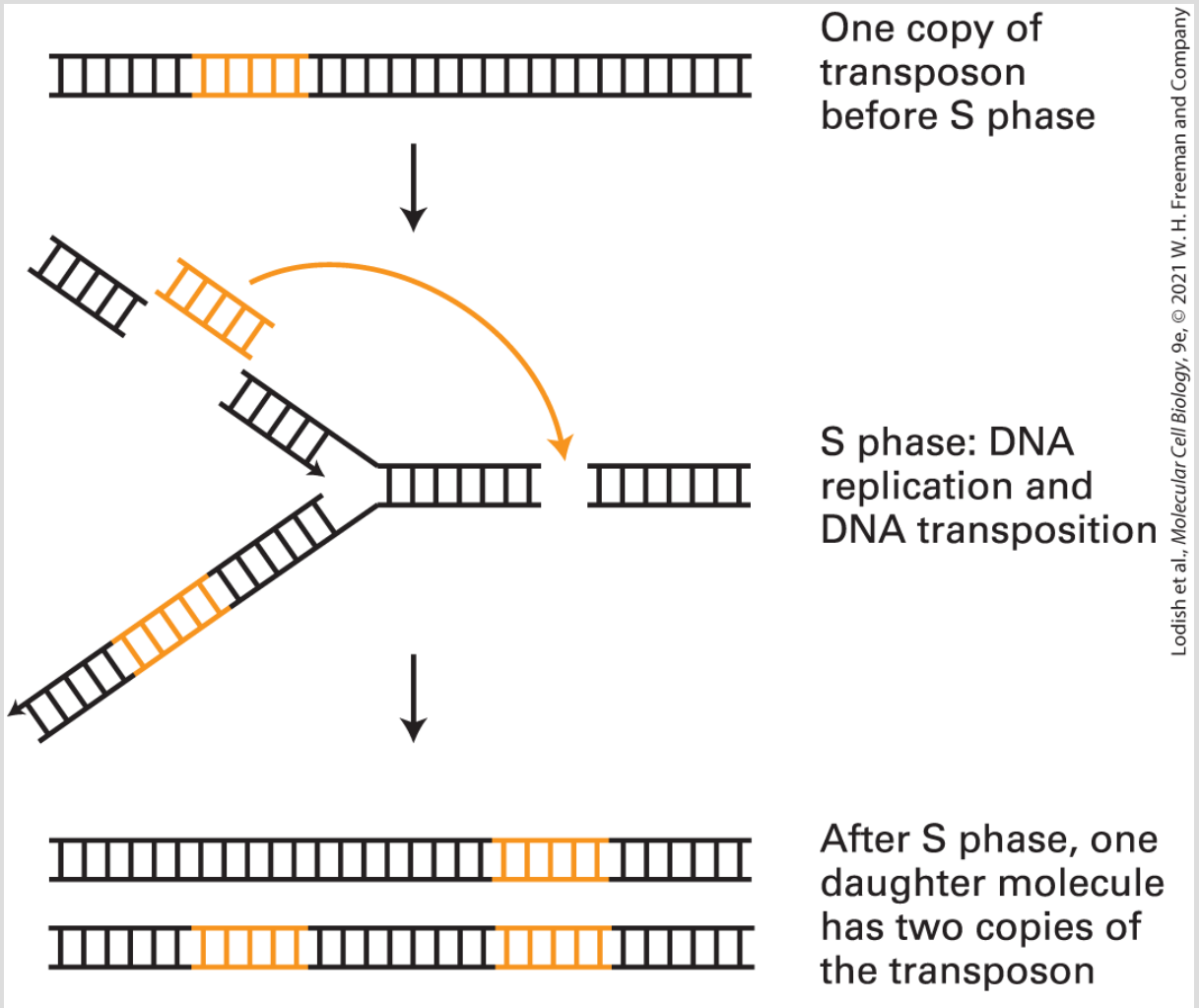


FIGURE 7-11 Mechanism for increasing DNA-transposon copy number. If a DNA transposon, which transposes by a cut-and-paste mechanism (see Figure 7-10), transposes during S phase from a region of the chromosome that has replicated to a region that has not yet replicated, then when chromosomal replication is completed, one of the two daughter chromosomes will have a net increase of one copy of the transposon.

LTR Retrotransposons Behave Like Intracellular Retroviruses

The genomes of all eukaryotes studied, from yeast to humans, contain retrotransposons, mobile DNA elements that transpose through an RNA intermediate using a reverse transcriptase (see Figure 7-8b). These mobile elements are divided into two major categories: those containing and those lacking **long terminal repeats (LTRs)**. LTR retrotransposons, which we discuss here, are common in yeast (e.g., Ty elements) and in *Drosophila* (e.g., *copia* elements). In humans, LTR retrotransposons constitute about 8 percent of the genomic DNA. Non-LTR retrotransposons are the most common type of mobile element in mammals; these retrotransposons are described in the next section.

The general structure of LTR retrotransposons found in eukaryotes is depicted in Figure 7-12. In addition to the short direct repeats that are typical of all transposons, these retrotransposons are marked by the presence of LTRs flanking the central protein-coding region. The LTRs in LTR retrotransposons are 250–600 bp long. Unlike the inverted repeats that flank IS elements and DNA transposons, LTRs are *direct* repeats.

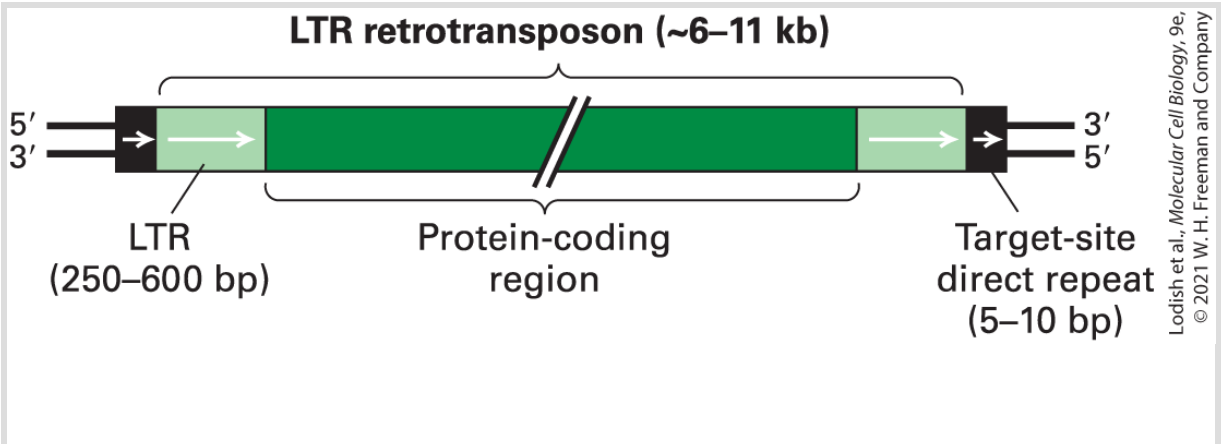


FIGURE 7-12 General structure of eukaryotic LTR retrotransposons. The central protein-coding region is flanked by two long terminal repeats (LTRs), which are element-specific direct repeats. Like other mobile elements, integrated retrotransposons have short target-site direct repeats at each end. Note that the different regions are not drawn to scale. The protein-coding region constitutes 80 percent or more of a retrotransposon and encodes reverse transcriptase, integrase, and other proteins required for transposition.

LTR retrotransposons and infectious retroviruses have much in common. The LTRs in retrotransposons are characteristic of integrated retroviral DNA and are critical to the life cycle of retroviruses and retrotransposons. In addition to sharing LTRs with retroviruses, LTR retrotransposons encode all the proteins of the most common type of retroviruses, except for the envelope proteins. Lacking these envelope proteins, LTR retrotransposons cannot bud from their host cell and infect other cells; however, they can transpose to new sites in the DNA of their host cell. Because of their clear relationship with retroviruses, LTR retrotransposons are often called *retrovirus-like elements*.

A key step in the retroviral life cycle is the formation of retroviral genomic RNA from integrated retroviral DNA (see Figure 5-43). We describe this process in some detail here because it serves as a model for the generation of the RNA intermediate during the transposition of LTR retrotransposons. As depicted in Figure 7-13, the leftward retroviral LTR functions as a promoter that directs host-cell RNA polymerase to initiate transcription at the 5′ nucleotide of the roughly 20-base R sequence that is repeated at each end of the

retroviral RNA. After the entire downstream retroviral DNA has been transcribed, the RNA sequence corresponding to the rightward LTR directs host-cell RNA-processing enzymes to cleave the primary transcript and add a poly(A) tail at the 3′ end of the R sequence. The resulting retroviral RNA genome, which lacks a complete LTR, exits the nucleus and is packaged into a virion that buds from the host cell.

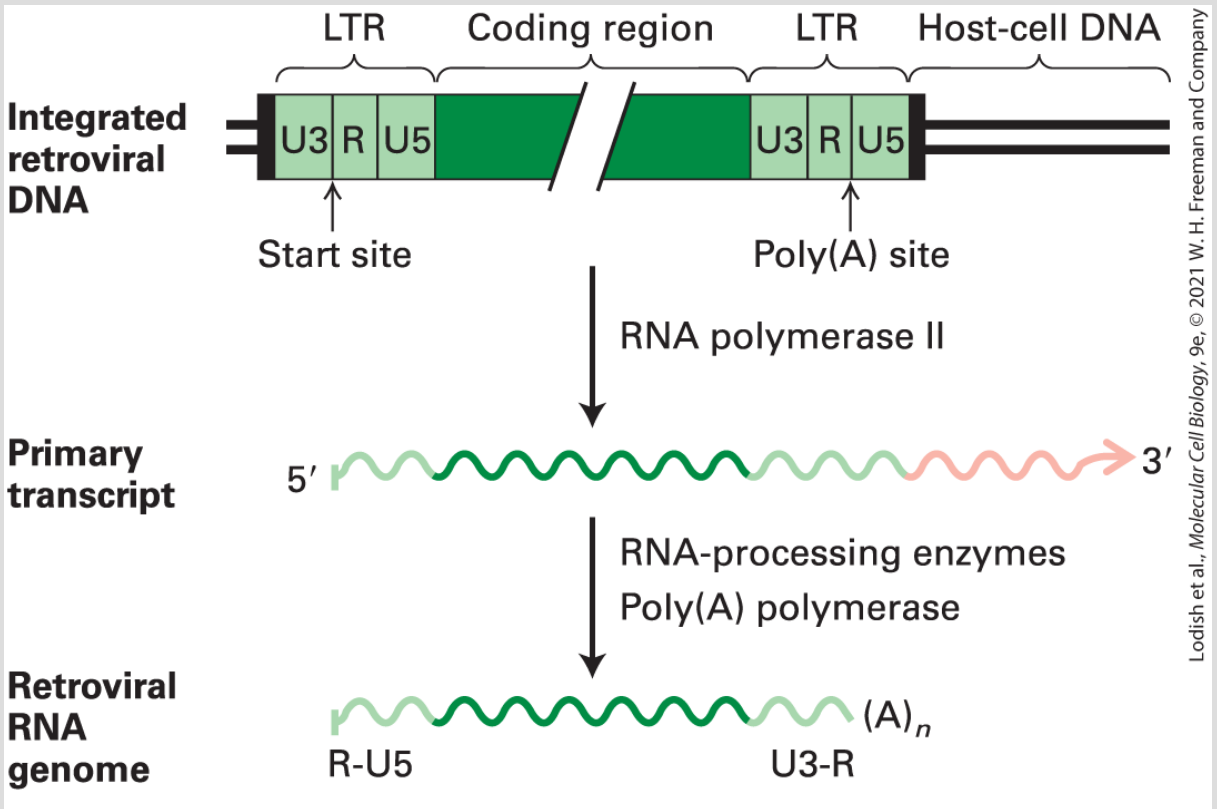


FIGURE 7-13 Generation of retroviral genomic RNA from integrated retroviral DNA. The left LTR of the integrated provirus DNA functions as a promoter, and the right LTR sequence functions as a poly(A) site. (See text.) R represents an ~20 base sequence that is precisely repeated at each end of the viral genomic RNA (bottom). U5 and U3 refer to sequences immediately adjacent to the R sequence at the 5′- and 3′-ends of the genomic RNA, respectively. While U5 and U3 become repeated in the sequence of the integrated provirus DNA (see Figure 7-14), they are not repeated in the genomic retroviral RNA (hence the names

U5 and U3 where “U” indicates a unique sequence in the viral RNA). Short direct repeat sequences (black) flanking the integrated provirus are generated during integration of the retroviral DNA into the host-cell genome. A similar mechanism is thought to generate the RNA intermediate during transposition of LTR retrotransposons.

After a retrovirus infects a cell, reverse transcription of its RNA genome by the retrovirus-encoded reverse transcriptase yields a double-stranded DNA containing complete LTRs (Figure 7-14). This DNA synthesis takes place in the cytosol. The double-stranded DNA, with an LTR at each end, is then transported into the nucleus in a complex with integrase, another enzyme encoded by retroviruses and retrotransposons. Retroviral integrases are closely related to the transposases encoded by DNA transposons and use a similar mechanism to insert the double-stranded retroviral DNA into the host-cell genome (see Figure 7-10), generating short direct repeats of the target-site sequence at either end of the inserted viral DNA sequence. This complex mechanism of reverse transcription is a critical aspect of the retrovirus life cycle because it generates the complete 5′ LTR that functions as a promoter for initiation of transcription precisely at the 5′ nucleotide of the R sequence (Figure 7-13), and the complete 3′ LTR that functions as a poly(A) site leading to polyadenylation precisely at the 3′ nucleotide of the R sequence. Consequently, no nucleotides are lost from an LTR retrotransposon as it undergoes successive rounds of insertion, transcription, reverse transcription, and reinsertion at a new site.

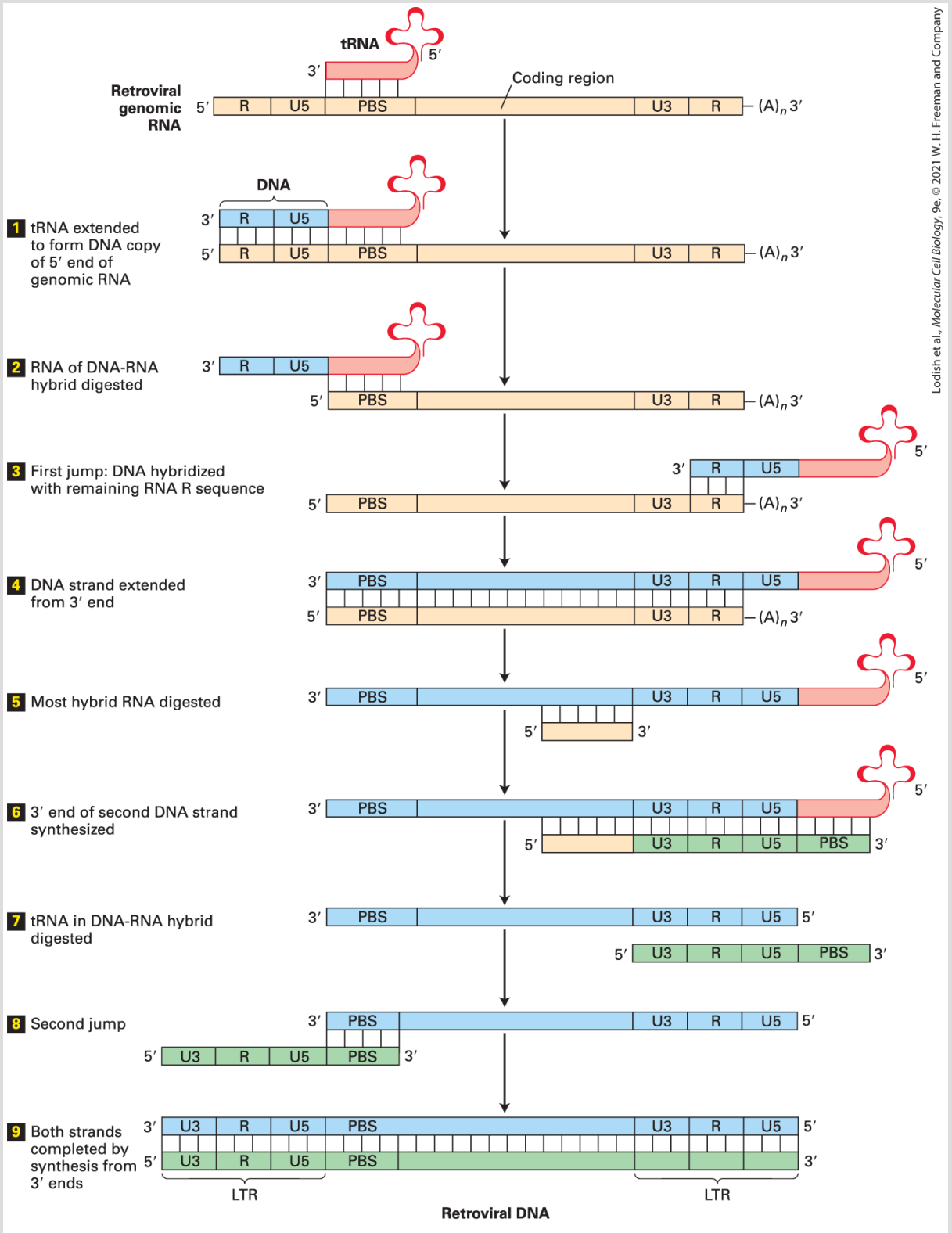
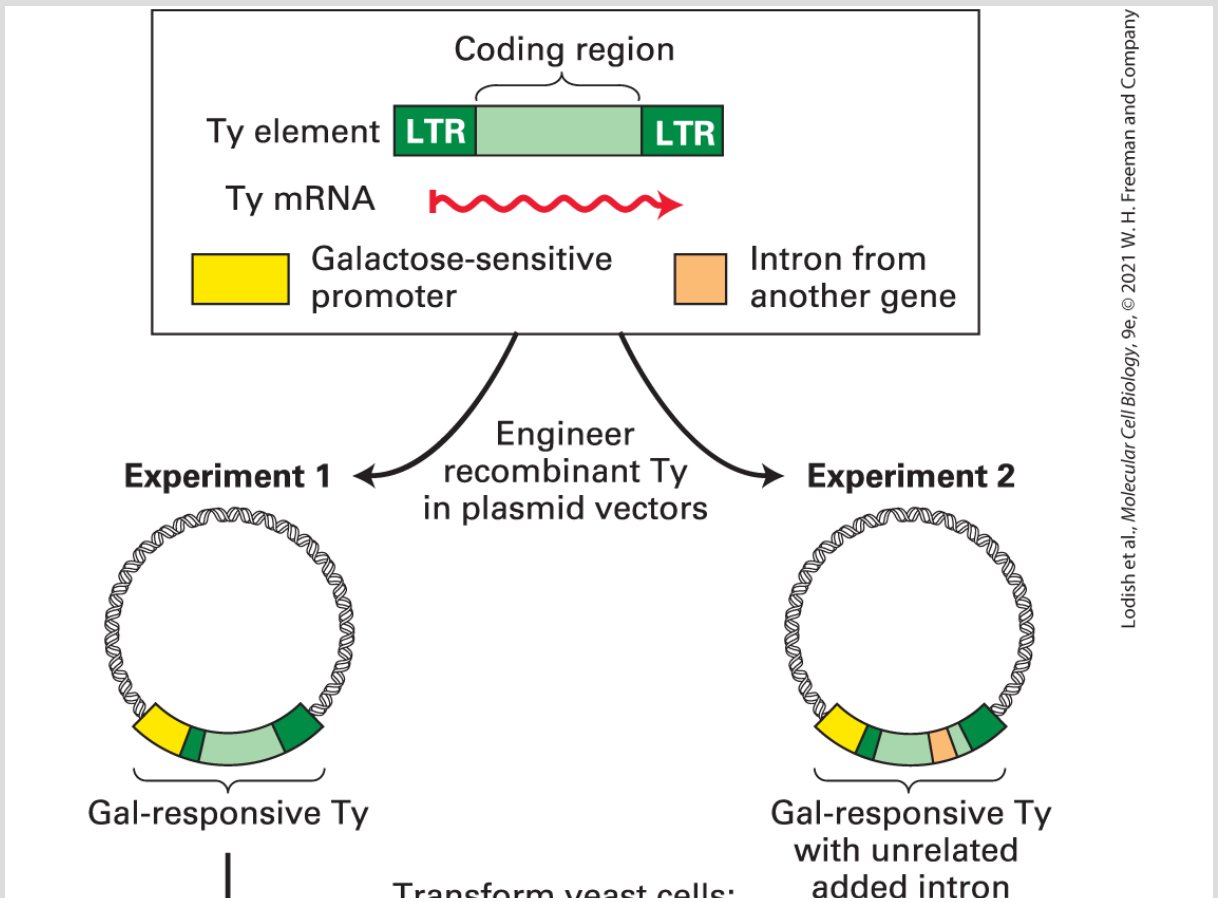


FIGURE 7-14 Model for reverse transcription of retroviral genomic RNA into DNA. (Figure caption is on the facing page.) In this model, a complicated series of nine events generates a

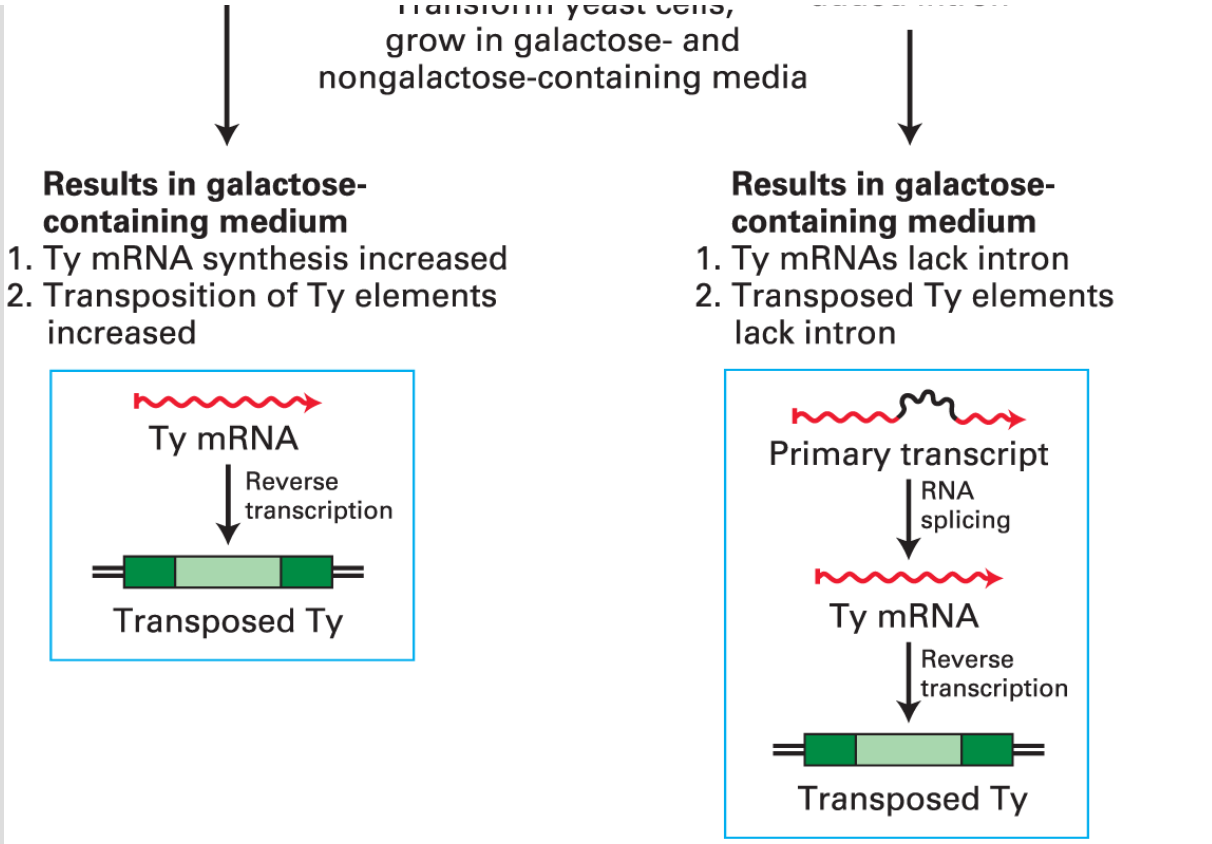
double-stranded DNA copy of the single-stranded RNA genome of a retrovirus. The genomic RNA is packaged in the virion with a retrovirus-specific cellular tRNA hybridized to a complementary sequence near its $5'$ end, called the *primer-binding site* (PBS). The retroviral RNA has a short direct repeat terminal sequence (R) at each end. The overall reaction is carried out by reverse transcriptase, which catalyzes polymerization of deoxyribonucleotides. RNase H, also encoded in the viral RNA and packaged into the virion particle, digests the RNA strand in a DNA-RNA hybrid. The entire process yields a double-stranded DNA molecule that is longer than the template RNA and has a long terminal repeat (LTR) at each end. The different regions are not shown to scale. The PBS and R regions are actually much shorter than the U5 and U3 regions, and the central coding region is very much longer than the other regions. See E. Gilboa et al., 1979, *Cell* **18**:93.

As noted above, LTR retrotransposons encode reverse transcriptase and integrase. By analogy with retroviruses, these mobile elements move by a copy-and-paste mechanism whereby reverse transcriptase converts an RNA copy of a donor element into DNA, which is inserted into a target site by integrase. The experiments depicted in Figure 7-15 provided strong evidence for the role of an RNA intermediate in the transposition of Ty LTR retrotransposons in yeast. When yeast cells are transformed with a Ty-containing plasmid, the Ty element can transpose to new sites, although normally this occurs at a low rate. Using the elements diagrammed at the top of the figure, researchers engineered two different recombinant plasmid vectors containing recombinant Ty elements adjacent to a strong galactose-activated promoter. Yeast cells transformed with these plasmids were grown in a galactose-containing medium to induce transcription of the Ty element, or in a galactose-free medium where Ty transcription is not induced. In experiment 1, growth of cells in galactose-containing medium

resulted in many more transpositions than in galactose-free medium, indicating that transcription into an mRNA intermediate is required for Ty transposition. In experiment 2, an intron from an unrelated yeast gene was inserted into the putative protein-coding region of the recombinant galactose-responsive Ty element. This intron sequence was lost from Ty elements transposed to new sites. This absence of the intron in transposed Ty elements was strong evidence that transposition involves an mRNA intermediate from which the intron was removed by RNA splicing, as depicted in the box on the right. In contrast, eukaryotic DNA transposons, such as the Ac element of maize, contain introns within the transposase gene, indicating that they do not transpose via an RNA intermediate.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company



EXPERIMENTAL FIGURE 7-15 The yeast Ty element transposes through an RNA intermediate. Ty transcription from a plasmid was experimentally manipulated by fusing the Ty element to a *GAL1* transcription control region. When transcription was increased by culturing the cells in medium with galactose (experiment 1), the frequency of Ty transposition was greatly increased. When an intron from another yeast gene was inserted into the middle of the Ty sequence in this plasmid (experiment 2), addition of galactose to the media stimulated transposition of Ty elements lacking the intron. This indicates that transposition likely involves an RNA intermediate from which the intron was removed by RNA splicing, as depicted in the box on the right. See J. Boeke et al., 1985, *Cell* **40**:491.

The most common LTR retrotransposons in humans are called *ERVs*, for *e*ndogenous *r*etro*v*iruses. Most of the 443,000 ERV-related DNA sequences in the human genome consist only of isolated LTRs. These sequences are derived from full-length proviral DNA by

homologous recombination between two LTRs, resulting in deletion of the internal retroviral sequences. Isolated LTRs such as these cannot be transposed to a new position in the genome, but recombination between homologous LTRs at different positions in the genome has probably contributed to the chromosomal DNA rearrangements leading to gene and exon duplications, the evolution of proteins with new combinations of exons, and, as we will see in Chapter 8, the evolution of complex control of gene expression.

Non-LTR Retrotransposons Transpose by a Distinct Mechanism

The most abundant mobile elements in mammals are retrotransposons that lack LTRs, sometimes called *nonviral retrotransposons*. These moderately repeated DNA sequences form two classes in mammalian genomes: **long interspersed elements (LINEs)** and **short interspersed elements (SINEs)**. In humans, full-length LINEs are about 6 kb long, and SINEs are about 300 bp long (see Table 7-1). Repeated sequences with the characteristics of LINEs have been observed in protozoans, insects, and plants, but for unknown reasons, they are particularly abundant in the genomes of mammals. SINEs too are found primarily in mammalian DNA. Large numbers of LINEs and SINEs have accumulated in mammalian genomes over evolutionary time by repeated copying of sequences

at a few positions in the genome and insertion of the copies into new positions.

LINEs

Human DNA contains three major families of LINEs, L1, L2, and L3, that are similar in their mechanism of transposition but differ in their sequences. Only members of the L1 family transpose in the contemporary human genome; apparently there are no remaining functional copies of L2 or L3 to provide the proteins required for their transposition. LINE sequences are present at roughly 900,000 sites in the human genome, accounting for a staggering 21 percent of total human DNA. The general structure of a complete LINE is diagrammed in Figure 7-16. LINEs are usually flanked by short direct repeats, the hallmark of mobile elements, and contain two long open reading frames (ORFs, which are protein-coding regions; see Section 6.3). ORF1, about 1 kb long, encodes an RNA-binding protein. ORF2, about 4 kb long, encodes a protein that has a long region of homology with the reverse transcriptases of retroviruses and LTR retrotransposons, but also exhibits DNA endonuclease activity.

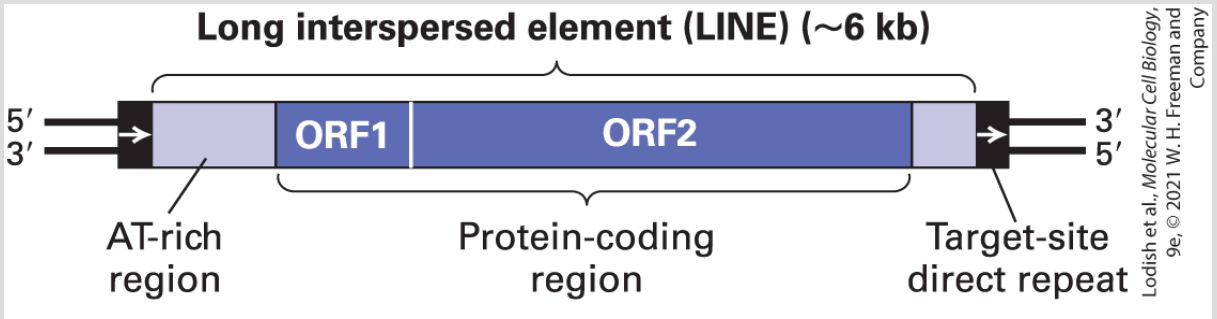
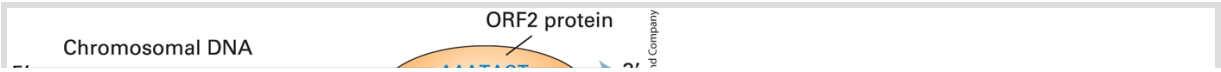


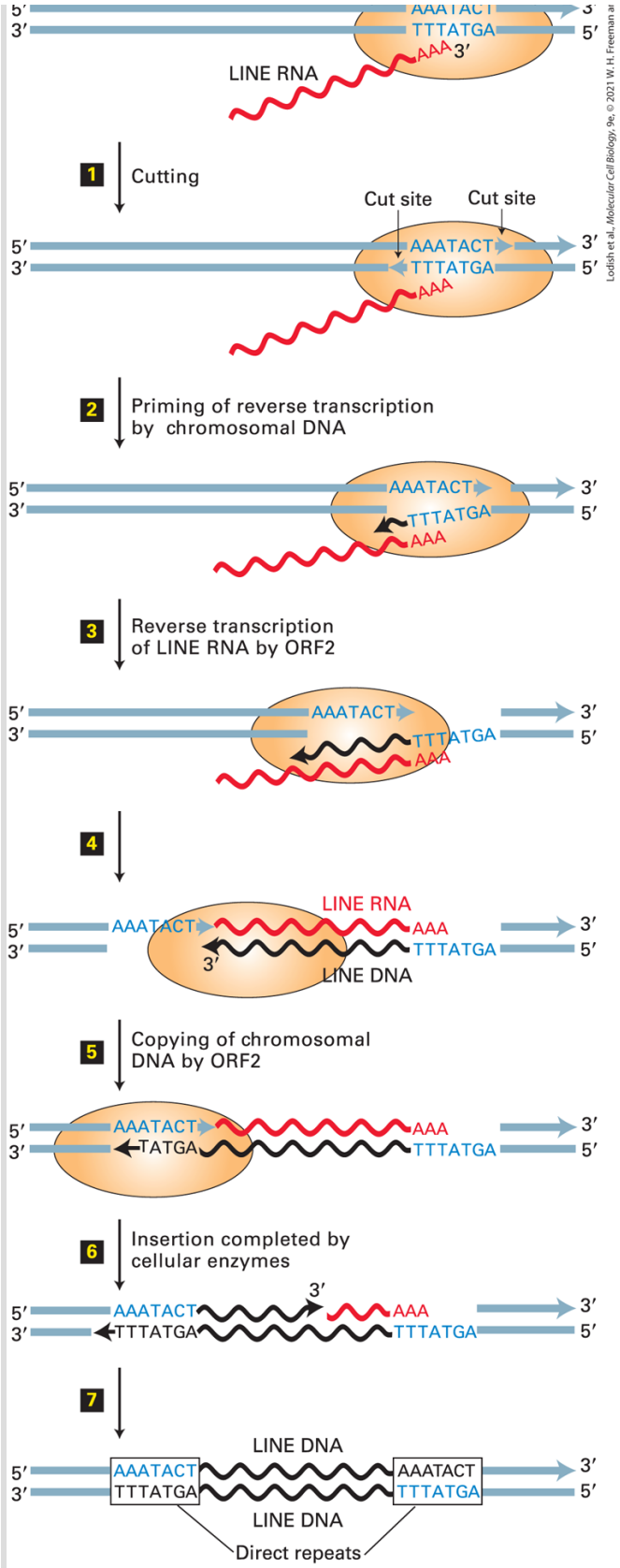
FIGURE 7-16 General structure of a LINE. The length of the target-site direct repeats varies among the copies of a LINE at different sites in the genome. Although the full-length LINE element is about 6 kb long, variable amounts of the left end are absent at over 90 percent of the sites where this mobile element is found. The shorter open reading frame (ORF1), about 1 kb in length, encodes an RNA-binding protein. The longer ORF2, about 4 kb in length, encodes a bifunctional protein with reverse transcriptase and DNA endonuclease activity. Note that LINEs lack the long terminal repeats found in LTR retrotransposons.

Evidence for the mobility of L1 elements first came from analysis of DNA cloned from patients with certain genetic diseases such as hemophilia and myotonic dystrophy. DNA from these patients was found to carry mutations resulting from insertion of an L1 element into a gene, whereas no such element occurred within that gene in either parent. About 1 in 600 mutations that cause significant disease in humans are due to L1 transpositions or SINE transpositions that are catalyzed by L1-encoded proteins. Later experiments similar to those just described with yeast Ty elements (see Figure 7-15) confirmed that L1 elements transpose through an RNA intermediate. In these experiments, an intron was introduced into a cloned mouse L1 element and the recombinant L1 element was stably transfected into cultured hamster cells. After several cell doublings, a DNA fragment corresponding to the L1 element but

lacking the inserted intron was detected in the cells. This finding strongly suggests that, over time, the recombinant L1 element containing the inserted intron had transposed to new sites in the hamster genome through an RNA intermediate that underwent RNA splicing to remove the intron. ■

Since LINEs do not contain LTRs, their mechanism of transposition through an RNA intermediate differs from that of LTR retrotransposons. In vitro studies indicate that transcription by RNA polymerase is directed by promoter sequences at the left end of integrated LINE DNA. LINE RNA is polyadenylated by the same post-transcriptional mechanism that polyadenylates other mRNAs. The LINE RNA is then exported into the cytosol, where it is translated into ORF1 and ORF2 proteins. Multiple copies of ORF1 protein bind to the LINE RNA, and ORF2 protein binds to the poly(A) tail. The LINE RNA is then transported back into the nucleus as a complex with ORF1 and ORF2 proteins, where it is reverse-transcribed into LINE DNA by ORF2. The mechanism involves staggered cleavage of cellular DNA at the insertion site, followed by priming of reverse transcription by the resulting cleaved cellular DNA, as detailed in Figure 7-17. The complete process results in insertion of a copy of the original LINE retrotransposon into a new site in chromosomal DNA. A short direct repeat is generated at the insertion site because of the initial staggered cleavage of the two chromosomal DNA strands.





Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and

FIGURE 7-17 Proposed mechanism of LINE reverse transcription and integration. Only ORF2 protein is represented here. Newly synthesized LINE DNA is shown in black. ORF1 and ORF2 proteins, produced by translation of LINE RNA in the cytoplasm, bind to LINE RNA and transport it into the nucleus. Step **1**: In the nucleus, ORF2 makes staggered cuts in AT-rich target-site DNA, generating the DNA $3'$-OH ends indicated by blue arrowheads. Step **2**: The $3'$ end of the T-rich DNA strand hybridizes to the poly(A) tail of the LINE RNA and primes DNA synthesis by ORF2. Step **3**: ORF2 extends the DNA strand using the LINE RNA as a template. Steps **4** and **5**: When synthesis of the LINE DNA bottom strand reaches the $5'$ end of the LINE RNA template, ORF2 extends the newly synthesized LINE DNA using as a template the top-strand cellular DNA generated by the initial ORF2 staggered cleavage. Step **6**: A cellular DNA polymerase extends the $3'$ end of the top strand generated by the initial ORF2 staggered cut, using the newly synthesized bottom-strand LINE DNA as a template. The LINE RNA is digested as the DNA polymerase extends the upper-strand DNA (this also occurs during removal of lagging-strand primer RNA during cellular DNA synthesis; see Figure 5-11). The $3'$ ends of the newly synthesized DNA strands are ligated to the $5'$ ends of the cellular DNA strands as in lagging-strand cellular DNA synthesis. See D. D. Luan et al., 1993, *Cell* **72**:595.

As noted already, the DNA form of an LTR retrotransposon is synthesized from its RNA form in the cytosol using a cellular tRNA as a primer for reverse transcription of the first strand of DNA (see Figure 7-14). The resulting double-stranded DNA with long terminal repeats is then transported into the nucleus, where it is integrated into chromosomal DNA by a retrotransposon-encoded integrase. In contrast, the DNA form of a non-LTR retrotransposon is synthesized in the nucleus. The synthesis of the first strand of the non-LTR retroviral DNA by ORF2, a reverse transcriptase, is primed by the $3'$ end of cleaved chromosomal DNA, which base-pairs with the poly(A) tail of the non-LTR RNA (Figure 7-17, step **1**). Since its synthesis is primed by the cut end of a cleaved chromosome, and since synthesis

of the other strand of the non-LTR retrotransposon DNA is primed by the 3′ end of chromosomal DNA on the other side of the initial cut (step **6**), the mechanism of synthesis results in integration of the non-LTR retrotransposon DNA. There is no need for an integrase to insert the non-LTR retrotransposon DNA. Because its synthesis begins with reverse transcription of a poly(A) tail on the LINE RNA, one end of a non-LTR retrotransposon is AT-rich.

The vast majority of LINEs in the human genome are truncated at their 5′ end, suggesting that reverse transcription was terminated before completion and that the resulting fragments, extending variable distances from the poly(A) tail, were inserted. Because of this shortening, the average size of LINE elements is only about 900 bp, even though the full-length sequence is about 6 kb long. Truncated LINE elements, once formed, probably are not further transposed because they lack a promoter for transcription of the RNA intermediate. In addition to the fact that most L1 insertions are truncated, nearly all the full-length elements contain stop codons and frameshift mutations in ORF1 and ORF2; these mutations have probably accumulated in most LINE sequences over evolutionary time. As a result, only about 0.01 percent of the LINE sequences in the human genome, or about 60 in total number, are full-length, with intact open reading frames for ORF1 and ORF2.

SINEs

The most abundant class of mobile elements in the human genome, SINEs constitute about 13 percent of total human DNA. Varying in length from about 100 to 400 base pairs, these retrotransposons do not encode protein, but most contain a 3′ AT-rich sequence similar to that in LINEs. SINEs are transcribed by RNA polymerase III, the same nuclear RNA polymerase that transcribes genes encoding tRNAs, 5S rRNAs, and other small stable RNAs. Most likely, the ORF1 and ORF2 proteins expressed from full-length LINEs mediate reverse transcription and integration of SINEs by the mechanism depicted in Figure 7-17. Consequently, SINEs can be viewed as parasites of the LINE symbionts, competing with LINE RNAs for binding, reverse transcription, and integration by LINE-encoded ORF1 and ORF2.

SINEs occur at about 1.6 million sites in the human genome. Of these, about 1.1 million are *Alu elements*, so named because they contain a single recognition site for the restriction enzyme *Alu*I. *Alu* elements exhibit considerable sequence homology with, and probably evolved from, 7SL RNA, a cytosolic RNA in a ribonucleoprotein complex called the signal recognition particle. This abundant cytosolic ribonucleoprotein particle aids in targeting certain polypeptides to the membranes of the endoplasmic reticulum (see Chapter 13). *Alu* elements are scattered throughout the human genome at sites where their insertion has not disrupted gene expression: between genes, within introns, and in 3′ untranslated regions. For instance, nine *Alu* elements are located

within the 80-kb human β-globin gene cluster (see Figure 7-4a). Of the new germ-line non-LTR retrotranspositions that are estimated to occur about once in every eight individuals, about 40 percent involve L1 elements and 60 percent involve SINEs, of which about 90 percent of the latter are *Alu* elements. Like other mobile elements, most SINEs have accumulated mutations from the time of their insertion in the germ line of an ancient ancestor of modern humans. Like LINEs, many SINEs are truncated at their 5′ ends.

Other Retroposed RNAs Are Found in Genomic DNA

In addition to the mobile elements listed in Table 7-1, DNA copies of a wide variety of mRNAs appear to have become integrated into chromosomal DNA. Since these sequences lack introns and do not have flanking sequences similar to those of functional gene copies, they clearly are not simply duplicated genes that have drifted into nonfunctionality and become pseudogenes, as discussed earlier (see Figure 7-4a). Instead, these DNA segments appear to be retrotransposed copies of spliced and polyadenylated mRNA. Compared with normal genes encoding mRNAs, these inserted segments generally contain multiple mutations, which are thought to have accumulated since they were first reverse-transcribed and randomly integrated into the genome of a germ cell in an ancient ancestor. These nonfunctional genomic copies of mRNAs are

referred to as *processed pseudogenes*. Most processed pseudogenes are flanked by short direct repeats, supporting the hypothesis that they were generated by rare retrotransposition events involving cellular mRNAs.

Other interspersed repeats representing partial or mutant copies of genes encoding small nuclear RNAs (snRNAs) and tRNAs are found in mammalian genomes. Like processed pseudogenes derived from mRNAs, these nonfunctional copies of small RNA genes are flanked by short direct repeats and most likely result from rare retrotransposition events that have accumulated through the course of evolution. Enzymes expressed from a LINE are thought to have carried out these retrotransposition events involving mRNAs, snRNAs, and tRNAs.

Mobile DNA Elements Have Significantly Influenced Evolution

Although most mobile DNA elements appear to have no direct function other than to maintain their own existence, their presence has had a profound effect on the evolution of modern-day organisms. About half the spontaneous mutations in *Drosophila* result from insertion of a mobile DNA element into or near a transcription unit. In mammals, mobile elements cause a much smaller proportion of spontaneous mutations: about 10 percent in

mice and 0.1–0.2 percent in humans. Still, mobile elements have been found in mutant alleles associated with several human genetic diseases. For example, insertions into the clotting factor IX gene cause hemophilia, and insertions into the gene encoding the muscle protein dystrophin lead to Duchenne muscular dystrophy. The genes encoding clotting factor IX and dystrophin are both on the X chromosome. Consequently, disease resulting from a transposition into these genes occurs primarily in males, in which there is no second copy of the normal gene to complement the resulting mutation.

In lineages leading to multicellular eukaryotes, homologous recombination between mobile DNA elements dispersed throughout ancestral genomes may have generated gene duplications and other DNA rearrangements during evolution (see Figure 7-2b). For instance, cloning and sequencing of the β-globin gene clusters from various primate species has provided strong evidence that the human *HGB1* and *HGB2* genes encoding fetal β-globins arose from an unequal homologous crossover between two L1 elements flanking an ancestral globin gene. Subsequent divergence of the duplicated genes led to the acquisition of distinct, beneficial functions by each duplicated gene. Unequal crossing over between mobile elements located within introns of a particular gene could lead to the duplication of exons within that gene (see Figure 7-2a). This process most likely influenced the evolution of genes that

contain multiple copies of similar exons encoding similar protein domains, such as the fibronectin gene (see Figure 5-28).

Some evidence suggests that during the evolution of multicellular eukaryotes, recombination between mobile DNA elements (e.g., *Alu* elements) in introns of *two separate* genes also occurred, generating new genes made from novel combinations of preexisting exons (Figure 7-18). This evolutionary process, termed **exon shuffling**, may have occurred during the evolution of the genes encoding tissue plasminogen activator, the Neu receptor, and epidermal growth factor, all of which contain an EGF domain (see Figure 3-11). In this case, exon shuffling presumably resulted in the insertion of an EGF domain–encoding exon into an intron of the ancestral form of each of these genes.

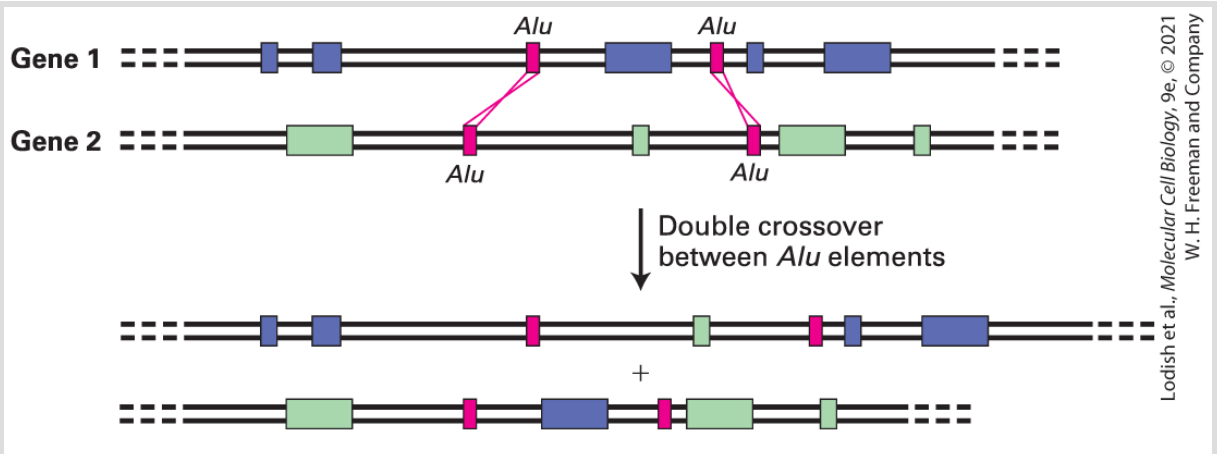


FIGURE 7-18 Exon shuffling via recombination between homologous interspersed repeats. Recombination between interspersed repeats in the introns of separate genes produces transcription units with new combinations of exons. In the example shown here, a double crossover between two sets of *Alu* repeats results in an exchange of exons between two genes.

Both DNA transposons and LINE retrotransposons have been shown to occasionally carry unrelated flanking sequences — including exons — when they transpose to new sites by the mechanisms diagrammed in Figure 7-19. This occasional transposition of flanking DNA sequences probably also contributed to exon shuffling during the evolution of contemporary genes.

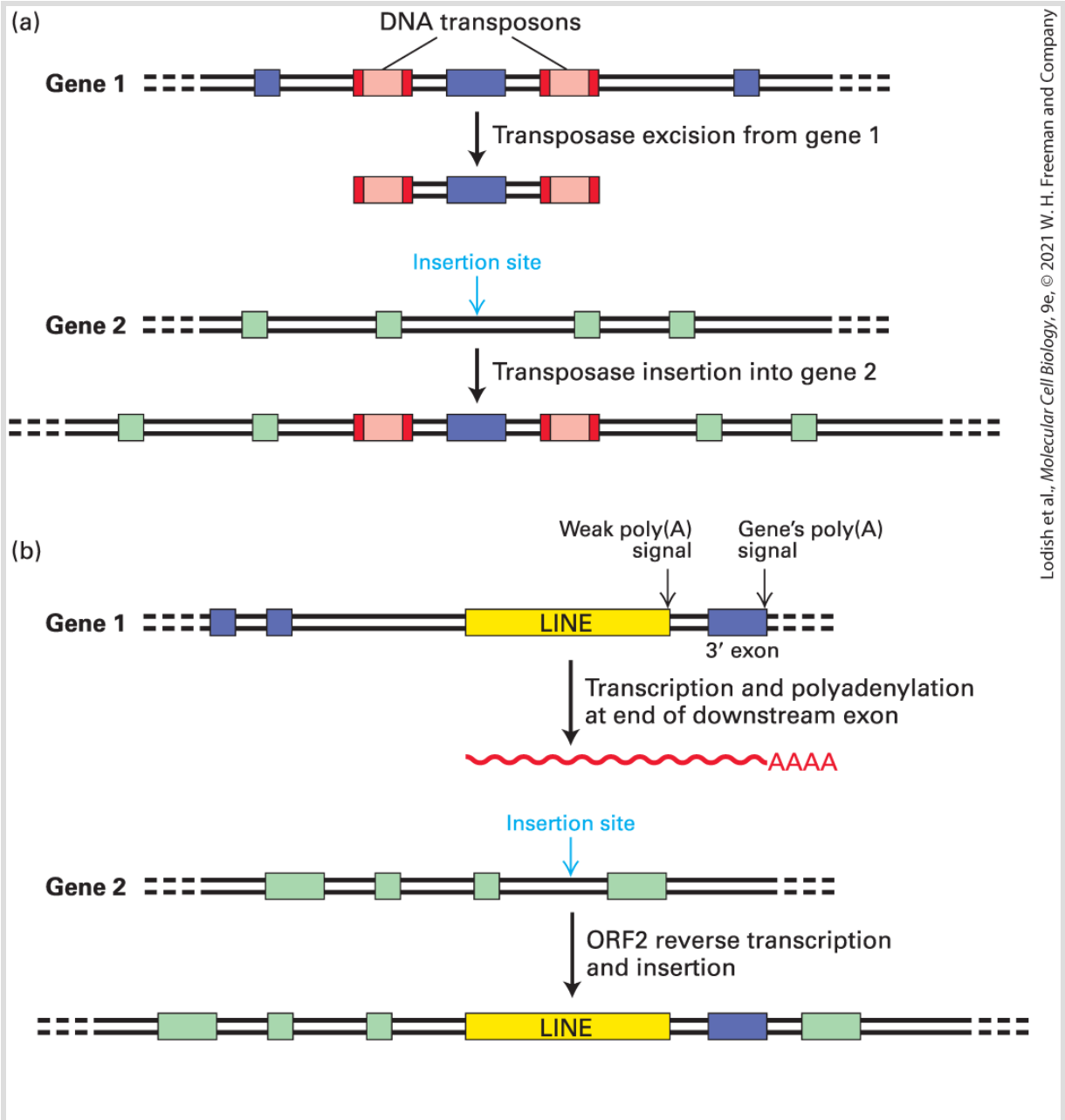


FIGURE 7-19 Exon shuffling by transposition. (a) Transposition of an exon flanked by homologous DNA transposons into an intron on a second gene. As we saw in Figure 7-10, step **1**, transposase can recognize and cleave the DNA at the ends of the transposon inverted repeats. In gene 1, if the transposase cleaves at the left end of the transposon on the left and at the right end of the transposon on the right, it can transpose all the intervening DNA, including the exon from gene 1, to a new site in an intron of gene 2. The net result is an insertion of the exon from gene 1 into gene 2. (b) Integration of an exon into another gene via LINE transposition. Some LINEs have weak poly(A) signals. If such a LINE is in the $3'$-most intron of gene 1, during transposition its transcription may continue beyond its own poly(A) signals and extend into the $3'$ exon, transcribing the cleavage and polyadenylation signals of gene 1 itself. This RNA can then be reverse-transcribed and integrated by the LINE ORF2 protein (see Figure 7-17) into an intron on gene 2, introducing a new $3'$ exon (from gene 1) into gene 2.

In addition to causing changes in coding sequences, recombination between mobile elements and transposition of DNA adjacent to mobile elements probably played a significant role in the evolution of regulatory sequences that control gene expression. As noted earlier, eukaryotic genes have transcription-control regions called enhancers that can operate over distances of tens of thousands of base pairs. The transcription of many genes is controlled through the combined effects of several enhancer elements. Insertion of mobile elements near such transcription-control regions probably contributed to the evolution of new combinations of enhancer sequences. These combinations, in turn, control which specific genes are expressed in particular cell types and the amount of the encoded protein produced in modern organisms, as we discuss in the next chapter.

These considerations suggest that the early view of mobile DNA elements as completely selfish molecular parasites misses the mark. Rather, these elements have contributed profoundly to the evolution of multicellular organisms by promoting (1) the generation of gene families via gene duplication, (2) the creation of new genes via shuffling of preexisting exons, and (3) the formation of more complex regulatory regions that provide multifaceted control of gene expression. Today researchers are attempting to harness transposition mechanisms to insert therapeutic genes into patients as a form of gene therapy.

A process analogous to that shown in Figure 7-19a is largely responsible for the rapid spread of antibiotic resistance among pathogenic bacteria, a major problem in modern medicine. Bacterial genes encoding enzymes that inactivate antibiotics (drug resistance genes) have become flanked by insertion sequences, generating drug resistance transposons. The widespread use of antibiotics in medicine and agriculture has led to positive selection on drug resistance transposons. These transposons are often found incorporated into conjugating plasmids. Conjugating plasmids encode proteins that result in their replication and transfer to other bacterial cells — even cells of other related bacterial species — through a complex macromolecular tube called a pilus. These plasmids, called *R factors* (for drug *r*esistance), can carry multiple drug resistance genes introduced by transposition and selected in environments, such as hospitals, where antibiotic-laced cleaning

solutions are used to sterilize surfaces. R factors have led to the rapid spread of resistance to multiple antibiotics among pathogenic bacteria. Coping with the spread of R factors is a major challenge for modern medicine. ■

KEY CONCEPTS OF SECTION 7.3

Transposable (Mobile) DNA Elements

- Transposable DNA elements are moderately repeated sequences interspersed at multiple sites throughout the genomes of multicellular eukaryotes. They are present less frequently in unicellular eukaryotic and prokaryotic genomes.
- DNA transposons move to new sites directly as DNA; retrotransposons are first transcribed into an RNA copy of the element, which is then reverse-transcribed into DNA (see Figure 7-8).
- A common feature of all mobile elements is the presence of short direct repeats flanking the element, generated as the result of staggered cuts in the target-site DNA that are filled in by a DNA polymerase during transposition (see Figure 7-10).
- Enzymes called *transposases* encoded by transposons themselves catalyze insertion of these sequences at new sites in genomic DNA.
- Although DNA transposons, similar in structure to bacterial IS elements, occur in eukaryotes (e.g., the *Drosophila* P element), retrotransposons are generally much more abundant, especially in vertebrates.
- LTR retrotransposons are flanked by long terminal repeats (LTRs) similar to those in retrovirus proviral DNA; like retroviruses, they encode reverse transcriptase and integrase. They move in the genome by being transcribed into RNA, which then undergoes reverse transcription in the cytosol, nuclear import of the resulting DNA with LTRs, and integration into a host-cell chromosome (see Figure 7-14).
- Non-LTR retrotransposons, including long interspersed elements (LINEs) and short interspersed elements (SINEs), lack LTRs and have an AT-rich stretch at one end. They are thought to move by a nonviral retrotransposition mechanism mediated by LINE-encoded proteins involving priming of reverse transcription by chromosomal DNA (see Figure 7-17).
- SINE sequences exhibit extensive homology with 7SL small cellular RNA and are transcribed by RNA polymerase III. *Alu* elements, the most common SINEs in humans,

are sequences of about 300 bp found scattered throughout the human genome at about 1.6 million sites (see Figure 7-4a).

- Some interspersed repeats are derived from cellular RNAs that were reverse-transcribed and inserted into genomic DNA at some time in evolutionary history. Processed pseudogenes derived from mRNAs lack introns, a feature that distinguishes them from pseudogenes, which arose by sequence drift of duplicated genes.
- Mobile DNA elements most likely influenced evolution significantly by serving as sites for homologous recombination during unequal crossing over, leading to gene and exon duplication (see Figure 7-2) and exon shuffling (see Figure 7-18), and by mobilizing adjacent DNA sequences (see Figure 7-19).

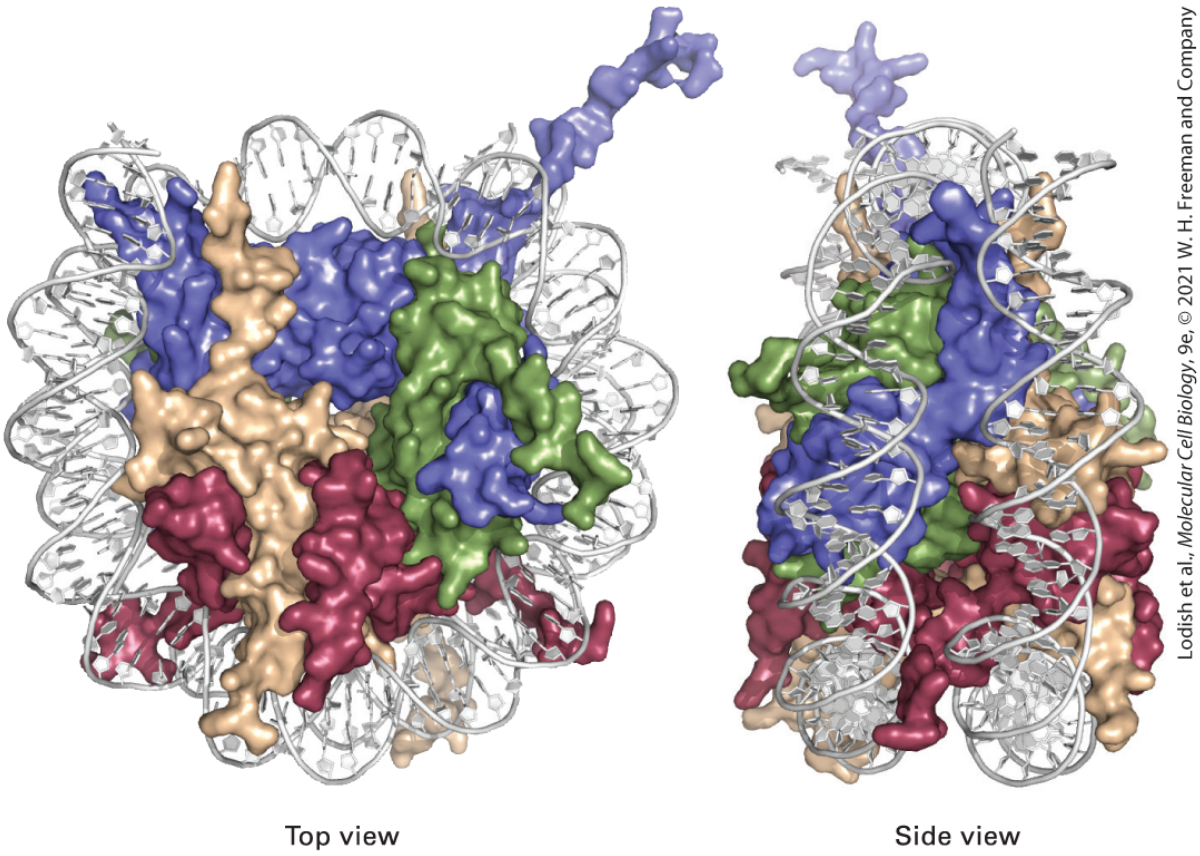
7.4 Structural Organization of Eukaryotic Chromatin and Chromosomes

Now that we have examined the various types of DNA sequences found in eukaryotic genomes and how they are organized within the single, extremely long DNA molecules contained in each chromosome, we turn to the question of how DNA molecules of this length are organized within eukaryotic cells. Because the total length of cellular DNA (~2 meters in humans) is roughly 100,000 times the diameter of the nucleus (~20 μm for most human cells), the packing of DNA is crucial to cell architecture. It is also essential to prevent the long DNA molecules from getting knotted or tangled with each other during cell division, when they must be precisely segregated to daughter cells.

Packing chromosomal DNA into the nucleus requires a number of proteins, in particular, histones. In this section, we consider the structural organization of histones and DNA to form chromosomes. We also examine the dynamic nature of these structures as they respond to gene expression signals and the changes that occur throughout the cell cycle.

Chromatin Is Made of Nucleosomes

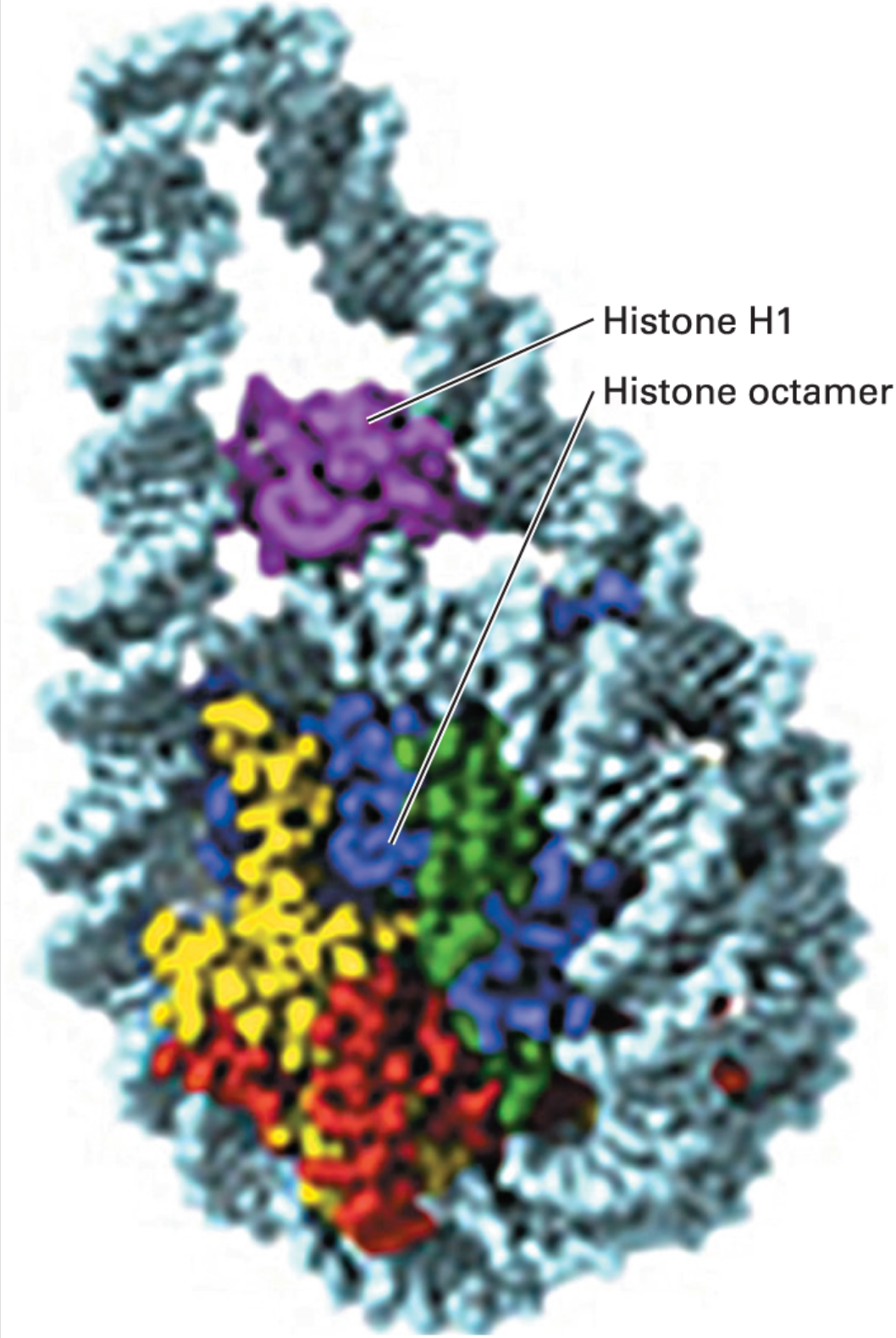
The key molecular structure that allows chromosomal DNA to undergo this incredible compaction is the **nucleosome** (Figure 7-20), which consists of 145–147 base pairs of DNA wrapped tightly ~1 2/3 turns around a symmetrical, roughly globular protein complex called the **histone octamer**. The histone octamer is made of two copies each of four small (~15 kDa), basic proteins. There are five major types of **histone** proteins in eukaryotes termed *H2A, H2B, H3*, and *H4*, for the four histones that make up the histone octamer (Figure 7-20), and histone *H1*, a *linker histone*. One copy of H1 binds linker DNA at the entry and exit sites of the 1 2/3 DNA wraps of 145–147 base pairs around the histone octamer core of the nucleosome (Figures 7-20 and 7-21). The histones are rich in positively charged basic amino acids, which interact with the negatively charged phosphate groups in DNA. The complex of DNA plus histones and other less abundant proteins is called **chromatin**.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 7-20 Structure of the nucleosome based on x-ray crystallography. Nucleosome shown from the top (*left*) and from the side (*right*, rotated clockwise 90°). DNA is shown as a gray ribbon diagram. Histones are shown as space-filling models with H2A subunits in yellow; H2Bs in red; H3s in blue; and H4s in green. One H2A, H2B heterodimer projects out of the page on the lower right of the side view, while the other H2A, H2B heterodimer projects into the page, on the lower left of the side view. Only one H2A, H2B heterodimer is visible in the top view. The other H2A, H2B dimer is not visible in this view because it is behind the H3, H4 tetramer on the upper right. (See also the ribbon diagram of the histone polypeptide chains in Figure 7-26a, in which an H2A, H2B heterodimer is clearly visible on the lower left of the top view of the nucleosome.)

[Data from K. Luger et al., 1997, *Nature* **389**:251, PDB ID 1aoi.]



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 7-21 Structure of histone H1 interacting with both a nucleosome and linker DNA. Histones H2A, H2B, H3, and H4 are colored as in Figure 7-20. Histone H1 is purple.

[Republished with permission of Nature Publishing Group, from K. Zhou, G. Gaullier, and K. Luger, 2019, *Nat. Struct. Mol. Biol.* **26**: 3–13; permission conveyed through Copyright Clearance Center, Inc.]

Examination of isolated chromatin by standard electron microscopy using staining with heavy metals revealed a "beads-on-a-string" organization of nucleosomes separated by a variable length of DNA of ~50–150 bp called **linker DNA** (Figure 7-22). Until recently, it has not been possible to trace the path of a chromatin fiber in electron micrographs of nuclei of fixed cells because chromatin fibers could not be distinguished from other nuclear macromolecules of similar electron density after staining with standard EM methods. However, a recently developed technique, called *ChromEMT*, has now made this possible. ChromEMT specifically labels DNA in thin sections of fixed cells with the electron opaque compound OsO_4. The resolution is sufficient to reveal chains of nucleosomes that can be traced through EM thin sections (Figure 7-23). Nearly all nuclear DNA was found to be associated with nucleosomes connected by flexible linker DNA. In regions visualized as **euchromatin** in standard electron micrographs, where most expressed genes are located, the linker DNA is extended between nucleosomes (Figure 7-23a) compared to more condensed regions of chromatin, known as **heterochromatin**, where few genes are transcribed. In heterochromatin, the linker DNA is curved and bent so that the

nucleosomes are much closer together than in euchromatin (Figure 7-23a). The regions in the nucleus between nucleosomes in Figure 7-23a that were not stained are not empty but contain RNA as well as thousands of different nuclear proteins that are not stained by ChromEMT. As discussed below, some of these nuclear proteins are thought to control the density of chromatin fibers and the distance between nucleosomes in different regions of the nucleus.

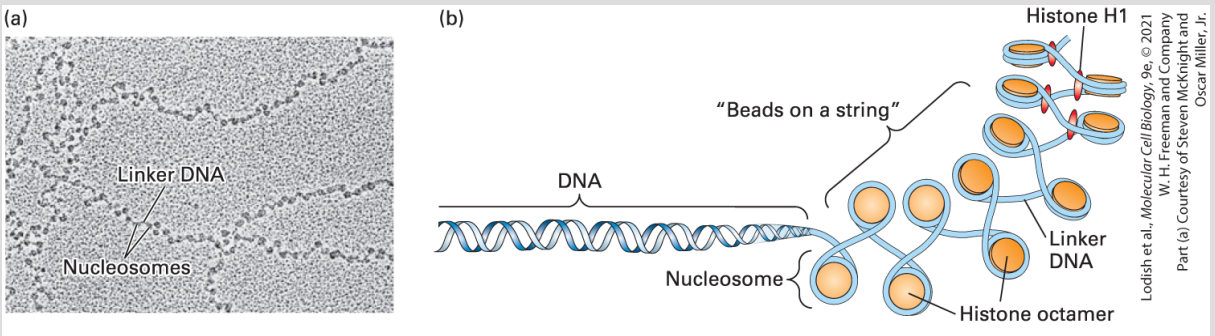


FIGURE 7-22 "Beads-on-a-string" arrangement of nucleosomes in extracted chromatin. (a) Electron micrograph of chromatin extracted from isolated nuclei in a low ionic-strength buffer. The "beads" are nucleosomes (10 nm in diameter) and the "string" is connecting (linker) DNA. (b) A diagram of the beads-on-a-string configuration.

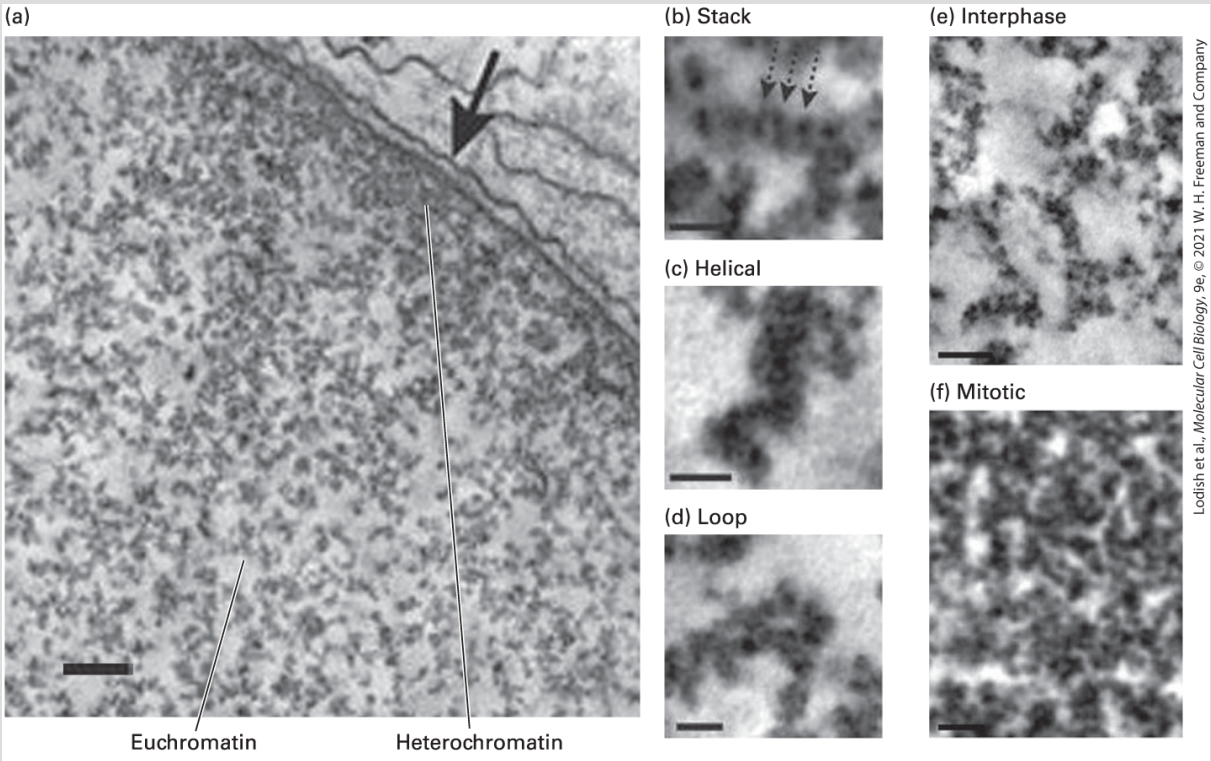


FIGURE 7-23 Heterogeneity in nucleosome interactions and compaction. (a) Human small airway epithelial cells were fixed with glutaraldehyde, stained with ChromEMT (see text), and an EM tomogram was constructed. Scale bar = 100 nm. Arrow indicates the outer nuclear membrane. Regions of euchromatin and heterochromatin are indicated. (b)–(d) DNA and nucleosomes form disordered chromatin chains that have different nucleosome arrangements, conformations, and densities. Three categories of conformations are shown: stack, helical, and loop. Scale bar = 20 nm. (e), (f) Regions of the nucleus from interphase cells (e), and mitotic cells (f). Scale bar = 40 nm.

[Republished with permission from American Association for the Advancement of Science, from H. D. Ou et al., 2017, "ChromEMT: Visualizing 3D Chromatin Structure and Compaction in Interphase and Mitotic Cells," *Science* **357**(6349):eaag0025; https://doi.org/10.1126/science.aag0025; permission conveyed through Copyright Clearance Center, Inc.]

Chromatin is dispersed throughout much of the nucleus in interphase cells (those that are not undergoing mitosis). Further compaction of chromatin during mitosis (see Figure 6-3) produces

the visible *metaphase chromosomes* whose morphology and staining characteristics were detailed by early cytogeneticists. ChromEMT of cells in mitosis reveals that the chromatin fibers are closely packed, making it difficult to trace the chromatin fiber (Figure 7-23f). Although every eukaryotic chromosome includes millions of individual protein molecules, each chromosome contains just one, extremely long, linear DNA molecule. The longest DNA molecules in human chromosomes, for instance, are 2.8×10^8 bp, or almost 10 cm in length. The structural organization of chromatin allows this vast length of DNA to be compacted into the microscopic constraints of a cell nucleus (see Figure 7-1). Despite this degree of compaction, chromatin is organized in such a way that specific DNA sequences within the chromatin are available for cellular processes such as transcription, replication, repair, and recombination.

Chromatin Structure Is Conserved Among Eukaryotes

The general structure of chromatin is remarkably similar in the cells of all eukaryotes, including fungi, plants, and animals, indicating that the structure of chromatin was optimized early in the evolution of eukaryotic cells. The amino acid sequences for the four core histones that make up the nucleosome core (H2A, H2B, H3, and H4) are highly conserved among distantly related species. For example, the sequences of histone H3 from sea urchin and human tissue

differ by only three amino acids, and H3 from the garden pea and human differ by only two amino acids. Apparently, significant deviations from the histone amino acid sequences were selected against during evolution. The similarity in sequence among histones from all eukaryotes indicates that they fold into very similar three-dimensional conformations. The amino acid sequence of H1, however, varies more from organism to organism than do the sequences of the other major histones.

Minor histone variants encoded by genes that differ from the highly conserved major types also exist, particularly in vertebrates. For example, a special form of H2A, designated H2AX, is incorporated in place of H2A in a small fraction of nucleosomes throughout chromatin. At sites of double-stranded breaks in chromosomal DNA, H2AX becomes phosphorylated and participates in the chromosome repair process, probably by functioning as a binding site for repair proteins. In the nucleosomes within centromeres, H3 is replaced by a variant histone called *CENP-A*, which anchors the proteins that bind spindle microtubules during mitosis. Another histone variant, known as *H3.3*, replaces histone H3 in regions of DNA undergoing transcription. The replacement probably occurs when the histone octamer must be moved out of the way by histone chaperones as RNA polymerase is transcribing DNA. These and other minor histone variants differ only slightly in sequence from the major histones. These slight changes in histone sequence may influence the interactions of nucleosomes with other nuclear proteins, the

stability of the nucleosome, and the tendency of chromatin fibers containing the variant histone to condense.

Chromatin Is a Disordered Chain of Nucleosomes Packed Together at Different Concentration Densities in the Nucleus

As mentioned previously, the ChromEMT technique has made it possible to trace chains of nucleosomes separated by linker DNA in thin sections of fixed cells (Figure 7-23). A model for the organization of nucleosomes in a chromatin fiber derived from ChromEMT images of an interphase human cell is shown in Figure 7-24a. These chains of nucleosomes appear to range in diameter from 5 to 24 nm. The narrower dimension is thought to result from a projection of the fiber that shows primarily linker DNA, and the larger dimension from a projection that includes more than one nucleosome. Nucleosomes often contact each other; sometimes appearing to stack on each other (Figure 7-23b), sometimes packing in a helical conformation (Figure 7-23c), and sometimes forming a chain that loops back on itself (Figure 7-23d). Nucleosomes appear to interact with multiple alternative orientations and interacting surfaces. In darker staining regions of heterochromatin adjacent to the inner nuclear membrane (Figure 7-23a, arrow), the 5–24 nm chromatin fibers loop back on themselves frequently and are closer

together than in euchromatin. Figures 7-24b, c show models of heterochromatin and euchromatin based on the ChromEMT images.

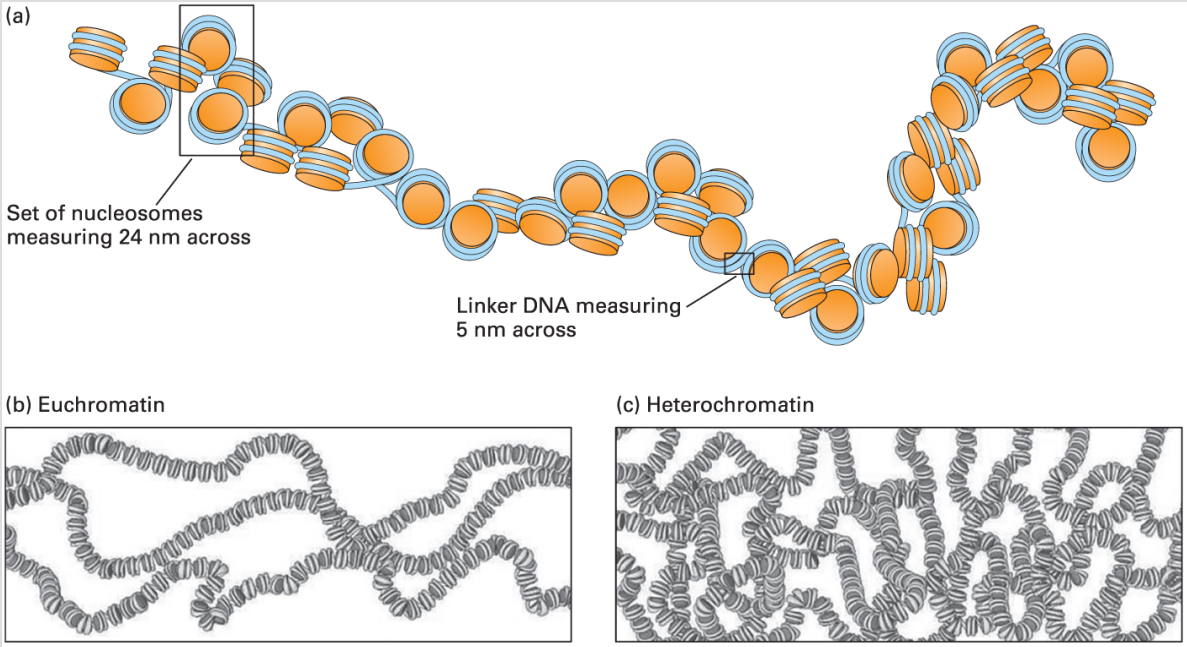


FIGURE 7-24 Chromatin is composed of structurally disordered 5- to 24-nm-diameter chains that are packed together at different densities in euchromatic and heterochromatic regions. (a) Model of the chromatin fiber based on ChromEMT of small airway epithelial cells. Histone octamers are represented as orange discs and DNA as a blue tube. Large rectangle frames a region visualized as 24 nm in diameter. Small rectangle frames a region of 5-nm diameter. (b) and (c) Representation of the density of chromatin fibers in euchromatin and heterochromatin, respectively.

[Part (a) republished with permission from American Association for the Advancement of Science, from H. D. Ou et al., 2017, "ChromEMT: Visualizing 3D Chromatin Structure and Compaction in Interphase and Mitotic Cells," *Science* **357**(6349):eaag0025; https://doi.org/10.1126/science.aag0025; permission conveyed through Copyright Clearance Center, Inc. Parts (b) and (c) republished with permission from American Association for the Advancement of Science, from Daniel R. Larson and Tom Mistel, 2017, "The Genome — Seeing It Clearly Now," *Science* **357**(6349):354–355; https://doi.org/10.1126/science.aao1893; permission conveyed through Copyright Clearance Center, Inc.]

Condensed metaphase chromosomes appear to have a chromatin structure similar to heterochromatin. In addition, in the condensed metaphase chromosome the chromatin fiber loops back on itself frequently compared to euchromatin and there are more interactions between nucleosomes, resulting in a high density of closely packed chromatin chains (Figure 7-23e, f, and Figure 7-24b, c). Scaffolding proteins are thought to constrain and compact the flexible chromatin chains into the mitotic chromosome architecture. One family of these chromatin scaffolding proteins is called *structural maintenance of chromosome (SMC) proteins* (Figure 7-25). Significantly, using the ChromEMT technique, higher order folding of regular helical structures of ~30 nm, once thought to account for chromatin condensation in heterochromatin and during mitosis, was not observed.

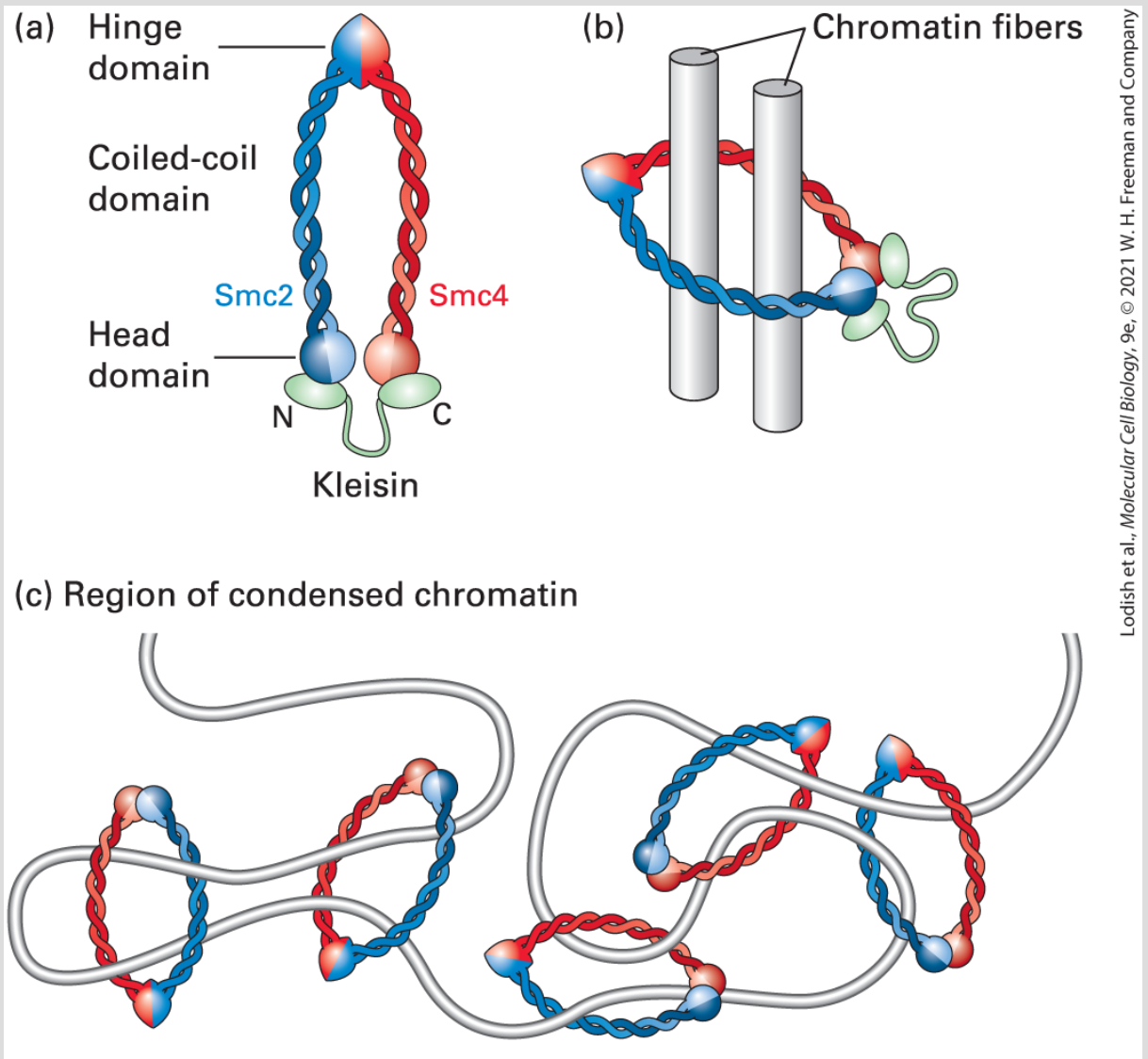


FIGURE 7-25 Model of SMC complexes bound to chromatin. (a) Model of an SMC protein complex. (b) Model of SMC complex topologically linking two chromatin fibers. (c) Model for chromosome condensation by SMC complexes. See K. Nasmyth and C. H. Haering, 2005, *Ann. Rev. Biochem.* **74**:595.

Ringlike Structure of SMC Protein Complexes

SMC proteins are critical for maintaining the structure of condensed chromosomes during mitosis. In extracts prepared from the large

nuclei of *Xenopus laevis* (African frog) eggs, chromosomes can be induced to condense as they do in intact cells as they enter prophase of mitosis. This condensation fails to occur when one type of SMC protein called *condensin* is depleted from the extract with specific antibodies. Yeast with mutations in a related SMC protein called *cohesin* fail to properly associate sister chromatids following DNA replication in the S phase. As a result, chromosomes do not properly segregate to daughter cells during mitosis (see Chapter 19). Structurally related SMC proteins are required for proper segregation of chromosomes in bacteria and archaea, indicating that SMCs are an ancient class of proteins vital to chromosome structure and segregation in all domains of life.

In the general structure of SMC proteins, each SMC monomer contains a hinge region where the polypeptide folds back on itself, forming a very long coiled-coil region and bringing the N- and C-termini together so they can interact to form a globular head domain (Figure 7-25a). The hinge domain of one monomer (blue in Figure 7-25) binds to the hinge domain of a second monomer (red), forming a roughly U-shaped dimeric complex. The head domains of the monomers have ATPase activity and are linked by members of another small protein family called *kleisins*. The overall SMC complex is a ring with a diameter large enough to accommodate two chromatin fibers (Figure 7-25b) and is capable of linking two circular DNA molecules in vitro. Rings of SMC proteins are proposed to fold the chromatin fiber into loops, as diagrammed in Figure 7-

25c. This model can explain why cleavage of the DNA at a relatively small number of sites leads to rapid dissolution of condensed metaphase chromosome structure, whereas protease cleavage of proteins has only a minor effect on chromosome structure until most of the protein is digested. When the DNA is cut anywhere in a long region of chromatin containing several chromatin loops, the broken ends can slip through the SMC protein rings, untying the topological knots that constrain the loops of chromatin. In contrast, a protease must cut most of the individual rings of SMC proteins before the topological constraints forming the loops are released completely.

Modifications of Histone Tails Control Chromatin Condensation and Function

Each of the histone proteins making up the nucleosome core contains an intrinsically disordered, flexible N-terminus of 19–39 residues extending from the globular structure of the nucleosome (Figure 7-26a). The H2A and H2B proteins also contain a flexible C-terminus extending from the globular histone octamer core. These termini, called *histone tails,* are represented as dashed lines in Figure 7-26a in order to represent their lengths relative to the rest of the nucleosome.

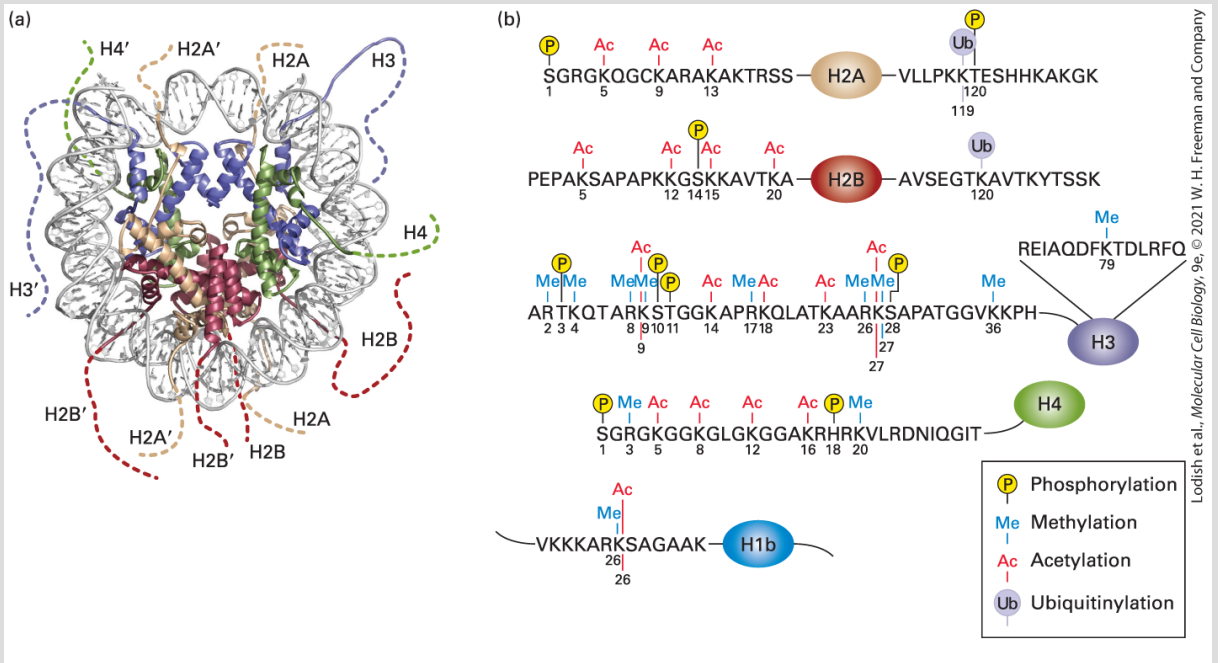


FIGURE 7-26 Post-translational modifications observed on human histones. (a) Model of a nucleosome viewed from the top with histones shown as ribbon diagrams. This model depicts the lengths of the histone tails (dotted lines), which are not visible in the crystal structure (see Figure 7-20). The H2A N-terminal tails are at the bottom, and the H2A C-terminal tails are at the top. The H2B N-terminal tails are toward the bottom on the right and left, and the H2B C-terminal tails are at the bottom center. The H3 N-terminal tails extend between the gyres of DNA near the entry and exit point of DNA wrapping around the histone octamer core, and the H4 N-terminal tails extend downward and upward from the roughly disc-shaped nucleosome. Histones H3 and H4 have short C-terminal tails that are not modified. (b) Summary of post-translational modifications observed in human histones. Histone-tail sequences are shown in the one-letter amino acid code, numbered beginning with 1 at the N-terminus (see Figure 2-14). The main portion of each histone that is part of the histone octamer core is depicted as an oval. These modifications do not all occur simultaneously on a single histone molecule. Rather, specific combinations of a few of these modifications are observed on any one histone. See K. Luger and T. J. Richmond, 1998, *Curr. Opin. Genet. Devel.* **8**:140.

[Part (a) Data from K. Luger et al., 1997, *Nature* **389**:251, PDB ID 1aoi. Part (b) Data from R. Margueron et al., 2005, *Curr. Opin. Genet. Devel.* **15**:163.]

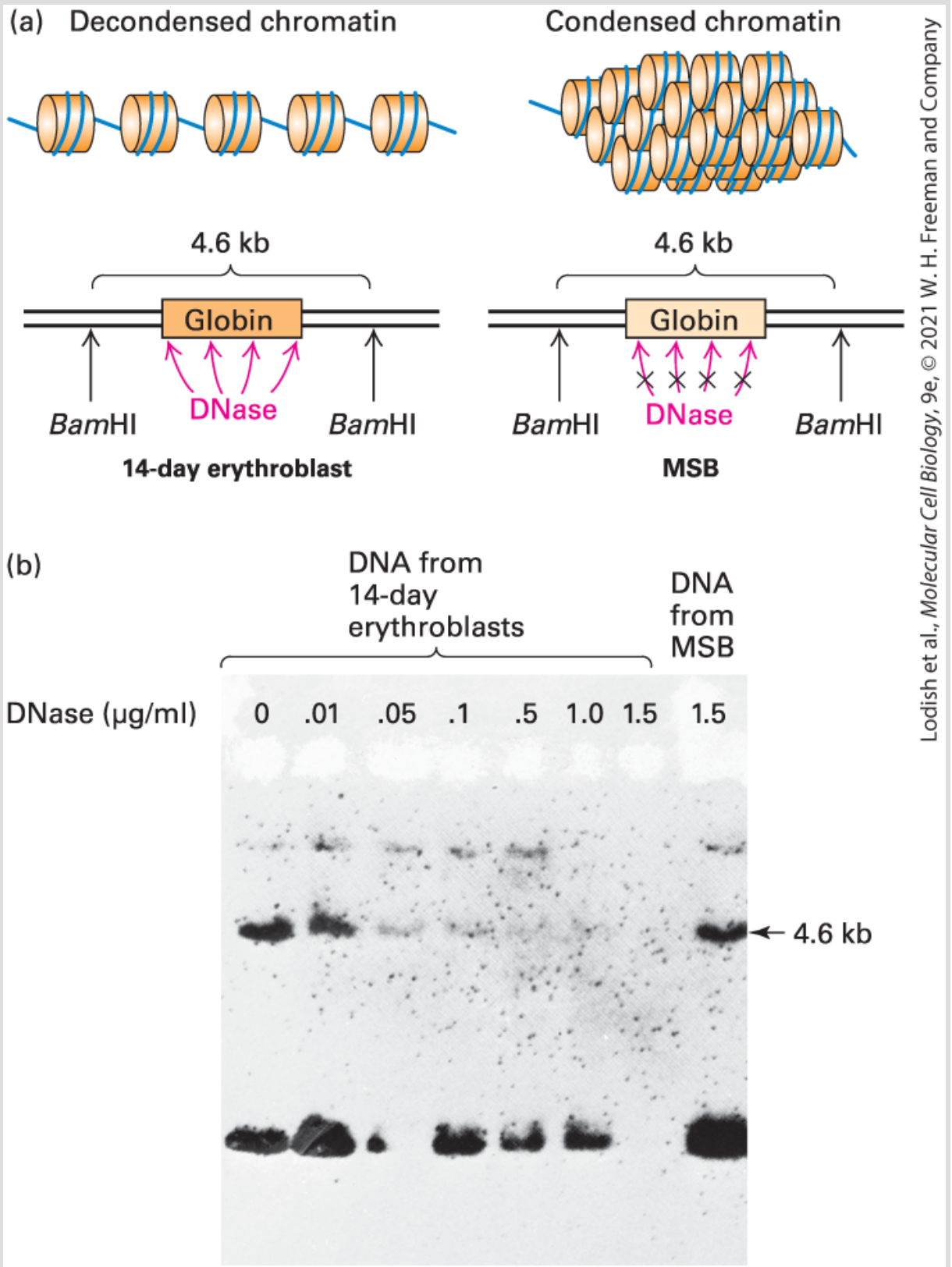
Histone-tail–amino acid residue side chains are subject to multiple types of reversible, post-translational covalent modifications such as acetylation, methylation, phosphorylation, and ubiquitinylation. Figure 7-26b summarizes the most common types of post-translational modifications observed in human histones. A particular histone protein never has all of these modifications simultaneously, but the histones in a single nucleosome may contain several of these modifications simultaneously. The particular combinations of histone-tail post-transcriptional modifications found in different regions of chromatin constitute a *histone code* that influences chromatin function by creating or removing binding sites for chromatin-associated proteins depending on the specific combinations of the modifications present. Here we describe the most abundant kinds of modifications found in histone tails and the association of these modifications with chromatin condensation and function. We end with a discussion of a special case of chromatin condensation, the inactivation of X chromosomes in female mammals.

Histone Acetylation

Histone-tail lysines in the amino termini undergo reversible acetylation and deacetylation. In the acetylated form, the positive charge of the lysine ε-amino group is neutralized. Histone acetylation correlates with increased sensitivity of chromatin DNA to nuclease digestion — an observation indicating that histone

acetylation is associated with a more open form of chromatin. This correlation can be demonstrated by conducting experiments in which isolated nuclei are digested with DNase I. Next the DNA is separated from chromatin protein, digested with a restriction enzyme, and analyzed by Southern blotting. An intact gene treated with a restriction enzyme yields fragments of characteristic sizes. But when the DNA is exposed first to DNase I, there are two possible results seen in the Southern blot. One, the gene is cleaved by DNase I at random sites within the boundaries of the restriction enzyme cut sites. Consequently, any Southern blot bands normally seen with that gene will be lost. Two, the gene is hidden within the chromatin so that DNase I molecules could not access it and cut it up. In that case, the Southern blot would show the characteristic banding pattern for that gene. This method was first used to compare the structure of chromatin in the region of the chicken β-globin gene in cells in which the gene is not transcribed and cells in which the gene is actively transcribed (Figure 7-27). The β-globin gene is transcriptionally inactive in non-erythroid cells (MSB cells), but is actively transcribed in erythroid progenitor cells (erythroblasts). In non-erythroid cells, the histones around the β-globin gene have relatively low levels of acetylation, and the chromatin that contains the β-globin gene is resistant to DNase I. In contrast, in erythroid progenitor cells, the histones around the β-globin gene are acetylated, and the chromatin that contains the β-globin gene is sensitive to DNase I (Figure 7-27). These results indicate that the chromatin structure of non-transcribed DNA in *hypoacetylated*

chromatin makes the DNA less accessible to DNase I than it is in transcribed, *hyperacetylated* chromatin. The proposed model is that in non-erythroid cells, chromatin containing the β-globin gene is folded into condensed structures with closely spaced nucleosomes that sterically inhibit access to the associated DNA by the nuclease, as for the region of heterochromatin near the nuclear envelope indicated in Figure 7-23a. In contrast, the transcribed gene is associated with a more unfolded form of chromatin, which allows the nuclease better access to the associated DNA, as in the regions designated euchromatin in Figure 7-23a. More generally, it is presumed that the condensed chromatin structure around the β-globin gene in non-erythroid cells sterically inhibits the proteins involved in transcription from accessing the β-globin promoter and other transcription control sequences in the DNA. Thus histone acetylation and deacetylation appear to be one means of regulating transcription (see Chapter 8).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 7-27 Nontranscribed genes are less susceptible to DNase I digestion than active genes. Chick embryo erythroblasts at 14 days actively synthesize globin, whereas undifferentiated chicken lymphoblastic leukemia (MSB) cells do not. (a) Nuclei from each type of cell were isolated and exposed to increasing concentrations of DNase I. The nuclear DNA was then extracted and treated with the restriction enzyme *Bam*HI, which cleaves the DNA around the globin sequence and normally releases a 4.6-kb globin DNA fragment. (b) The DNase I- and *Bam*HI-digested DNA was subjected to Southern blot analysis with a probe of labeled cloned adult globin DNA, which hybridizes to the 4.6-kb *Bam*HI fragment. If the globin gene is susceptible to the initial DNase digestion, it should be cleaved repeatedly and would not be expected to show this fragment. As seen in the Southern blot, the transcriptionally active DNA from the 14-day globin-synthesizing cells was sensitive to DNase I digestion, indicated by the absence of the 4.6-kb band at high nuclease concentrations. In contrast, the inactive DNA from MSB cells was resistant to digestion. These results suggest that the inactive DNA is in a more condensed form of chromatin in which the globin gene is shielded from DNase digestion. See J. Stalder et al., 1980, *Cell* **19**:973.

[Part (b) Republished with permission of Elsevier, from J. Stalder et al., 1980, "Hb Switching in Chickens," *Cell* **19**(4):973–980; permission conveyed through Copyright Clearance Center, Inc.]

Genetic studies in yeast indicated that *histone acetyl transferases (HATs)*, the enzymes that acetylate specific lysine residues in histones, are required for activation of transcription of a number of genes. These enzymes are now known to have other substrates that influence gene expression in addition to histones. Consequently, they are more generally known as *nuclear lysine acetyl transferases*, or *KATs*, because *K* represents lysine in the single-letter code for amino acids (see Figure 2-14). Conversely, early genetic studies in yeast indicated that complete repression of many yeast genes requires the action of *histone deacetylases (HDACs)* that remove acetyl groups of

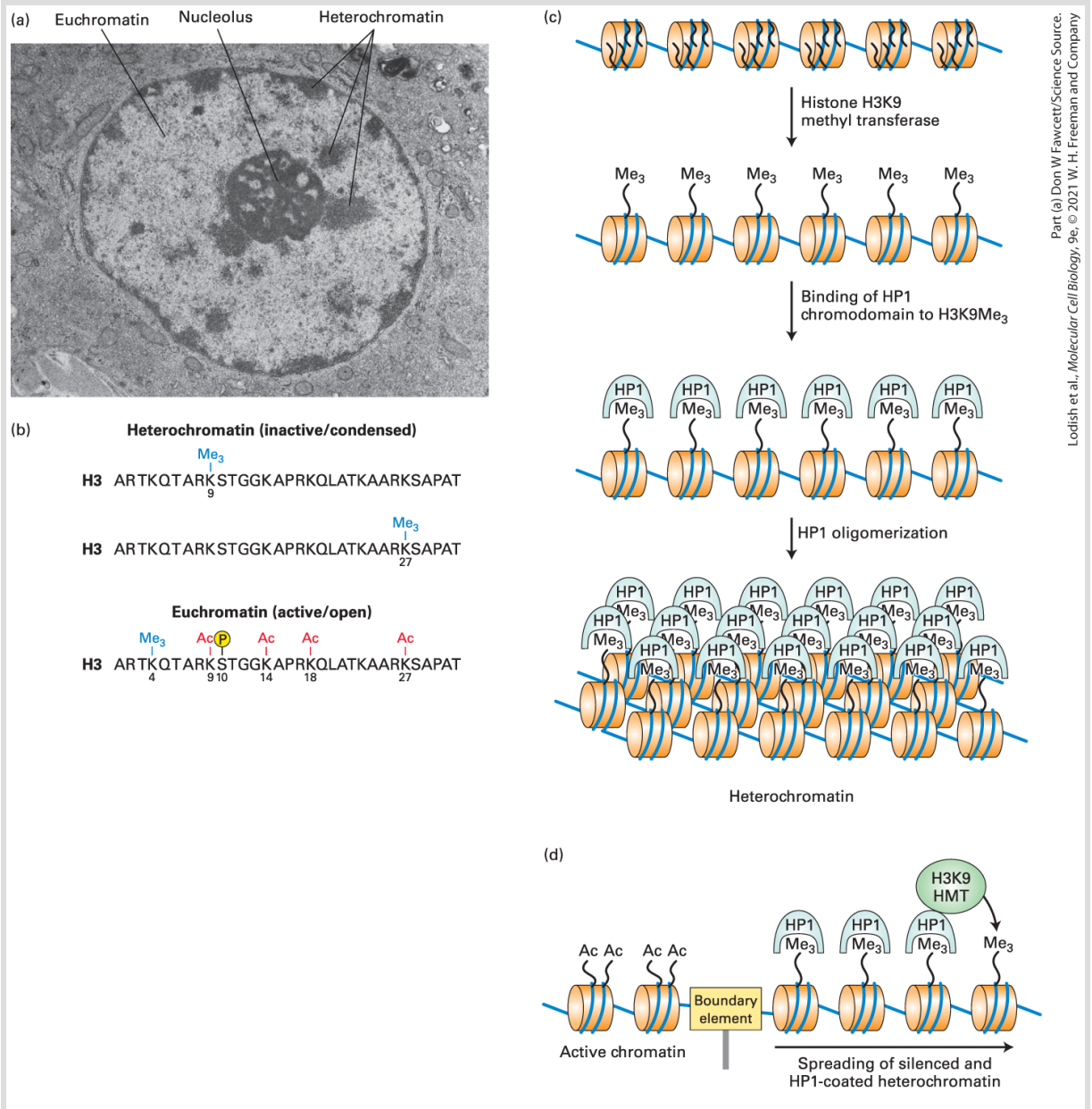
acetylated lysines from histone tails, as discussed further in Chapter 8.

Other Specific Histone-Tail Post-Translational Modifications

Histone tails in chromatin undergo a variety of other covalent modifications at specific amino acids in addition to lysine acetylation (see Figure 7-26b). Lysine ε-amino groups can be methylated, a process that prevents acetylation, thus maintaining their positive charge. Moreover, the nitrogen of lysine ε-amino groups can be methylated once, twice, or three times. Arginine side chains can also be methylated. The oxygen in hydroxyl groups of serine and threonine side chains can be reversibly phosphorylated, introducing two negative charges. Each of these post-translational modifications contributes to the binding of chromatin-associated proteins that participate in the control of chromatin condensation and the ability of DNA and RNA polymerases to replicate or transcribe the associated DNA. The last of the most common post-translational modifications of histone tails is monoubiquitinylation. In this process, a single 76–amino acid ubiquitin molecule can be reversibly added to a lysine in the C-terminal tails of H2A and H2B. Recall that addition of multiple, linked ubiquitin molecules to a protein can mark it for degradation by the proteasome (see Figure 3-32). In this case, however, the addition of a single ubiquitin molecule

does not greatly affect the stability of a histone, although it does influence chromatin compaction.

Heterochromatin does not fully decondense following mitosis, remaining in a compacted state during interphase and usually associating with the nuclear envelope, nucleoli, and additional distinct foci (Figure 7-28a). Heterochromatin includes centromeres and telomeres of chromosomes as well as transcriptionally inactive genes. Heterochromatin usually contains histone H3 modified by methylation of lysine 9. In contrast, euchromatin generally contains histone H3 extensively acetylated on lysines 9, 18, and 27, especially at promoters and enhancers where there is also phosphorylation of H3 serine 10 (Figures 7-26b and 7-28b). Methylation of H3 lysine 4 is observed at transcription start sites of most metazoan genes.



Part (a) Don W Fawcett/Science Source.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 7-28 Histone-tail post-translational modifications in heterochromatin versus euchromatin. (a) In this electron micrograph of a bone marrow stem cell, the dark-staining areas in the nucleus outside the nucleolus are heterochromatin. The light-staining, whitish areas are euchromatin. (b) The modifications of histone N-terminal tails in heterochromatin and euchromatin differ, as illustrated here for histone H3. Note in particular that histone tails are generally much more extensively acetylated in euchromatin than in heterochromatin. Heterochromatin is much more condensed (thus less accessible to proteins) and is much less transcriptionally active than is euchromatin. (c) HP1 contributes to the condensation of heterochromatin by binding to histone H3 N-terminal tails

trimethylated at lysine 9, then associating with other histone-bound HP1 molecules. (d) Heterochromatin condensation can spread along a chromosome because HP1 binds the histone methyltransferase (HMT) that methylates lysine 9 of histone H3. This creates a binding site for HP1 on the neighboring nucleosome. The spreading process continues until a boundary element is encountered. See G. Thiel et al., 2004, *Eur. J. Biochem.* **271**:2855, and A. J. Bannister et al., 2001, *Nature* **410**:120.

[Part (b) Data from T. Jenuwein and C. D. Allis, 2001, *Science* **293**:1074.]

Reading the Histone Code

The histone code of modified amino acids is read by proteins that bind to the modified histone tails and promote condensation or decondensation of chromatin. For example, eukaryotes express a number of proteins containing a domain called a *chromodomain* that binds to histone tails when they are methylated at specific lysines. One example of such a protein is *heterochromatin protein 1 (HP1*). The chromodomain of HP1 binds the H3 N-terminal tail only when it is di- or trimethylated at lysine 9 (see Figure 7-28b). HP1 also contains a second domain called a *chromoshadow domain* because it is frequently found in proteins that contain a chromodomain. The chromoshadow domain binds to chromoshadow domains in other proteins. Consequently, chromatin containing H3 di- and trimethylated at lysine 9 ($H3K9Me_{2/3}$) is assembled into a condensed chromatin structure by HP1 (see Figure 7-28c, *bottom*), although the structure of this chromatin is as yet not well understood.

In addition to binding to itself, the chromoshadow domain of HP1 also binds the enzyme that methylates H3 lysine 9, an H3K9 *histone methyl transferase (HMT)* (see Figure 7-28d). As a consequence, nucleosomes adjacent to a region of HP1-containing heterochromatin become methylated at lysine 9 (see Figure 7-28d). This methylation creates a binding site for another HP1 that can bind the H3K9 HMT; so in this way, the heterochromatin structure spreads along the chromosome until a *boundary element* is encountered that blocks further spreading. Boundary elements are regions in chromatin where several nonhistone proteins bind to DNA, possibly blocking histone methylation on the other side of the boundary.

Epigenetic Memory

The model of heterochromatin formation in Figure 7-28d provides an explanation for how heterochromatic regions of a chromosome are reestablished following DNA replication during S phase of the cell cycle. When DNA in heterochromatin is replicated, the histone octamers that are di- or trimethylated at H3 lysine 9 are distributed to both daughter chromosomes along with an equal number of newly assembled histone octamers. The H3K9 HMT associated with the H3K9 di- and trimethylated nucleosomes methylates lysine 9 of the newly assembled, neighboring nucleosomes, regenerating the heterochromatin covering the same DNA sequence in both daughter chromosomes. Consequently, heterochromatin is marked with an

epigenetic code that maintains the repression of associated genes in replicated daughter cells. It is referred to as "epi-" genetic because it is not determined by DNA sequence, but rather by the post-translational modifications on the histone tails of the parental chromatin.

There are other protein domains that associate with histone-tail modifications typical of euchromatin. Several proteins involved in stimulating gene transcription, such as the largest subunit of TFIID, contain protein domains of ~62 amino acids called *bromodomains* (see Chapter 8). A bromodomain binds to acetylated histone tails and therefore is associated with transcriptionally active chromatin. TFIID and other bromodomain-containing proteins also have histone acetylase activity, which helps to maintain the chromatin in a hyperacetylated state that is conducive to transcription.

Thus we see that an epigenetic code associated with histone post-translational modifications helps to maintain the transcriptional activity of genes in euchromatin and the repression of genes in heterochromatin through successive cell divisions. These epigenetic codes for heterochromatin and euchromatin are essential during early embryonic development when they maintain the patterns of gene expression established in different cell types as these differentiated cells undergo multiple rounds of cell division. Abnormal alterations in these epigenetic codes have been found to

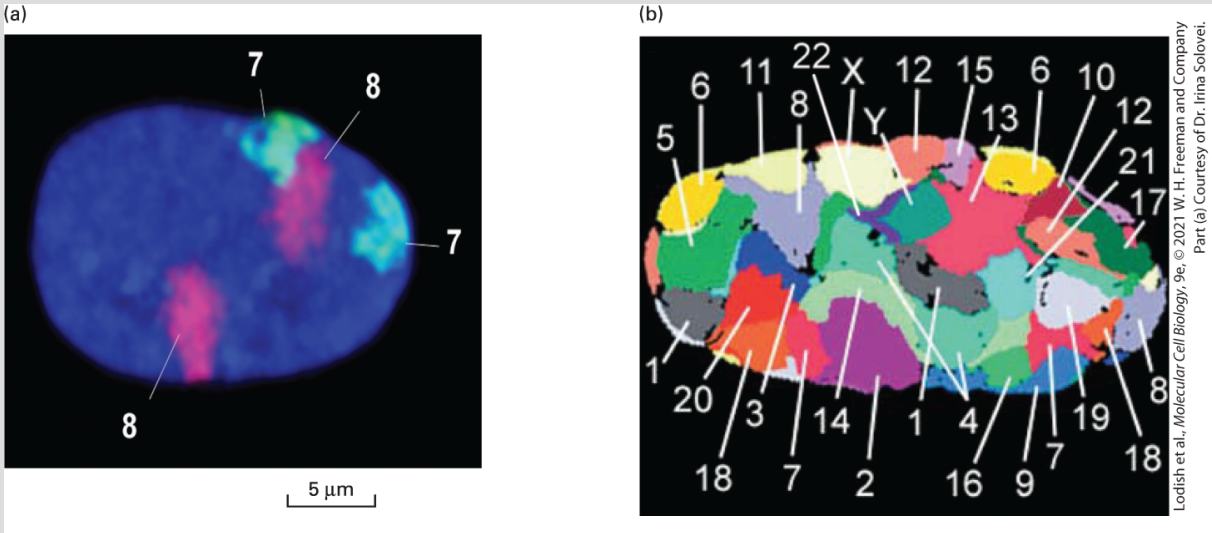
contribute to the pathogenic replication and behavior of cancer cells (see Chapter 25).

In summary, multiple types of covalent modifications to histone tails influence chromatin structure by altering nucleosome–nucleosome interactions and interactions with additional proteins that participate in or regulate processes such as transcription and DNA replication. The mechanisms and molecular processes governing chromatin modifications that regulate transcription are discussed in greater detail in the next chapter.

Interphase Chromosome Territories

Individual interphase chromosomes are less condensed than metaphase chromosomes (see Figure 7-23e, f). Nonetheless, interphase chromatin is not spread throughout the entire nucleus. In situ hybridization of interphase nuclei with chromosome-specific fluorescently labeled probes (called *chromosome paint probes*) has shown that the DNA of each interphase chromosome is localized within a region of the nucleus (Figure 7-29a). These localized regions are called *chromosome territories*. Note in Figure 7-29b how little overlap there is between chromosomes in an interphase nucleus. The precise position of each chromosome is not the same in different cells, although large chromosomes tend to lie at the periphery of the nucleus and small chromosomes toward the center. Also, repeats of rRNA transcription units on chromosomes 13, 14,

15, 21, and 22, known as *nucleolar organizers*, associate with nucleoli found near the center of the nucleus.



EXPERIMENTAL FIGURE 7-29 During interphase, human chromosomes remain in non-overlapping territories in the nucleus. (a) Fixed interphase human fibroblasts were hybridized in situ to fluorescently labeled probes specific for sequences along the full length of human chromosomes 7 (cyan) and 8 (purple). DNA is stained blue with DAPI. In the diploid cell, each of the two chromosome 7s and two chromosome 8s is restricted to a territory or domain within the nucleus, rather than stretching throughout the entire nucleus. (b) This image from a fixed interphase fibroblast from a human male was made with a method similar to that used in (a), except that chromosome paint probes specific for each chromosome were hybridized to the cell to reveal the locations of nearly all the chromosomes. Some of the chromosomes are not observed in this confocal slice through the nucleus.

[Part (b) From A. Bolzer et al., 2005, "Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes," *PLoS Biol.* **3**:826.]

X-Chromosome Inactivation in Mammalian Females

An important example of epigenetic repression of gene expression is the random inactivation and condensation of one X chromosome in female mammalian cells. Each female mammal has two X chromosomes, one contributed by the egg (X_m) and one contributed by the sperm (X_p). Early during embryonic development, random inactivation of either X_m or X_p occurs in each somatic cell. From that point forward during embryonic development and in adults, every daughter cell generated by cell division maintains the same inactive X chromosome as its parent cell. As a result, the adult female is a mosaic of clones, about half expressing the genes from X_m and the rest expressing the genes from X_p. Inactivation of one X chromosome is the mechanism of *dosage compensation* in female mammals, a process that ensures that cells of females express proteins encoded on the X chromosome at the same levels as the cells of males, which have only one X chromosome.

Histones associated with an inactive X chromosome have post-translational modifications characteristic of other regions of heterochromatin: hypoacetylation of lysines, di- and trimethylation of histone H3 lysine 9, and H3 lysine 27, and a lack of methylation at histone H3 lysine 4 (see Figure 7-28b). X-chromosome inactivation is controlled by the X-inactivation center, a complex locus containing two transcription units on the X chromosome that determines which of the two X chromosomes will be inactivated in each cell. The X-inactivation center includes the *XIST* gene, which encodes a remarkable, *long nonprotein-coding RNA* that associates with

multiple regions of the X chromosome from which it was transcribed, maintaining silencing of that chromosome.

Although the mechanism of X-chromosome inactivation is not fully understood, it involves several processes, including the action of *Polycomb* protein complexes, which are discussed further in Chapter 8. One subunit of the Polycomb complex contains a chromodomain that binds to histone H3 tails when they are trimethylated at lysine 27. The Polycomb complex also contains a histone methyl transferase specific for H3 lysine 27. This finding helps to explain how the X-inactivation process spreads along large regions of the X chromosome and how it is maintained through DNA replication, apparently in a manner similar to heterochromatization by the binding of HP1 to histone H3 tails methylated at lysine 9 (see Figure 7-28c, d).

The mechanism of X-chromosome inactivation is another example of an epigenetic process; that is, a process that affects the expression of specific genes and is inherited by daughter cells, but is not the result of a change in DNA sequence. In other words, the activity of genes on the X chromosome in female mammals is controlled by chromatin structure, rather than by the nucleotide sequence of the underlying DNA. And the inactivated X chromosome (either X_m or X_p) is maintained as the inactive chromosome in the progeny of all future cell divisions because the

histones are modified in a specific, repressing manner that is faithfully inherited through each cell division.

Topological Domains Within Chromosome Territories

It is now possible to gain information about the three-dimensional spatial organization of chromatin within nuclei of interphase cells using a group of related methods referred to as *chromosome conformation capture* methods. These methods utilize massively parallel DNA sequencing and the ability to sequence tens of millions of 50–100-bp DNA fragments (see Chapter 6). The general strategy of chromosome conformation capture is illustrated in Figure 7-30, where protein-dependent associations between distant regions of chromatin are represented in (a), and regions of chromatin not specifically associated with another region are shown in (b). Step **1**: Intact cells are treated with chemical cross-linkers such as formaldehyde that diffuse through cell membranes and covalently cross-link protein to protein and protein to DNA. Step **2**: The cross-linked chromatin is isolated and either digested with a restriction enzyme or subjected to intense sonication to mechanically shear the DNA into fragments of 200–600 bp. Step **3**: Short oligonucleotide linkers are then ligated to the ends of the DNA fragments. These linkers contain a cytosine linked to a biotin at the C-5 position (see Figure 2-17). Step **4**:After removal of excess linker, the preparation is diluted considerably and subjected to treatment with DNA ligase.

Because the preparation is diluted, ligations occur preferentially between the ends of fragments held in close proximity by cross-links. Step **5**: After ligation, the protein-DNA cross-linking reaction is reversed, the protein is digested with proteases, and the DNA is isolated and further sonicated. Ligated fragments are separated from other fragments using streptavidin, which binds the biotin in the linkers added in step **3**.

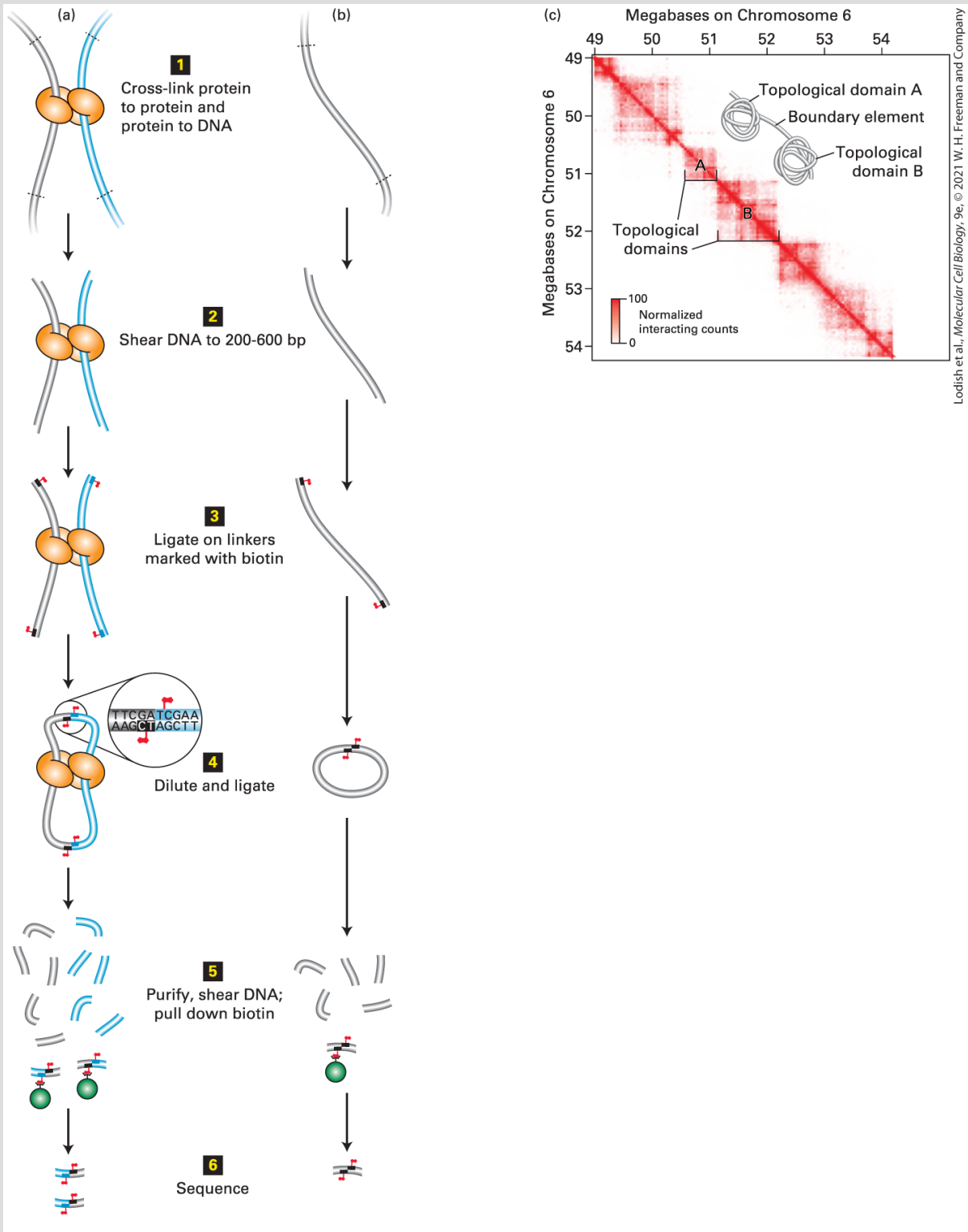


FIGURE 7-30 Chromosome conformation capture. (a, b) Strategy of chromatin conformation capture methods. See text for explanation. Gray and blue tubes represent regions of DNA separated by >10 kb in the genome sequence. Biotin is represented by red

flags. Data from E. Lieberman-Aiden, 2009, *Science* **326**:289. (c) Heat map of chromosome conformation capture data for a region of chromosome 6 in mouse embryonic stem cells. The sequence from 49 to 54 Mb from the left end of chromosome 6 is represented on both axes. A value of 100 (dark red) indicates that a sequence anywhere within the 10-kb region on the *x* axis was found ligated to a sequence from anywhere in the 10-kb region on the *y* axis 100 times. Since the probability that two ends generated by sonication will be ligated together is higher for ends that are close together than for ends that are far apart, the intensity of the red color in any pixel indicates the relative proximity of the sequences in the two 10-kb intervals in the nuclei at the time of cross-linking. Inset shows a model of chromatin folding that is consistent with these results.

[Part (c) Data from J. R. Dixon, 2012, *Nature* **485**:376.]

The purified, ligated fragments are then sequenced (step **6**). Ligation points are marked by the duplicated sequence of the oligonucleotide linker (step **4**). Sequences on each side of the ligation point are mapped to the genome. In cases in which the two ends of a single fragment were linked together (Figure 7-30b), the sequences on either side of the oligonucleotide linker sequence will map within a few hundred bases of each other on the genome, since the ligated fragments were only a few hundred bases long. But in cases in which fragments that are distant in the genomic sequence were ligated together because they were cross-linked to proteins holding them together (Figure 7-30a), the sequences will map far apart. The observation of distant (>10 kb) sequences repeatedly ligated to each other implies that their respective regions of chromatin were associated with each other in vivo.

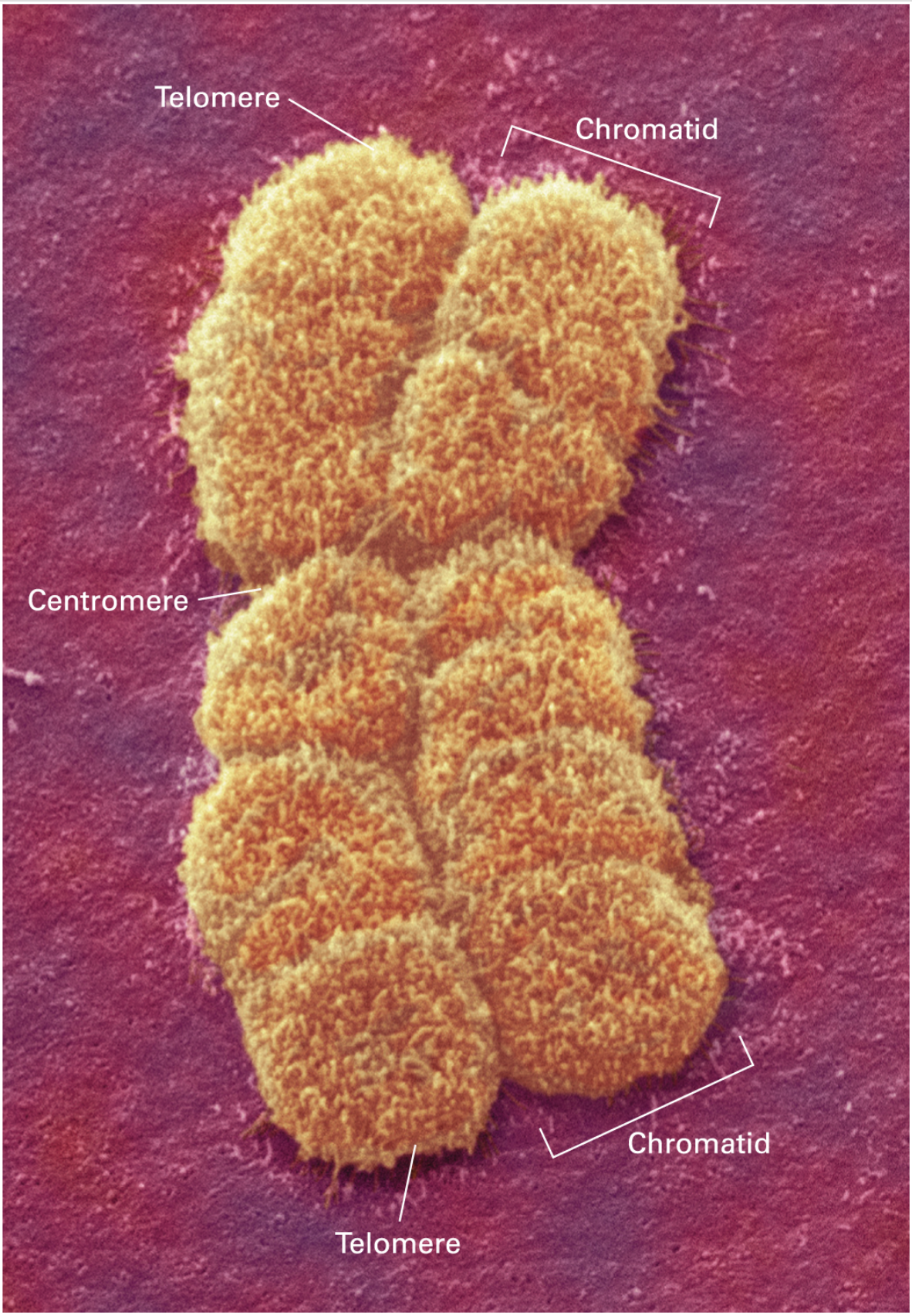
The data from a chromatin conformation capture assay can be plotted on a two-dimensional heat map (Figure 7-30c). In this plot, the sequence of the same portion of the genome is plotted on both the *x* and *y* axes, with each pixel equivalent to 10 kb. The pixel at coordinates *x, y* is colored red where a sequence at coordinate *x* was ligated to a sequence at coordinate *y*. The intensity of the red color is proportional to the numbers of ligation events observed that linked a sequence in the 10-kb interval *x* with a sequence in the 10-kb interval *y*.

Figure 7-30c shows the plot generated for a roughly 5.5-Mb region of chromosome 6 in mouse embryonic stem cells. It is immediately apparent that the genome is divided into regions called *topological domains*, in which a chromosomal region is far more likely to be ligated to another sequence within the same topological domain than it is to be ligated to a sequence in another topological domain. These topological domains are on the order of 200 kb to 1.5 Mb in length, with a median size of 880 kb. For example, sequences in the interval of chromosome 6 between 50.9 Mb and 51.3 Mb (Figure 7-30c, topological domain A) are much more likely to be ligated to each other than to sequences in the interval from 51.3 Mb to 52.2 Mb (topological domain B) or to sequences from any of the other topological domains. In situ hybridization studies showed that sequences within a topological domain lie much closer to each other in the fixed cell nucleus than to sequences the same distance away in base pairs, but in a neighboring topological domain. These results

have been interpreted to indicate that the chromatin fiber is folded into topological domains, as represented in the inset of Figure 7-30c. The topological domains are separated by shorter regions of chromatin, called *boundary elements*, that do not interact with distant regions of chromatin. Boundary elements such as this were discussed earlier where they were observed to block spreading of histone H3 lysine 9 tri-methylation and HP1-binding along chromatin tri-methylated on lysine 9 (see Figure 7-28d). Topological domains are long enough to contain several average-sized genes. One area of current research is exploring what protein-DNA interactions might be responsible for establishing boundary elements between topological domains. As we will see in Chapter 8, related chromosome conformation capture techniques have provided strong evidence that proteins bound to enhancers interact with proteins bound to promoters many kilobases away.

Metaphase Chromosome Structure

Condensation of chromosomes during prophase (see Figure 18-38) appears to involve the formation of many more loops of chromatin (as diagrammed for heterochromatin (see Figure 7-25c), so that the length of each loop is greatly reduced compared with chromatin loops in interphase cells. As a result, chromosomes condense into structures of much greater width than interphase chromosomes and decrease in length several-fold, generating the rod-shaped condensed chromosomes observed during metaphase (Figure 7-31).



Andrew Syred/Science Source.

FIGURE 7-31 Typical metaphase chromosome. As seen in this scanning electron micrograph, the chromosome has replicated and comprises two chromatids, each containing one of two identical DNA molecules. The centromere, where the chromatids are attached at a constriction, is required for their separation late in mitosis. Special telomere sequences at the ends function in preventing chromosome shortening.

Experiments with frog egg extracts have shown that a protein complex called *condensin*, composed of SMC subunits (see Figure 7-25 and Chapter 19), contributes to chromosome condensation using energy from ATP hydrolysis. Ultimately, the full lengths of the two associated daughter chromosomes generated by DNA replication during the previous S phase of the cell cycle (see Figure 1-22) condense into bar-shaped structures (chromatids) that in most eukaryotes are linked at the central constriction called the *centromere* (Figure 7-31).

Additional Nonhistone Proteins Regulate Transcription and Replication

The total mass of the histones associated with DNA in chromatin is about equal to that of the DNA. Interphase chromatin and metaphase chromosomes also contain many other proteins in addition to histones. For instance, thousands of different **transcription factors** are associated with interphase chromatin. The

structure and function of these proteins, which regulate transcription, are examined in Chapter 8. Other low-abundance nonhistone proteins associated with chromatin regulate DNA replication during the eukaryotic cell cycle (see Chapter 19).

A few other nonhistone DNA-binding proteins are present in much larger amounts than the transcription or replication factors. Some of these proteins exhibit high mobility during electrophoretic separation and thus have been designated *HMG (high-mobility group) proteins*. When genes encoding the most abundant HMG proteins are deleted from yeast cells, normal transcription is disturbed in most genes examined. Some HMG proteins have been found to assist in the cooperative binding of several transcription factors to specific DNA sequences that are close to each other, stabilizing multiprotein complexes that regulate transcription of a neighboring gene, as discussed in Chapter 8.

KEY CONCEPTS OF SECTION 7.4

Structural Organization of Eukaryotic Chromosomes and Chromatin

- Each eukaryotic chromosome contains a single extremely long DNA molecule associated with about an equal mass of proteins called histones, as well as less abundant amounts of thousands of different chromatin-associated proteins, forming a DNA-protein complex called *chromatin*.
- The key structural element of chromatin is the nucleosome formed from two copies each of four core histones (H2A, H2B, H3, and H4) that form a globular protein domain around which 147–149 base pairs of DNA are tightly wrapped (see Figure 7-20).

Nucleosomes are strung together by a variable length (~50–150 base pairs) of *linker DNA* between them (see Figure 7-22).

- Using a recently developed method called *ChromEMT* which stains DNA with the electron opaque compound OsO_4, chromatin in fixed cells is observed in the electron microscope as a disordered 5- to 24-nm-diameter granular chain of nucleosomes and linker DNA that is folded into different concentration densities in different regions of the nucleus (see Figures 7-23 and 7-24).
- The chromatin in transcriptionally active regions of DNA, called euchromatin, exists in a more extended form than the chromatin in regions called *heterochromatin* where nucleosomes are more densely packed. Most genes in heterochromatin are transcriptionally repressed. In heterochromatin the chromatin chain folds back on itself at short intervals (see Figure 7-23e, f and Figure 7-24b, c), and nucleosomes associate with each other through many alternative surfaces (Figure 7-23b–d) to a much greater extent than in interphase chromosomes. This results in a large increase in nucleosome density in heterochromatin compared to euchromatin (see Figure 7-23e, f).
- Flexible, intrinsically disordered N-terminal *tails* of 16–39 amino acids extend from the globular region of the histone octamer core (see Figure 7-26a). Histones H2A and H2B also have shorter, intrinsically disordered regions at their C-termini. Amino acid side chains in these histone tails are reversibly modified by acetylation, methylation, phosphorylation, and monoubiquitinylation (see Figure 7-26b), as well as other, less abundant post-translational modifications. These modifications influence chromatin structure by regulating the binding of histone tails to other, less abundant chromatin-associated proteins that control chromatin compaction, transcription, and DNA replication.
- Proteins involved in transcription, replication, and repair, and enzymes such as DNase I, can more easily access the DNA in euchromatin with hyperacetylated histone tails than in chromatin with hypoacetylated histone tails, as in heterochromatin.
- An abundant protein associated with heterochromatin, heterochromatin protein 1 (HP1) uses a chromodomain to bind to histone H3 trimethylated at lysine 9. The chromoshadow domain of HP1 associates with itself and with the histone methyl transferase that methylates H3 lysine 9. These interactions cause condensation of the chromatin fiber and spreading of the heterochromatic structure along the chromosome until a boundary element is encountered (see Figure 7-28c, d).
- One X chromosome in nearly every cell of mammalian females consists of highly condensed heterochromatin, resulting in repression of nearly all genes on that chromosome. This X chromosome inactivation is the dosage compensation mechanism used in mammals so that genes on the X chromosome are expressed at the same level in both males and females.

- In interphase cells, chromosomes are localized to largely non-overlapping territories in the nucleus (see Figure 7-29b).
- The results of chromosome conformation capture experiments indicate that chromatin is organized into topological domains of about 200 kb to 1.5 Mb separated by boundary elements (see Figure 7-30c). Chromatin within a topological domain is more densely packed than the chromatin between topological domains in *boundary elements.* As a result, the chromatin in a topological domain is much more likely to interact with another region of chromatin in the same topological domain than with a region of chromatin in another topological domain (see Figure 7-30c).
- During mitosis, chromosomes condense greatly, decreasing their lengths and increasing their diameters to generate metaphase chromosomes visible by light microscopy. Recent observations using ChromEMT indicate that chromatin fibers in condensed mitotic chromosomes fold back on themselves at shorter intervals than in euchromatin, and that nucleosomes associate with each other with multiple possible orientations and surfaces much more frequently in mitotic chromosomes compared to interphase chromosomes (see Figure 7-23b–d) resulting in the much higher density of nucleosomes in mitotic chromosomes.

7.5 Morphology and Functional Elements of Eukaryotic Chromosomes

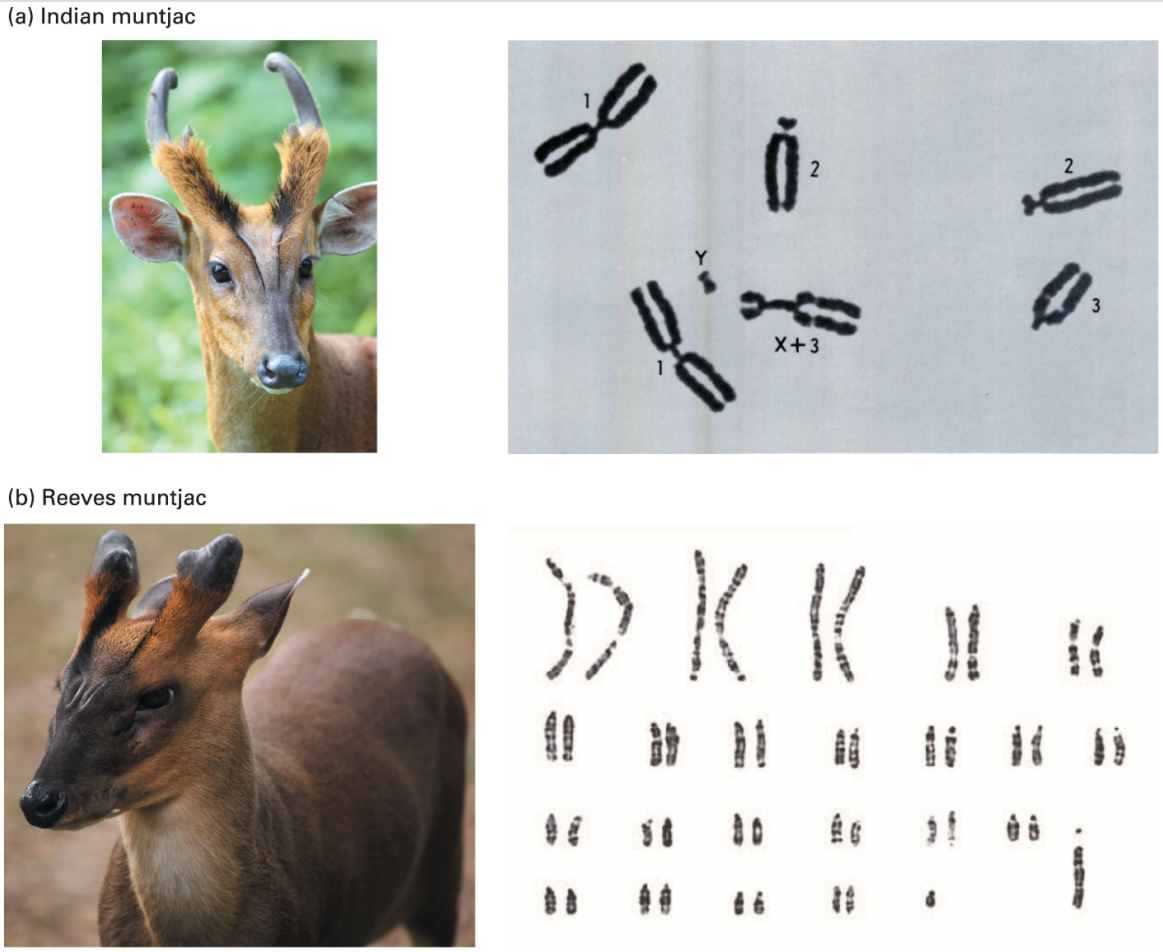
Having examined the detailed structural organization of chromosomes in the previous section, we now view them from a more global perspective. Early microscopic observations on the number and size of chromosomes and their staining patterns led to the discovery of many important general characteristics of chromosome structure. Researchers subsequently identified specific regions of chromosomes that are critical to their replication and segregation into daughter cells during cell division. In this section, we discuss these functional elements of chromosomes and consider how chromosomes evolved through rare rearrangements of ancestral chromosomes.

Chromosome Number, Size, and Shape at Metaphase Are Species-Specific

In interphase cells, as noted previously, chromosome territories can be visualized with fluorescently labeled hybridization probes (see Figure 7-29). Because chromosomes condense during mitosis and meiosis (see Figure 7-23e, f) and become visible in the light microscope (see Figure 7-31), almost all cytogenetic work (i.e.,

studies of chromosome morphology) has been done with condensed metaphase chromosomes obtained from dividing cells — either somatic cells in mitosis or dividing gametes during meiosis. At the time of mitosis and meiosis, cells have already progressed through the S phase of the cell cycle and have replicated their DNA (see Chapter 19). Consequently, the chromosomes that become visible during metaphase are *duplicated* structures. Each metaphase chromosome consists of two sister **chromatids**, which are linked at a constricted region, the centromere (see Figure 7-31).

The number, sizes, and shapes of the metaphase chromosomes constitute the **karyotype**, which is distinctive for each species. While somatic cells from one species generally have the same karyotype, species that appear quite similar can have very different karyotypes. Distinct karyotypes among closely related species indicate that similar genetic potential can be organized on chromosomes in very different ways. For example, two species of small deer — the Indian muntjac and Reeves muntjac — contain about the same total amount of genomic DNA. The Indian muntjac contains the smallest number of chromosomes of any mammal, only three pairs of autosomes; one sex chromosome is physically separate, but the other is joined to the end of one autosome (Figure 7-32a). In the Reeves muntjac, however, this DNA is organized into 22 pairs of homologous **autosomes** and two physically separate sex chromosomes (Figure 7-32b).



Part (a) Feathercollector/Shutterstock. (b) Vladimir Wrangel/Shutterstock. Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 7-32 Karyotypes of the Indian muntjac and the Reeves muntjac. These two species of small deer are quite similar, but do not interbreed. Despite the difference in the number of chromosomes in these animals, the two genomes contain about the same total amount of DNA. (The chromosomes of males of both species are shown at the same magnification.)

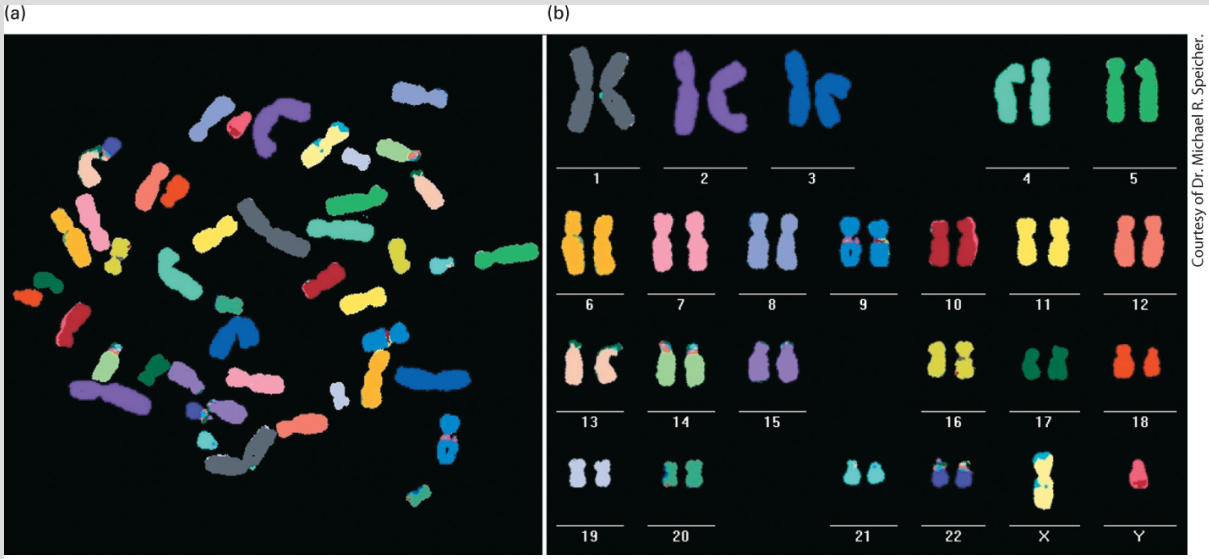
Republished with permission by SAGE Publications, from A. V. Carrano et al., 1976, "Purification of the Chromosomes of the Indian Muntjac by Flow Sorting," *J. Histochem. Cytochem*. **24**(1), 348–354; https://doi.org/10.1177/24.1.1254929; permission conveyed through Copyright Clearance Center, Inc. Republished with permission by Karger Publishers, from J, Fu B, Nie W, Wang J, Graphodatsky A, S, Yang F: New insights into the karyotypic relationships of Chinese muntjac (*Muntiacus reevesi*), forest musk deer (*Moschus berezovskii*) and gayal (*Bos frontalis*). *Cytogenet Genome Res* 2005; **108**:310–316. doi: 10.1159/000081520]

During Metaphase, Chromosomes Can Be Distinguished by Banding Patterns and Chromosome Painting

Certain dyes selectively stain some regions of metaphase chromosomes more intensely than other regions, producing characteristic banding patterns that are specific for individual chromosomes. The regularity of chromosome bands provides useful visible landmarks along the length of each chromosome and can help to distinguish chromosomes of similar size and shape.

Today the method of chromosome painting greatly simplifies the identification and differentiation of individual chromosomes within a karyotype, many of which have similar sizes and shapes. This technique, a variation of **fluorescence in situ hybridization (FISH)**, makes use of probes specific for sites scattered along the length of each chromosome. The probes are labeled with several different fluorescent dyes with distinct excitation and emission wavelengths. Probes specific for each chromosome are labeled with a predetermined fraction of each of the dyes. After the probes are hybridized to chromosomes and the excess removed, the sample is observed with a fluorescence microscope in which a detector determines the fraction of each dye present at each fluorescing position in the microscopic field. This information is conveyed to a computer, and a special program assigns a false-color image to each type of chromosome (Figure 7-33a). Computer graphics allows the

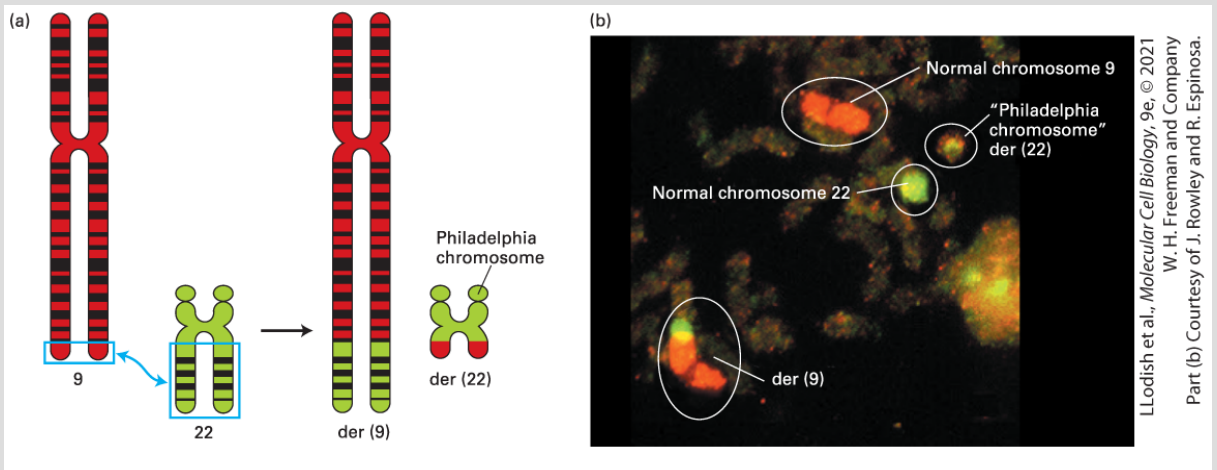
two homologs of each chromosome to be placed next to each other and numbered according to their decreasing size. Such an image clearly displays the cell's karyotype (Figure 7-33b).



EXPERIMENTAL FIGURE 7-33 Human chromosomes are readily identified by chromosome painting. (a) Image of human chromosomes from a male cell in mitosis made by fluorescence in situ hybridization (FISH) using chromosome paint probes. (b) Alignment of these painted chromosomes by computer graphics to reveal the normal human male karyotype.

Chromosome painting is a powerful method for detecting an abnormal number of chromosomes, such as chromosome 21 trisomy in patients with Down syndrome, or chromosomal translocations that occur in rare individuals and in cancer cells (Figure 7-34). The use of probes with different ratios of fluorescent dyes that hybridize to distinct positions along each normal human chromosome allows finer structural analysis of the chromosomes that can more readily reveal deletions or duplications of

chromosomal regions. The chapter-opening figure illustrates the use of such *multicolor FISH* in analysis of the karyotype of a normal human female.



EXPERIMENTAL FIGURE 7-34 Chromosomal translocations can be analyzed using chromosome painting. Characteristic chromosomal translocations are associated with certain genetic disorders and specific types of cancers. For example, in nearly all patients with chronic myelogenous leukemia, the leukemic cells contain the Philadelphia chromosome, a shortened chromosome 22 [der (22)], and an abnormally long chromosome 9 [der (9)] ("der" stands for derivative). These forms result from a translocation between normal chromosomes 9 and 22. This translocation can be detected (a) by classical banding analysis or (b) by chromosome painting.

Chromosome Painting and DNA Sequencing Reveal the Evolution of Chromosomes

Analysis of chromosomes from different species has provided considerable insight into how chromosomes evolved. For example,

hybridization of chromosome paint probes for chromosome 16 of the tree shrew (*Tupaia belangeri*) to tree shrew metaphase chromosomes revealed the two copies of chromosome 16, as expected (Figure 7-35a). However, when the same chromosome paint probes were hybridized to human metaphase chromosomes, most of the probes hybridized to the long arm of chromosome 10 (Figure 7-35b). Further, when multiple probes for the long arm of human chromosome 10 with different fluorescent dye labels were hybridized to tree shrew metaphase chromosomes, these probes bound to sequences along tree shrew chromosome 16 in the same order in which they bind to human chromosome 10.

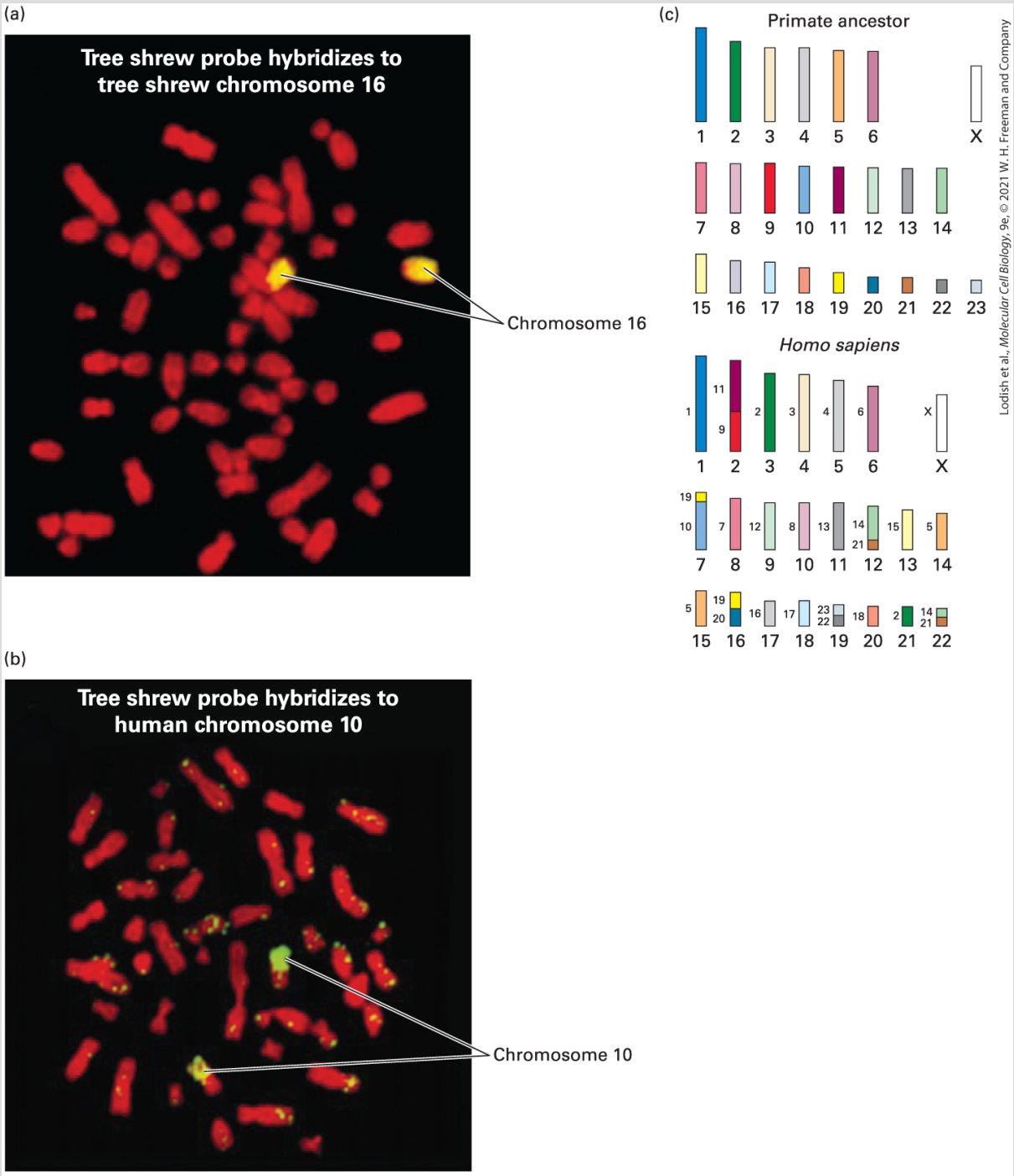


FIGURE 7-35 Evolution of primate chromosomes. (a) Chromosome paint probes (yellow) for chromosome 16 of the tree shrew (*T. belangeri*, distantly related to humans) hybridized to tree shrew metaphase chromosomes (red). (b) The same tree shrew chromosome 16 paint probes (green) hybridized to human metaphase chromosomes (red). (c) Proposed evolution of human chromosomes (*bottom*) from the chromosomes of the common

ancestor of all primates (*top*). The proposed common primate ancestor chromosomes are numbered according to their sizes, with each chromosome represented by a different color. The human chromosomes are also numbered according to their relative sizes and labeled with colors taken from the colors of the proposed common primate ancestor chromosomes from which they were derived. Small numbers to the left of the colored regions of the human chromosomes indicate the number of the ancestral chromosome from which the region was derived.

[Parts (a) and (b) Republished with permission of Springer, from S. Muller et al., 1999, "Defining the Ancestral Karyotype of all Primates by Multidirectional Chromosome Painting Between Tree Shrews, Lemurs and Humans," *Chromosoma* **108**(6):393–400; permission conveyed through Copyright Clearance Center, Inc. Part (c) Data from L. Froenicke, 2005, *Cytogenet. Genome Res.* **108**:122.]

Humans and tree shrews had a common ancestor that lived as recently as 85 million years ago. These results indicate that during the evolutionary process of divergence, a long, continuous DNA sequence on one of the ancestral chromosomes became chromosome 16 in tree shrews but evolved into the long arm of chromosome 10 in humans. The phenomenon of genes occurring in the same order on a chromosome in two different species is referred to as conserved **synteny** (derived from Latin for "on the same ribbon"). The presence of two or more genes in a common chromosomal region in two or more species indicates a conserved syntenic segment.

The relationships between the chromosomes of many primates have also been determined by cross-species application of chromosome paint probes. Using these relationships and higher resolution analyses of syntenic regions by DNA sequencing and other methods,

it has been possible to propose the karyotype of the common ancestor of all primates. This karyotype is based on the minimum number of chromosomal rearrangements necessary to generate the regions of synteny in chromosomes of contemporary primates (Figure 7-35c).

Human chromosomes are thought to have been derived by several different mechanisms from a common primate ancestor with 23 autosomes plus the X and Y sex chromosomes (Figure 7-35c). Some human chromosomes were derived without large-scale rearrangements of chromosome structure in the primate ancestor (e.g., human chromosome 1). Others are thought to have evolved by breakage of an ancestral chromosome into two chromosomes (e.g., human chromosomes 14 and 15 by breakage of ancestral chromosome 5) or, conversely, by fusion of two ancestral chromosomes (e.g., human chromosome 2 by fusion of ancestral chromosomes 9 and 11). Still other human chromosomes appear to have been generated by exchanges of parts of the arms of distinct chromosomes, that is, by reciprocal translocation involving two ancestral chromosomes (e.g., a reciprocal translocation between ancestral chromosomes 14 and 21 generated human chromosomes 12 and 22). Analysis of regions of conserved synteny between the chromosomes of many mammals indicates that chromosomal rearrangements by breakage, fusion, and translocation occurred rarely in mammalian evolution, about once every 5 million years. When such chromosomal rearrangements did occur, they very likely

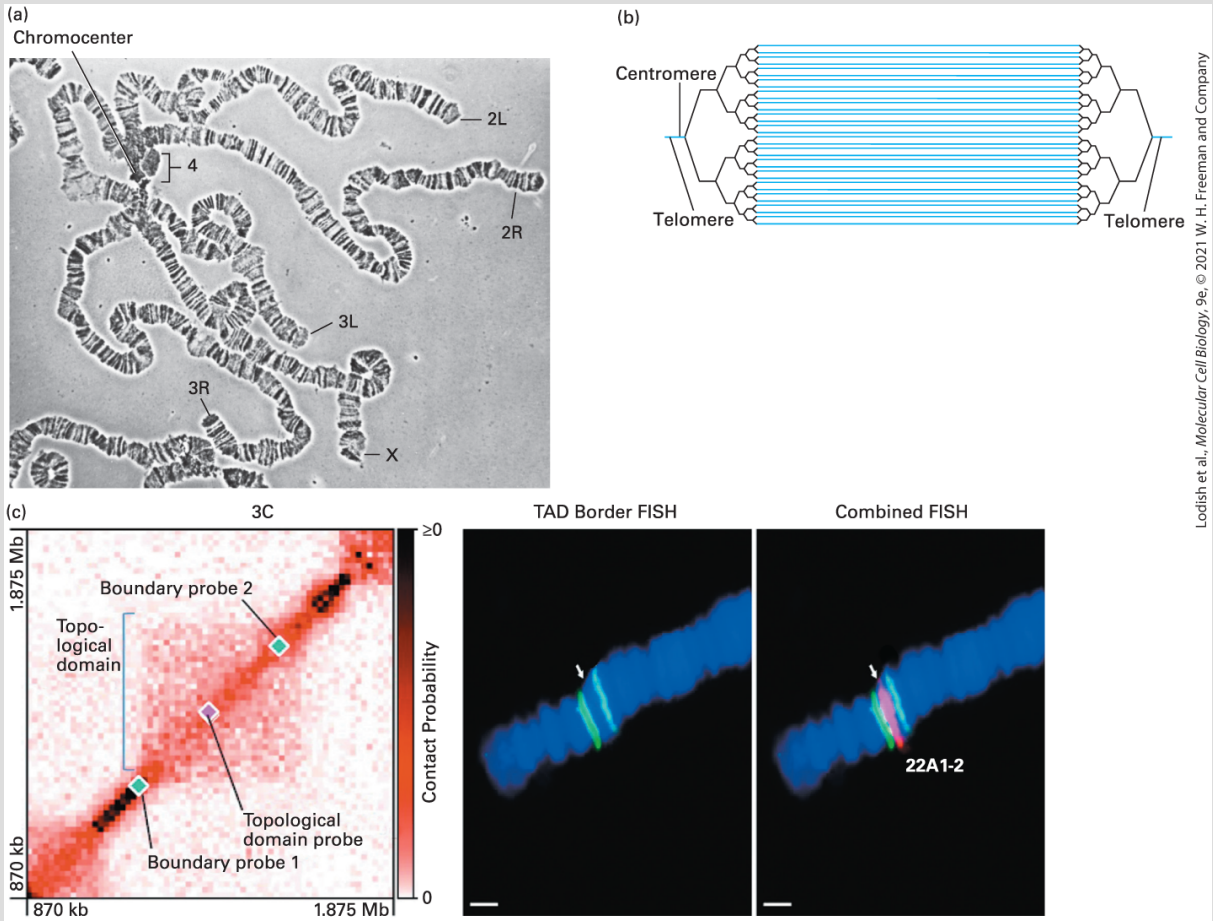
contributed to the evolution of new species that could not interbreed with the species from which they evolved.

Chromosomal rearrangements similar to those inferred for the primate lineage have been inferred for other groups of related organisms, including invertebrate, plant, and fungal lineages. There is excellent agreement between predictions of evolutionary relationships based on analysis of syntenic regions of chromosomes from organisms with related morphology (e.g., among mammals, or among insects with similar body organization) and evolutionary relationships based on the fossil record and on the extent of divergence of DNA sequences for homologous genes. This is a strong argument for the validity of evolution as the process that generated the diversity of contemporary organisms.

Interphase Polytene Chromosomes Arise by DNA Amplification

The larval salivary glands of *Drosophila* species and other dipteran insects contain enlarged interphase chromosomes that are visible in the light microscope. When fixed and stained with a dye that stains DNA, these **polytene chromosomes** are characterized by a large number of reproducible, well-demarcated bands (Figure 7-36a), which have been assigned standardized numbers. The densely staining bands represent regions where the chromatin is more

condensed, and the light interband areas are regions where the chromatin is less condensed. This highly reproducible banding pattern seen in *Drosophila* salivary gland chromosomes provides an extremely powerful method for locating specific DNA sequences along the chromosomes of this species. Chromosomal translocations and inversions are readily detectable by changes in the banding pattern of polytene chromosomes. Also, DNA sequences can be localized along the length of polytene chromosomes by *in situ* hybridization (see Figure 7-36c). Insect polytene chromosomes offer one of the only experimental systems in all of nature in which high resolution microscopic localization studies on decondensed interphase chromosomes are possible.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 7-36 Banding on *Drosophila* polytene salivary gland chromosomes. (a) In this light micrograph of *Drosophila melanogaster* larval salivary gland chromosomes, four chromosomes can be observed (X, 2, 3, and 4), with a total of approximately 5000 distinguishable bands. The banding pattern of light and dark bands results from reproducible differences in the extent of chromatin condensation (from euchromatin-like to heterochromatin-like condensation (see Figure 7-23a) along the length of each of the ~1000 aligned chromosomes in a polytene chromosome. Dark bands are regions of more highly compacted chromatin; light bands, regions of more extended chromatin. The centromeres of all four chromosomes often appear fused at the chromocenter. The telomeres of chromosomes 2 and 3 are labeled (L = left arm; R = right arm), as is the telomere of the X chromosome. (b) The pattern of amplification of chromosome 4 during five replications. Double-stranded DNA is represented by a single line. Telomere and centromere DNA are not amplified. In salivary gland polytene chromosomes, each parental chromosome undergoes about 10 replications $(2^{10} = 1024 \text{ strands})$. (c) Detection of specific DNA sequences by *in situ* hybridization of fluorescently labeled probes.

The left panel shows chromatin conformation (3C) assay data for a region near the telomere of chromosome 2L, with one topological domain at the center of the plot. The position of a fluorescently labeled DNA probe to a region near the center of the topological domain is indicated by a red diamond, and location of probes from the borders of the topological domain, in or near the neighboring boundary elements, are shown in cyan. The right two panels are micrographs of the left end of chromosome 2 in polytene chromosomes hybridized to the fluorescent probes from the borders of the topological domain (cyan) and the fluorescent probe from the middle of the topological domain (red), and stained with DAPI for DNA (blue). The white arrowhead indicates a small interband between bands 22A1-2 and 22A3 in the standardized *Drosophila* polytene chromosome map where the probe from the center of the topological domain hybridizes. Scale bar = 2 μm. See C. D. Laird et al., 1973, *Cold Spring Harbor Symp. Quant. Biol.* **38**:311.

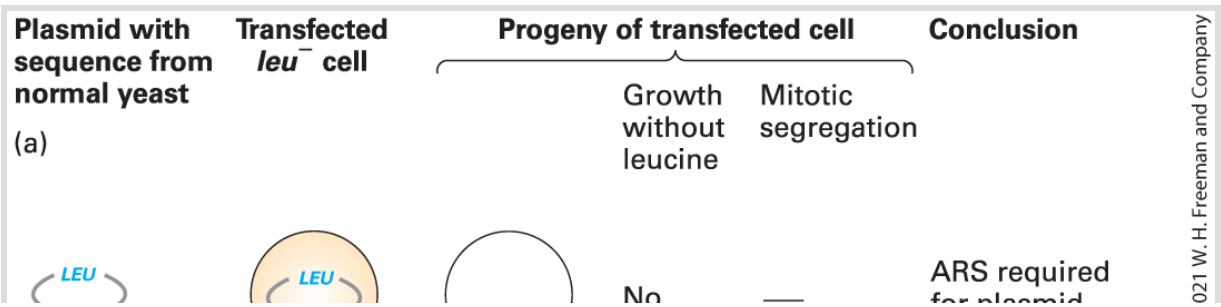
[Part (c) Republished with permission of Elsevier, from K. P. Eagen et al., 2015, "Stable Chromosome Condensation Revealed by Chromosome Conformation Capture," *Cell* **163**(4):934–46; https://doi.org/10.1016/j.cell.2015.10.026; permission conveyed through Copyright Clearance Center, Inc.]

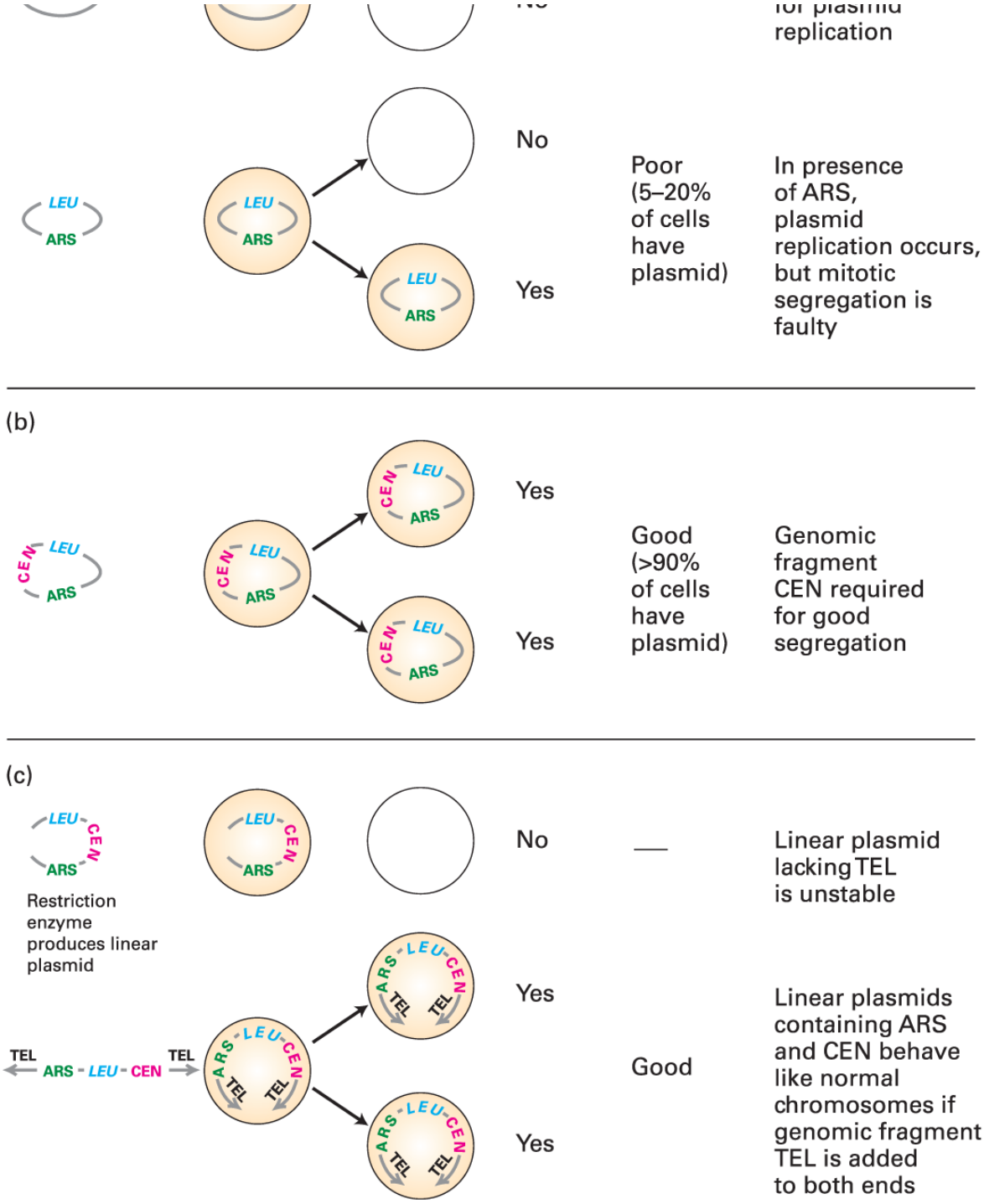
A generalized amplification of DNA gives rise to the polytene chromosomes found in the salivary glands of *Drosophila*. This process, termed *polytenization*, occurs when the DNA repeatedly replicates everywhere except at the telomeres and centromere, but the daughter chromosomes do not separate. The result is an enlarged chromosome composed of many parallel copies of itself, such as 1024 copies resulting from 10 such replications in *Drosophila melanogaster* salivary glands (see Figure 7-36b). The amplification of chromosomal DNA greatly increases gene copy number, presumably to supply sufficient mRNA for protein synthesis in the massive salivary gland cells. *In situ* hybridization studies (see Figure 7-36c) indicate that the dark bands in polytene chromosomes are due to

the alignment in the 1024 chromosomes of topological domains observed by chromatin conformation assays, and that the light bands correspond to alignment of boundary elements.

Three Functional Elements Are Required for Replication and Stable Inheritance of Chromosomes

Although eukaryotic chromosomes differ in length and number among species, cytogenetic studies have shown that they all behave similarly at the time of cell division. Moreover, any eukaryotic chromosome must contain three functional elements in order to replicate and segregate correctly: (1) **replication origins** at which DNA polymerases and other proteins initiate synthesis of DNA (see Figures 5-12 and 5-13); (2) the **centromere**, the constricted region required for proper segregation of daughter chromosomes (see Figure 7-31); and (3) the two ends, or **telomeres**. The yeast transformation studies depicted in Figure 7-37 demonstrated the functions of these three chromosomal elements and established their importance for chromosome function.





Lodish et al., *Molecular Cell Biology*, 9e, © 2

EXPERIMENTAL FIGURE 7-37 Yeast transformation experiments were used to identify the functional chromosomal elements necessary for normal chromosome replication and segregation. In these experiments, plasmids containing the *LEU* gene from normal yeast cells are constructed and introduced into leu^- cells by transfection. If the plasmid is maintained in the leu^- cells, they are transformed to LEU^+ cells by the *LEU* gene on the

plasmid and can form colonies on medium lacking leucine. (a) Sequences that allow autonomous replication (ARS) of a plasmid were identified because their insertion into a plasmid vector containing a cloned *LEU* gene resulted in a high frequency of transformation to LEU^+. However, even plasmids with an ARS exhibit poor segregation during mitosis and therefore do not appear in each of the daughter cells. (b) When randomly broken pieces of yeast DNA are inserted into plasmids containing ARS and *LEU*, some of the subsequently transfected cells produce large colonies, indicating that a high rate of mitotic segregation among their plasmids is facilitating the continuous growth of daughter cells. The DNA recovered from plasmids in these large colonies contains yeast centromere (CEN) sequences. (c) When leu^- yeast cells are transfected with linearized plasmids containing *LEU*, ARS, and CEN, no colonies grow. Addition of telomere (TEL) sequences to the ends of the linear DNA gives the linearized plasmids the ability to replicate as new chromosomes that behave very much like a normal chromosome in both mitosis and meiosis. See A. W. Murray and J. W. Szostak, 1983, *Nature* **305**:189, and L. Clarke and J. Carbon, 1985, *Ann. Rev. Genet.* **19**:29.

As discussed in Chapter 5, replication of DNA begins from sites that are scattered throughout eukaryotic chromosomes. The yeast genome contains many 100-bp sequences, called *autonomously replicating sequences (ARSs)*, that act as replication origins. The observation that insertion of an ARS into a circular plasmid allowed the plasmid to replicate in yeast cells provided the first functional identification of replication origins in eukaryotic DNA (Figure 7-37a).

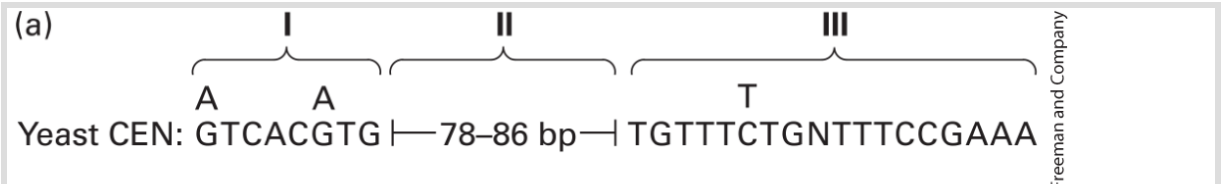
Even though circular ARS-containing plasmids can replicate in yeast cells, only about 5–20 percent of progeny cells contain the plasmid, because mitotic segregation of the plasmids is faulty. However, plasmids that also carry a CEN sequence, derived from the

centromeres of yeast chromosomes, segregate equally, or nearly so, to both mother and daughter cells during mitosis (Figure 7-37b).

If circular plasmids containing an ARS and a CEN sequence are cut once with a restriction enzyme, the resulting linear plasmids do not transform yeast cells generating LEU^+ colonies that grow on medium lacking leucine unless they contain special telomeric (TEL) sequences ligated to their ends (Figure 7-37c). The first successful experiments involving transfection of yeast cells with linear plasmids were achieved by using the ends of a DNA molecule that was known to replicate as a linear molecule in the ciliated protozoan *Tetrahymena*. During part of the life cycle of *Tetrahymena*, much of the nuclear DNA is repeatedly copied in short pieces to form a so-called *macronucleus*. One of these repeated fragments was identified as a dimer of ribosomal DNA, the ends of which contained a repeated sequence $(G_4T_2)_n$. When a section of this repeated TEL sequence was ligated to the ends of linear yeast plasmids containing ARS and CEN, replication and good segregation of the linear plasmids occurred. The initial cloning and characterization of telomeres garnered the Nobel Prize in Physiology or Medicine in 2009, which was awarded to Elizabeth Blackburn, Carol Greider, and Jack Szostak.

Centromere Sequences Vary Greatly in Length and Complexity

Once the yeast centromere regions that confer mitotic segregation were cloned, their sequences could be determined and compared. The results revealed three regions (I, II, and III) that are conserved among the centromeres on different yeast chromosomes (Figure 7-38a). Short, fairly well-conserved nucleotide sequences are present in regions I and III. Region II does not have a specific sequence, but is AT-rich with a fairly constant length, probably so that regions I and III will lie on the same side of the specialized centromere-associated histone octamer. As mentioned earlier in the discussion of variant histones, this specialized centromere-associated histone octamer contains the usual histones H2A, H2B, and H4, but a variant form of histone H3. Centromeres from all eukaryotes similarly contain nucleosomes with a specialized, centromere-specific form of histone H3, called CENP-A in humans. In the simple kinetochore of *S. cerevisiae*, a protein complex called CBF3 associates with this specialized nucleosome. The CBF3 complex, in turn, associates with several copies of an elongated multiprotein complex called Ndc80 (Figure 7-38b). The Ndc80 complexes initially make lateral interactions with a spindle microtubule and subsequently interact with a Dam1 complex, which forms a ring around the end of the microtubule (Figure 7-38c). This interaction results in an end-on attachment of the centromere to the spindle microtubule. *S. cerevisiae* has by far the simplest centromere known in nature.



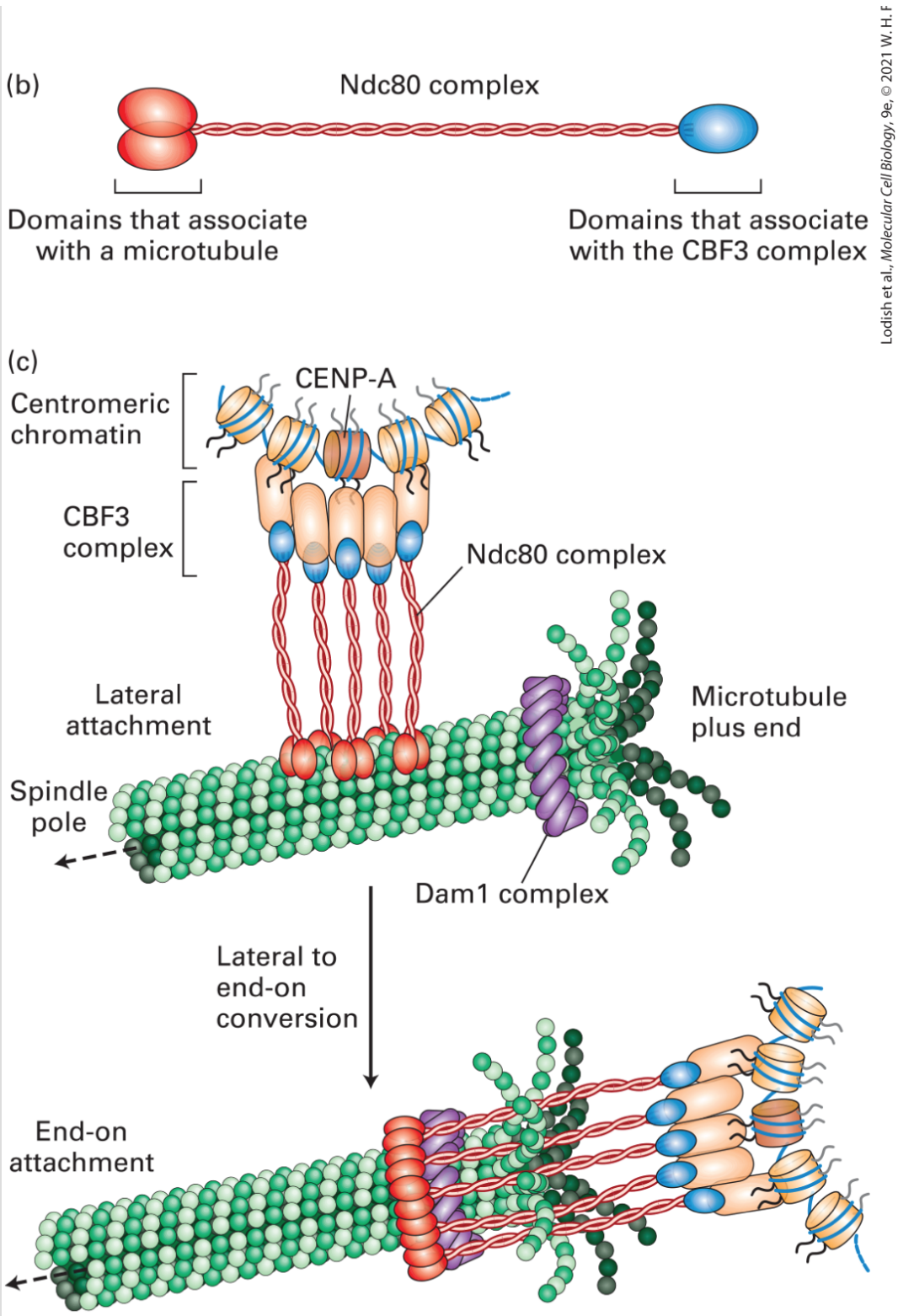


FIGURE 7-38 Kinetochore-microtubule interaction in *S. cerevisiae*. (a) Sequence of the simple centromeres of *S. cerevisiae*. Data from L. Clarke and J. Carbon, 1985, *Ann. Rev. Genet.*

19:29. (b) Ndc80 complexes associate with both the microtubule and the CBF3 complex. (c) Diagram of the centromere-associated CBF3 complex and its associated Ndc80 complexes, which associate with a ring of Dam1 proteins at the end of a spindle microtubule. The Ndc80 complexes initially make lateral interactions with the side of a spindle microtubule (*top*) and then associate with the Dam1 ring, making an end-on attachment (*bottom*) to the microtubule. See T. U. Tanaka, 2010, *EMBO J.* **29**:4070.

In the fission yeast *S. pombe*, centromeres are 40–100 kb in length and are composed of repeated copies of sequences similar to those in *S. cerevisiae* centromeres. Multiple copies of proteins homologous to those that interact with *S. cerevisiae* centromeres bind to these much larger *S. pombe* centromeres, and in turn bind the much longer *S. pombe* chromosomes to several microtubules of the mitotic spindle apparatus. In plants and animals, centromeres are megabases in length and are composed of multiple repeats of simple-sequence DNA (see Table 7-1 and Figure 7-6). In humans, centromeres contain several types of repeated simple-sequence DNAs including 2–4 Mb arrays of a 171-bp simple-sequence DNA called *alphoid* DNA, which is bound by nucleosomes containing the CENP-A histone H3 variant.

In multicellular animals and plants, centromeres are many megabases in length. A complex protein structure called the *kinetochore* assembles at these large centromeres and associates with multiple mitotic spindle fibers during mitosis (see Figure 18-43). Homologs of many of the centromere-associated proteins found in the yeasts occur in humans and other multicellular eukaryotes.

Where homologs are not evident in multicellular organisms based on amino acid sequence comparisons (such as the Dam1 complex), alternative complexes with similar properties to the yeast proteins have been proposed to function at kinetochores. The functions of the centromere and of the kinetochore proteins that bind to it during the segregation of sister chromatids in mitosis and meiosis are described in Chapters 18 and 19.

Addition of Telomeric Sequences by Telomerase Prevents Shortening of Chromosomes

Sequencing of telomeres from multiple organisms, including humans, has shown that most are repetitive oligomers with a high G content located in the strand with its 3′ end at the end of the chromosome. The telomere repeat sequence in humans and other vertebrates is TTAGGG. These simple sequences are repeated at the ends of chromosomes for a total of a few hundred base pairs in yeasts and protozoans and a few thousand base pairs in vertebrates. The 3′ end of the G-rich strand extends 12–16 nucleotides beyond the 5′ end of the complementary C-rich strand. This short single-stranded region is bound by proteins that protect the ends of linear chromosomes from attack by exonucleases.

The need for a specialized region at the ends of eukaryotic chromosomes is apparent when we consider that all known DNA polymerases elongate DNA chains at the 3′ end, and all require an RNA or DNA primer. As the replication fork approaches the end of a linear chromosome, synthesis of the leading strand continues to the end of the DNA template strand, completing one daughter DNA double helix. However, because the lagging-strand template is copied in a discontinuous fashion, it cannot be replicated in its entirety (Figure 7-39). When the final RNA primer is removed, there is no upstream strand onto which DNA polymerase can build to fill the resulting gap. Without some special mechanism, the daughter DNA strand resulting from lagging-strand synthesis would be shortened at each cell division.

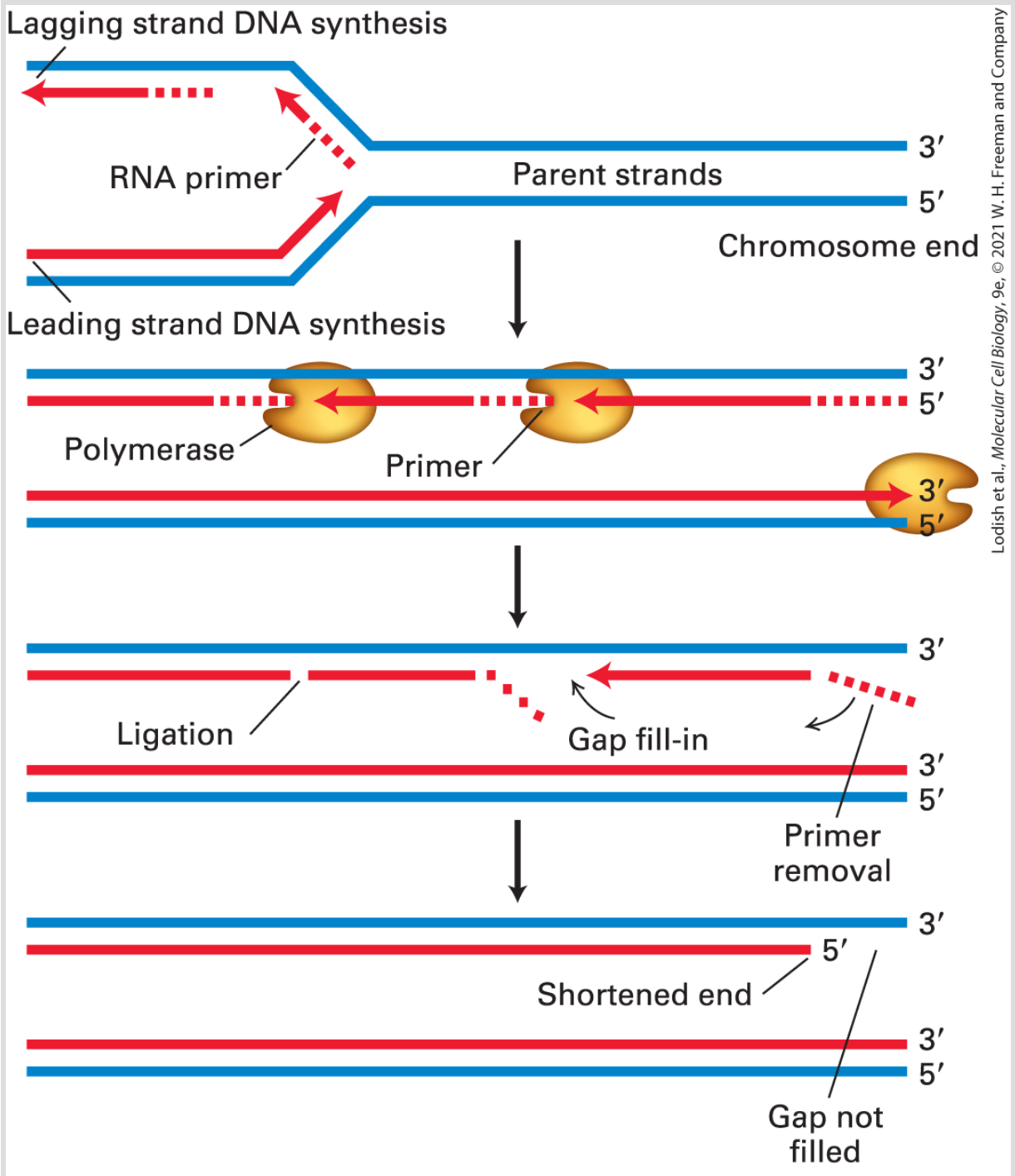


FIGURE 7-39 Standard DNA replication leads to loss of DNA at the 5′ end of each strand of a linear DNA molecule. Replication of the right end of a linear DNA is shown; the same process occurs at the left end (as can be shown by inverting the figure). As the replication fork approaches the end of the parental DNA molecule, the leading strand can be synthesized all the way to the end of the template strand without the loss of

deoxyribonucleotides. However, since synthesis of the lagging strand requires RNA primers, the right end of the lagging daughter DNA strand would remain as ribonucleotides, which are removed and therefore cannot serve as the template for a replicative DNA polymerase. Alternative mechanisms must be used to prevent successive shortening of the lagging strand with each round of replication.

The problem of telomere shortening is solved by an enzyme that adds telomeric repeat sequences to the ends of each chromosome. The enzyme is a protein-RNA complex called *telomere terminal transferase*, or *telomerase*. Because the sequence of the telomerase-associated RNA serves as the template for addition of deoxyribonucleotides to the ends of telomeres, the source of the enzyme, and not the source of the telomeric DNA primer, determines the sequence added. This was proved by transforming *Tetrahymena* with a mutated form of the gene encoding the telomerase-associated RNA. The resulting telomerase added a DNA sequence complementary to the mutated RNA sequence to the ends of telomeric primers. Thus telomerase is a specialized form of a reverse transcriptase that carries its own internal RNA template to direct DNA synthesis. These experiments also earned the Nobel Prize in Physiology or Medicine for the structure and function of telomeres in 2009.

Figure 7-40 depicts how telomerase, by reverse transcription of its associated RNA, elongates the 3′ end of the single-stranded DNA at the end of the G-rich strand mentioned above. Cells from knockout mice that cannot produce the telomerase-associated RNA exhibit no

telomerase activity, and their telomeres shorten successively with each cell generation. Such mice can breed and reproduce normally for three generations before the long telomere repeats become substantially eroded. Then, the absence of telomere DNA results in adverse effects, including fusion of chromosome termini and chromosome loss. By the fourth generation, the reproductive potential of these knockout mice declines, and they cannot produce offspring after the sixth generation.

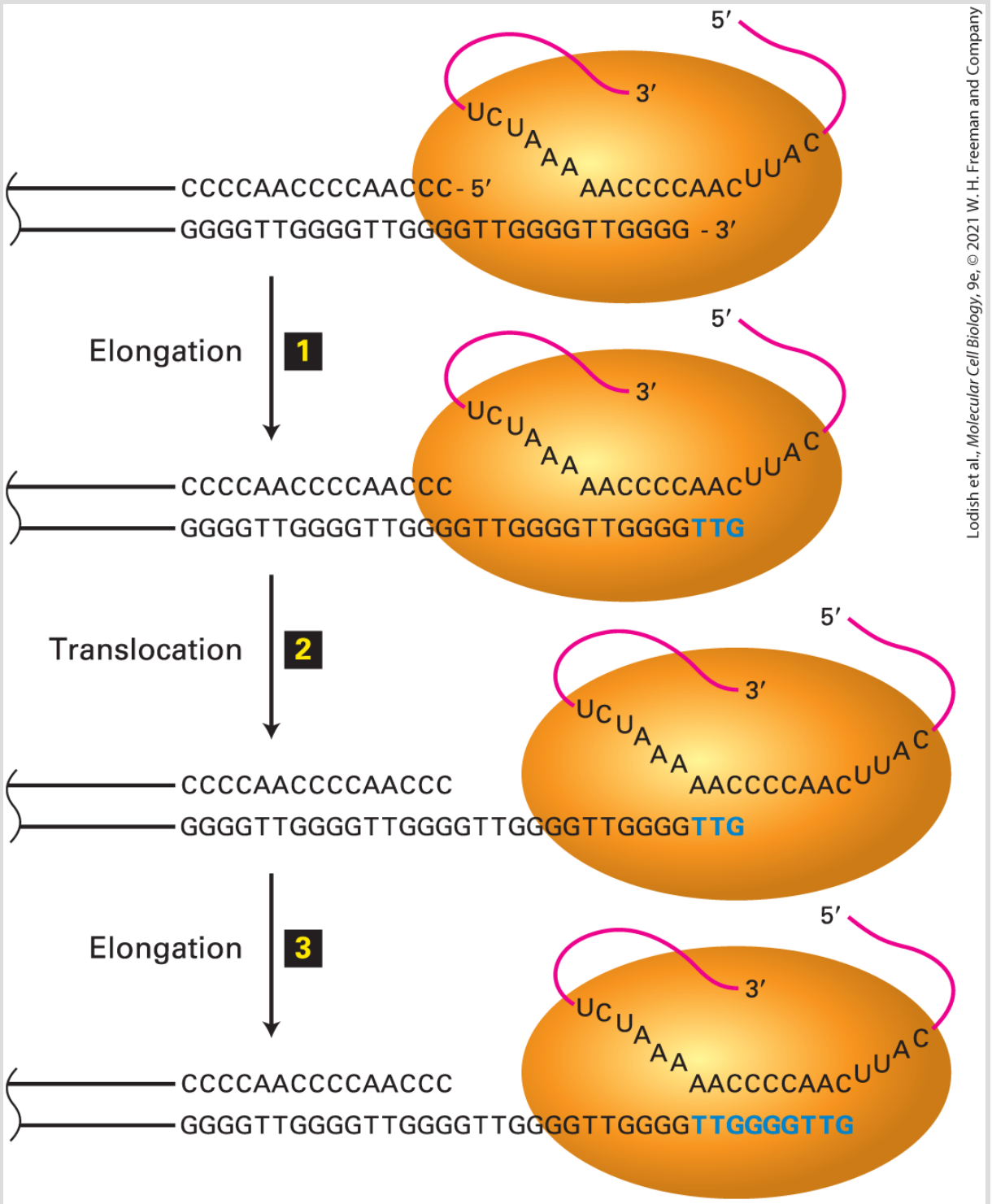


FIGURE 7-40 Mechanism of action of telomerase. The single-stranded 3′ terminus of a telomere is extended by telomerase, counteracting the inability of the DNA replication mechanism to synthesize the extreme terminus of linear DNA. Telomerase elongates this single-stranded end by a reiterative reverse-transcription mechanism. The action of the

telomerase from the protozoan *Tetrahymena*, which adds a T_2G_4 repeat unit, is depicted here; other telomerases add slightly different sequences. The telomerase contains an RNA template (red) that base-pairs to the $3'$ end of the lagging-strand template. The telomerase catalytic site then adds deoxyribonucleotides TTG (blue), using the RNA molecule as a template (step **1**). The strands of the resulting DNA-RNA duplex are then thought to slip (translocate) relative to each other so that the TTG sequence at the $3'$ end of the replicating DNA base-pairs to the complementary RNA sequence in the telomerase RNA (step **2**). The $3'$ end of the replicating DNA is then again extended by telomerase (step **3**). Telomerases can add multiple repeats by repetition of steps **2** and **3**. DNA polymerase α-primase can prime synthesis of new Okazaki fragments on this extended template strand. The net result prevents shortening of the lagging strand at each cycle of DNA replication. See C. W. Greider and E. H. Blackburn, 1989, *Nature* **337**:331.

The human genes expressing the telomerase protein and the telomerase-associated RNA are active in germ cells and stem cells, but are turned off in most cells of adult tissues that replicate only a limited number of times or will never replicate again (such cells are called *postmitotic*). However, these genes are activated in most human cancer cells, where telomerase is required for the multiple cell divisions necessary to form a tumor. This phenomenon has stimulated a search for inhibitors of human telomerase as potential therapeutic agents for treating cancer. ■

While telomerase prevents telomere shortening in most eukaryotes, some organisms use alternative strategies. *Drosophila* species maintain telomere lengths by the regulated insertion of non-LTR retrotransposons into telomeres. This is one of the few instances in which a mobile element has a specific function in its host organism.

KEY CONCEPTS OF SECTION 7.5

Morphology and Functional Elements of Eukaryotic Chromosomes

- During metaphase, eukaryotic chromosomes become sufficiently condensed that they can be visualized individually in the light microscope.
- The chromosomal karyotype is characteristic of each species. Closely related species can have dramatically different karyotypes, indicating that similar genetic information can be organized on chromosomes in different ways.
- Banding analysis and chromosome painting are used to identify the different human metaphase chromosomes and to detect translocations and deletions (see Figures 7-33 and 7-34).
- Analysis of chromosomal rearrangements and regions of conserved synteny between related species allows scientists to make predictions about the evolution of chromosomes (see Figure 7-35c). The evolutionary relationships between organisms indicated by these studies are consistent with proposed evolutionary relationships based on the fossil record and DNA sequence analysis.
- The highly reproducible banding patterns of polytene chromosomes make it possible to visualize chromosomal deletions and rearrangements as changes in the normal pattern of bands.
- Three types of DNA sequences are required for a long linear DNA molecule to function as a chromosome: a replication origin, called ARS in yeast; a centromere (CEN) sequence; and two telomere (TEL) sequences at the ends of the DNA (see Figure 7-37).
- Telomerase, a protein-RNA complex, has a special reverse transcriptase activity that completes replication of telomeres during DNA synthesis (see Figure 7-40). In the absence of telomerase, the daughter DNA strand resulting from lagging-strand synthesis is shortened at each cell division (see Figure 7-39).

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

alternative splice site
autosome
centromere
chromatid
chromatin
DNA transposon
enhancer
euchromatin
exon shuffling
fluorescence in situ hybridization (FISH)
gene family
genetic complementation

heterochromatin
histones
histone octamer
interspersed repeats
karyotype
linker DNA
long interspersed elements (LINEs)
long terminal repeats (LTRs)
micro-RNA (miRNA)
microsatellite
nucleosome
polytene chromosome
promoters
protein family
pseudogene
replication origin
retrotransposon
reverse transcriptase
short interspersed elements (SINEs)
simple-sequence (satellite) DNA
small nuclear RNA (snRNA)
small nucleolar RNA (snoRNA)
synteny
telomere
transcription factor
transcription unit
transposable (mobile) DNA element
transposition

Review the Concepts

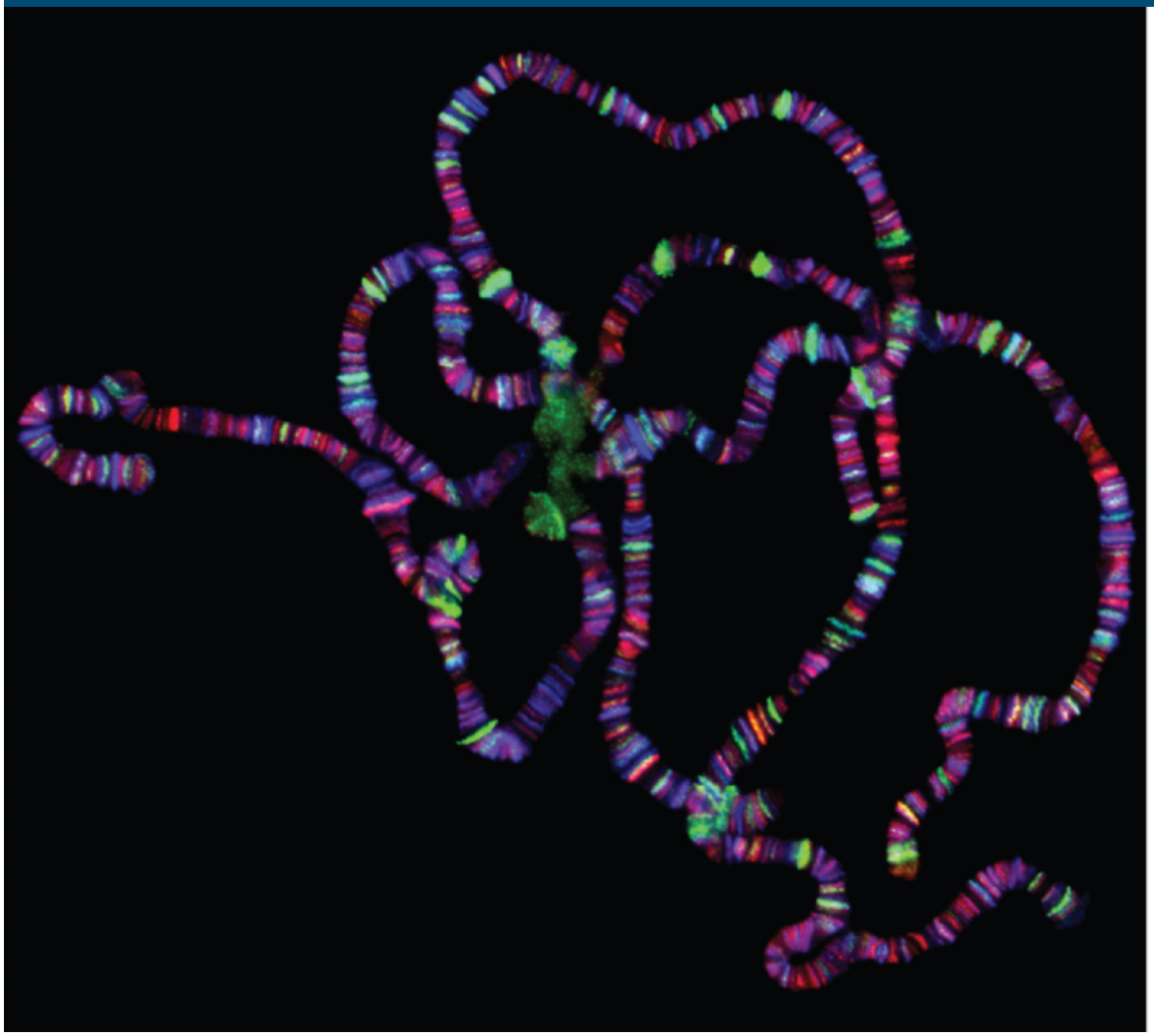
1. Genes can be transcribed into mRNA, in the case of protein-coding genes, or into RNA, in the case of genes such as those that encode ribosomal or transfer RNAs. Define a gene. For the following characteristics, state whether they apply to (a) simple or (b) complex transcription units.
   i. Found in eukaryotes
   ii. Contain introns
   iii. Capable of making multiple proteins from a given gene
2. Sequencing of the human genome has revealed much about the organization of genes. Describe the differences between solitary genes, gene families, pseudogenes, and tandemly repeated genes.
3. Much of the human genome consists of repetitive DNA. Describe the difference between microsatellite and minisatellite DNA. How is this repetitive DNA useful for identifying individuals by the technique of DNA fingerprinting?
4. Mobile DNA elements that can move or transpose to a new site directly as DNA are called DNA transposons. Describe the mechanism by which a bacterial DNA transposon, called an insertion sequence, can transpose.
5. Retrotransposons are a class of mobile elements that transpose via an RNA intermediate. Contrast the mechanism of transposition between retrotransposons that

contain long terminal repeats (LTRs) and those that lack LTRs.

6. Discuss the role that transposons may have played in the evolution of modern organisms. What is exon shuffling? What role do transposons play in the process of exon shuffling?
7. The DNA in a cell associates with proteins to form chromatin. What is a nucleosome? What role do histones play in nucleosomes? How are nucleosomes arranged in condensed 30-nm fibers?
8. How do chromatin modifications regulate transcription? What modifications are observed in regions of the genome that are being actively transcribed? In regions that are not actively transcribed?
9. What is chromosome painting, and how is this technique useful? How can chromosome paint probes be used to analyze the evolution of mammalian chromosomes?
10. Certain organisms contain cells that possess polytene chromosomes. What are polytene chromosomes, where are they found, and what function do they serve?
11. Replication and segregation of eukaryotic chromosomes require three functional elements: replication origins, a centromere, and telomeres. How would a chromosome be affected if it lacked (a) replication origins or (b) a centromere?
12. Describe the problem that occurs during DNA replication at the ends of chromosomes. How are telomeres related to this problem?

CHAPTER 8

Transcriptional Control of Gene Expression



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Drosophila polytene chromosomes stained with antibodies against a chromatin-remodeling ATPase called *Kismet* (blue), RNA polymerase II with low CTD phosphorylation (red), and RNA polymerase II with high CTD phosphorylation (green).

[Reproduced with permission from The Company of Biologists, from S. Srinivasan et. al., 2005, "The *Drosophila* Trithorax Group Protein Kismet Facilitates an Early Step in Trascriptional Elongation by RNA Polymerase II," *Development* **132**:7, 1623–1635; permission conveyed through Copyright Clearance Center, Inc.]

OUTLINE

8.1 Overview of Eukaryotic Transcription
8.2 RNA Polymerase II Promoters and General Transcription Factors
8.3 Regulatory Sequences for Protein-Coding Genes and the Proteins Through Which They Function
8.4 Molecular Mechanisms of Transcription Repression and Activation
8.5 Regulation of Transcription-Factor Activity
8.6 Epigenetic Regulation of Transcription
8.7 Other Eukaryotic Transcription Systems

In previous chapters, we have seen that the properties and functions of each cell type are determined by the proteins it contains. In this chapter and the next, we consider how the kinds and amounts of the various proteins produced by a particular cell type in a multicellular organism are regulated. This regulation of *gene expression* is the fundamental process that controls the development of multicellular organisms such as ourselves from a single fertilized egg cell into the thousands of cell types of which we are made. When gene expression goes awry, cellular properties are altered, a process that all too often leads to the development of cancer. As discussed further in Chapter 25, genes encoding proteins that restrain cell growth are abnormally repressed in cancer cells, whereas genes encoding proteins that promote cell growth and

replication are inappropriately activated in cancer cells. Abnormalities in gene expression also result in developmental defects such as cleft palate, tetralogy of Fallot (a serious developmental defect of the heart that can be treated surgically), and many others.

The basic steps in gene expression — that is, the entire process whereby the information encoded in a particular gene is decoded into a particular protein — are reviewed in Chapter 5. Synthesis of mRNA requires that an *RNA polymerase* initiate transcription (**initiation**), polymerize ribonucleoside triphosphates complementary to the DNA-coding strand (**elongation**), and then terminate transcription (**termination**). In bacteria, ribosomes and translation initiation factors have immediate access to newly formed RNA transcripts, which function as mRNA without further modification. In eukaryotes, however, the initial RNA transcript is subjected to processing that yields a functional mRNA (see Figure 5-27). The mRNA then is transported from its site of synthesis in the nucleus to the cytoplasm, where it is translated into protein with the aid of ribosomes, tRNAs, and translation factors.

Regulation occurs at several of the steps in gene expression outlined above: transcription initiation, elongation, RNA processing, and mRNA export from the nucleus, as well as through control of mRNA degradation, mRNA translation into protein, and protein degradation. The diversity of options for regulating gene expression

results in *differential* protein expression that varies among cell types, developmental stages, or in response to external conditions. Although examples have been found of regulation at each step in gene expression, in multicellular animals (metazoans), the control of transcription initiation and elongation are the most important mechanisms (Figure 8-1). The molecular mechanisms that regulate transcription initiation and elongation are critical for normal development of a multicellular organism, for proper immune responses that protect us from pathogenic microorganisms, and for neurological processes such as learning and memory. When these regulatory mechanisms controlling transcription function improperly, pathological processes may occur. For example, mutations of the *HOXD13* gene result in *polydactyly*, the embryological development of extra digits of the feet and hands (Figure 8-2a). *HOXD13* encodes a *transcription factor* that regulates transcription of multiple genes involved in development of the extremities. Other mutations affecting the function or expression of transcription factors cause an extra pair of wings to develop in *Drosophila* (Figure 8-2b) or alter the structures of flowers in plants (Figure 8-2c), among many other developmental abnormalities.

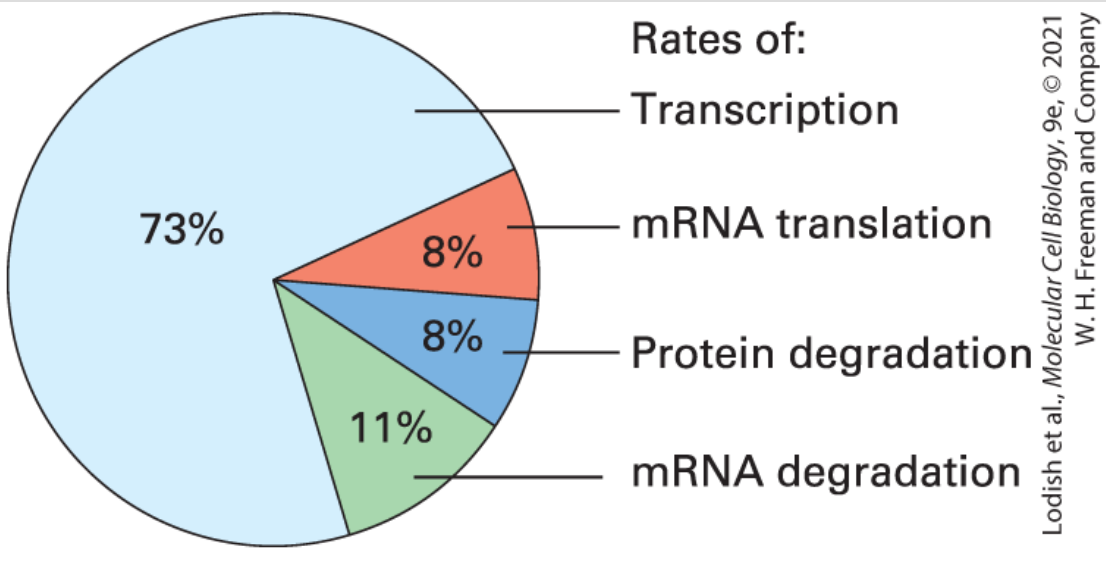


FIGURE 8-1 Contributions of the major processes that regulate protein concentrations. The concentration of a protein is controlled by regulation of the frequency with which the mRNA encoding the protein is synthesized (gene transcription), the rate at which that mRNA is degraded, the rate at which that mRNA is translated into protein, and the rate at which that protein is degraded. The relative contributions of these four rates to determining the concentrations of thousands of proteins in cultured mouse fibroblasts were determined by mass spectrometry to measure protein concentrations (see Chapter 3), mRNA sequencing (RNA-seq) to measure mRNA levels (see Chapter 6), protection of mRNA from ribonuclease digestion by associated ribosomes (ribosome footprinting) to estimate translation rates, stable isotope labeling to determine degradation rates, and statistical analysis of the data to correct for inherent biases and errors in these methods.

[Data from J. J. Li and M. D. Biggin, 2014, *Science* **347**:1066.]

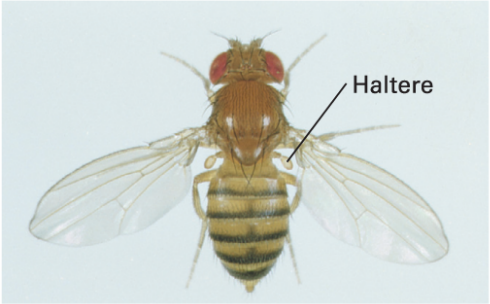
(a)

Normal

Dominant *HOXD13* mutation

Part (a) *left*: Lightvision, LLC/Getty Images

(b)



Normal

Ubx mutation

(c)

Normal

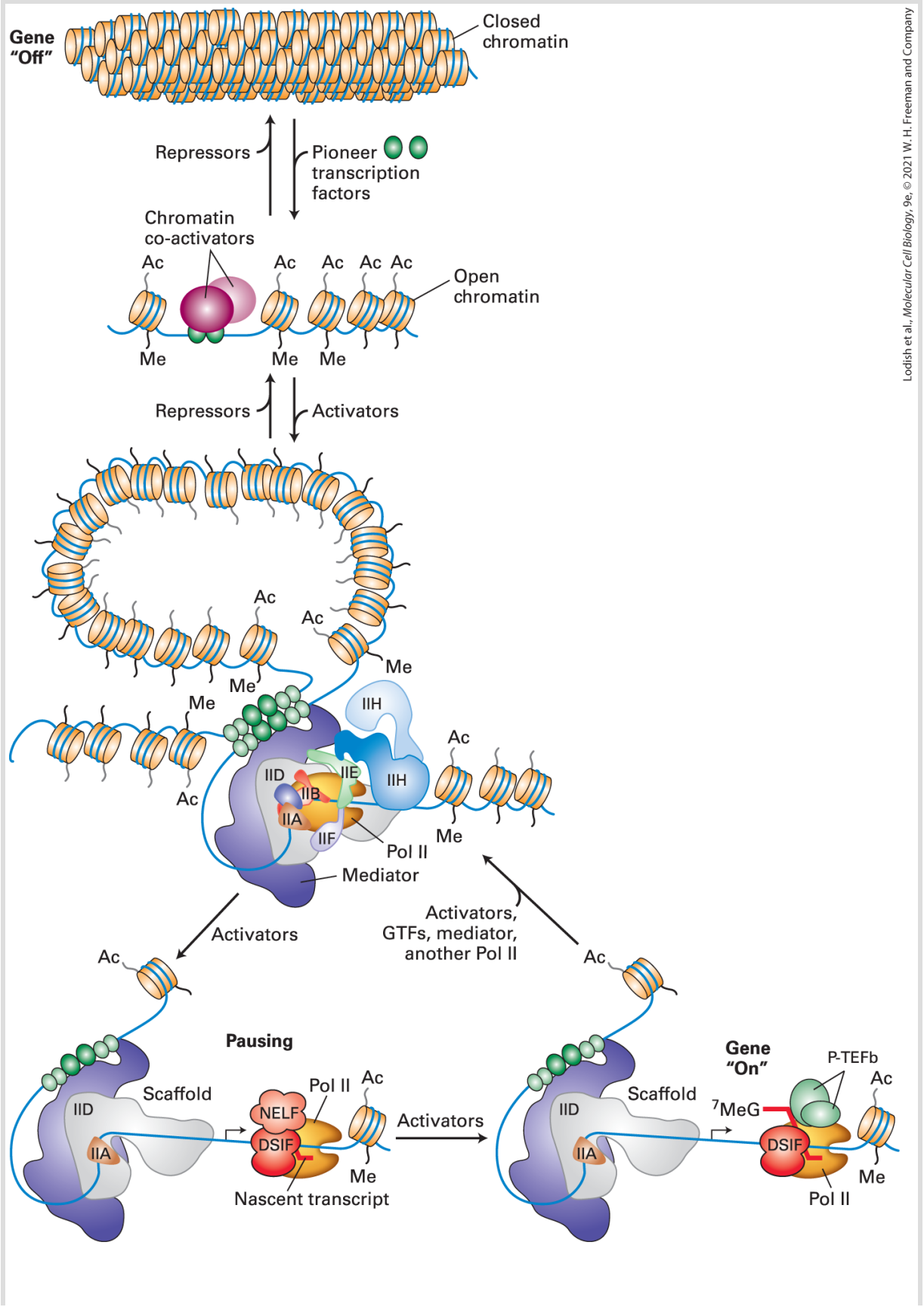
Homozygous recessive mutations in *ap2-1*, *pi-1*, and *ag-1* genes

FIGURE 8-2 Phenotypes of mutations in genes encoding transcription factors. (a) A dominant mutation in the human *HOXD13* gene results in the development of extra digits, a condition known as polydactyly. (b) Homozygous recessive mutations that prevent expression of the *Ubx* gene in the third thoracic segment of *Drosophila* result in transformation of that segment, which normally has a balancing organ called a *haltere*, into a second copy of the thoracic segment that develops wings. (c) Mutations in *Arabidopsis thaliana* that inactivate both copies of three floral organ–identity genes transform the normal parts of the flower into leaflike structures. In each case, these mutations affect master regulatory transcription factors that regulate multiple genes, including many genes encoding other transcription factors.

[Part (a) *right*: Republished with permission of John Wiley & Sons, Inc., from F. R. Goodman and P. J. Scrambler, 2001, "Human Hox Gene Mutations," *Clin. Genet.* **59**(1):1–11. Part (b) Reproduced with permission from E. B. Lewis, "The Bithorax Complex: The First Fifty Years," *Int. J. Dev. Biol. 42*:403–415, Figures 4a and 4b. Part (c) Republished with permission of Elsevier, from D. Weigel and M. Meyerowitz, 1994, "The ABCs of Floral Homeotic Genes," *Cell* **78**(2):203–209; permission conveyed through Copyright Clearance Center, Inc.]

Transcription is a complex and highly regulated process that results in the expression of specific genes in specific cell types, despite the fact that nearly all cells contain the same chromosomes with the same DNA sequences. In this chapter, we focus on the molecular events in eukaryotic cells that determine when transcription of a gene occurs. Eukaryotic transcription regulation mechanisms make use of the association of DNA with histone octamers, by forming chromatin structures with varying degrees of condensation, and through post-translational modifications of histone tails (Figures 7-26 and 7-28). Figure 8-3 provides an overview of transcription regulation in metazoans and of the processes outlined in this

chapter. We discuss how the RNA polymerases responsible for transcription of different classes of eukaryotic genes bind to promoter sequences to initiate the synthesis of an RNA molecule and how specific DNA sequences function as **transcription-control regions** by serving as the binding sites for transcription factors that regulate transcription. Next we consider how eukaryotic activators and repressors influence transcription through interactions with large multiprotein complexes. Some of these multiprotein complexes modify chromatin condensation, altering the accessibility of chromosomal DNA to transcription factors and RNA polymerases. Other complexes directly influence the frequency at which RNA polymerases bind to promoters and initiate transcription. In metazoans, the RNA polymerase pauses after transcribing a short RNA, and one mechanism for regulating transcription of a gene involves release of the paused polymerase, allowing it to transcribe the rest of the gene. We discuss how transcription of specific genes can be specified by particular combinations of the roughly 1600 transcription factors encoded in the human genome, giving rise to cell-type-specific gene expression. We consider the various ways in which the activities of transcription factors themselves are controlled to ensure that genes are expressed only in the correct cell types and at the appropriate time during their differentiation.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 8-3 Overview of eukaryotic transcriptional control. Inactive genes are assembled into regions of condensed chromatin that inhibit RNA polymerases and their initiation factors from binding to promoters. A pioneer transcription factor is able to bind to a specific regulatory sequence in DNA within the condensed chromatin and interact with chromatin co-activators that decondense, acetylate (Ac), and methylate (Me) the chromatin, thus making the DNA more accessible to binding by other activators (green), RNA polymerase II (the polymerase that transcribes protein coding genes), and its general transcription factors (IIA, IIB, … IIH). Additional activator proteins bound to both promoter-proximal sites and to distant enhancers interact with one another and with the multisubunit Mediator complex to assemble RNA polymerase II (Pol II) and general transcription factors on promoters. Alternatively, repressor proteins bind to other transcription-control elements to inhibit transcription initiation by Pol II. They do this by associating with multiprotein co-repressor complexes to condense chromatin. During transcriptional activation, Pol II initiates transcription, but pauses after transcribing fewer than 100 nucleotides due to the action of an elongation inhibitor (NELF) associated with a protein that clamps the polymerase on the DNA template (DSIF). Activators promote the association of the paused complex with an elongation factor (P-TEFb) that releases NELF and allows productive elongation through the gene. See S. Malik and R. G. Roeder, 2010, *Nat. Rev. Genet*. **11**:761.

We also discuss recent studies revealing that RNA-protein complexes in the nucleus can regulate transcription. New methods for sequencing DNA, coupled with reverse transcription of RNA into DNA in vitro, have revealed that much of the genome of eukaryotes is transcribed into low-abundance RNAs that do not encode proteins. Several nuclear *long noncoding RNAs (lncRNAs)* regulate the transcription of protein-coding genes. This finding raises the possibility that transcriptional control by noncoding RNAs may be much more general than is currently understood. Recent advances in mapping the association of transcription factors with specific regions of chromatin across the entire genome in a variety of cell

types have provided the first glimpses of how transcription factors regulate embryonic development from the pluripotent stem cells of the early embryo to the fully differentiated cells that make up most of our tissues. RNA processing and various post-transcriptional mechanisms for controlling eukaryotic gene expression are covered in Chapter 9. Subsequent chapters, particularly Chapters 15, 16, and 21, provide examples of how transcription is regulated by interactions between cells and how the resulting **gene control** contributes to the development and function of the thousands of specific types of cells in multicellular organisms.

8.1 Overview of Eukaryotic Transcription

In bacteria, gene control serves mainly to allow a single cell to adjust to rapid changes in its environment, such as the availability of nutrients, so that its growth and division can be optimized. In contrast, in multicellular organisms the cellular environment is relatively stable. As with bacteria, though, when sudden changes in the environment do occur, eukaryotic cells rapidly induce sets of genes that protect themselves, as we will see in the cellular responses to heat-shock and hypoxia discussed in Chapter 21. However, the most characteristic and biologically far-reaching purpose of gene control in multicellular organisms is execution of the genetic program that underlies embryological development. Generation of the many different cell types that collectively form a multicellular organism depends on the right genes being activated in the right cells at the right time during the developmental period.

In most cases, once a developmental step has been taken by a cell, it is not reversed. In executing their genetic programs, many differentiated cells (e.g., skin cells, red blood cells, and antibody-producing cells) march down a pathway to final cell death, leaving no progeny behind. The fixed patterns of gene control leading to differentiation serve the needs of the whole organism and not the survival of an individual cell.

Bacteria and eukaryotes share two key features of transcriptional control. First, transcription-control regions are associated with genes. Second, specific proteins that bind to a gene's transcription-control regions determine where and how frequently transcription will initiate. Eukaryotic cells exploit chromatin structure to regulate transcription; a mechanism of transcriptional control that is not available to bacteria. In multicellular eukaryotes, many inactive genes are assembled into condensed regions of chromatin, which inhibit binding of the RNA polymerases and general transcription factors required for transcription initiation (see Figure 8-3). **Activator** proteins, which bind to transcription-control regions near the transcription start site of a gene, and also bind to transcription control regions called *enhancers* located kilobases from the transcription start site, promote chromatin decondensation, binding of RNA polymerase to the promoter, and transcriptional elongation. **Repressor** proteins, which bind to alternative control elements, cause condensation of chromatin and inhibition of polymerase binding or elongation. In this section, we discuss general principles of eukaryotic gene control. Subsequent sections of this chapter will address specific aspects of eukaryotic transcription regulation in greater detail.

Regulatory Elements in Eukaryotic DNA Are Found Both Close to and

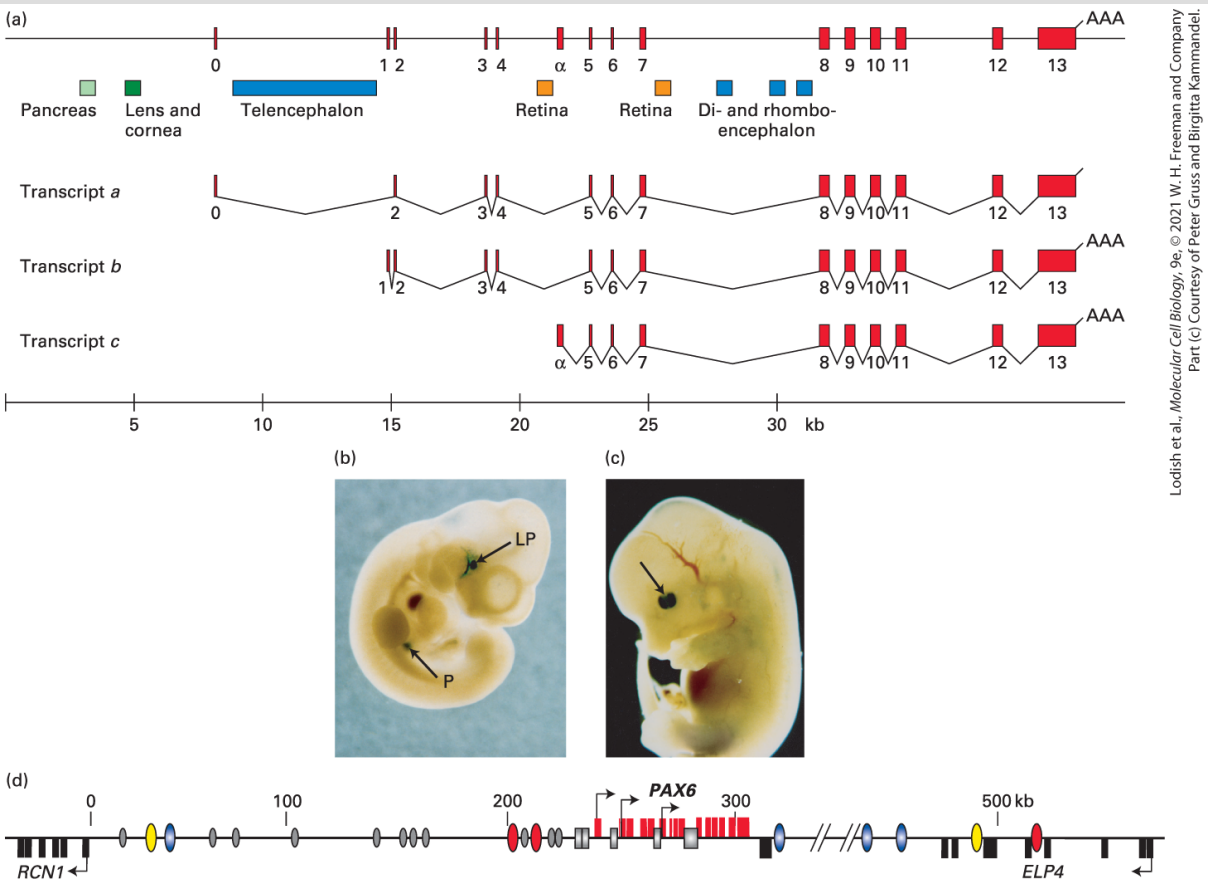
Many Kilobases Away from Transcription Start Sites

Direct measurements of the transcription rates of multiple genes in different cell types have shown that regulation of transcription, either at the initiation step or during elongation in the promoter proximal region, is the most widespread form of gene control in eukaryotes (as it is in bacteria). In eukaryotes, as in bacteria, a DNA sequence that specifies where RNA polymerase binds and initiates transcription of a gene is called a **promoter**. Transcription from a particular promoter is controlled by DNA-binding proteins that bind to transcription control regions. Individual eukaryotic transcriptional regulatory proteins often can function to either activate or repress transcription, depending on their associations with other proteins. Consequently, they are referred to by the general term **transcription factors**.

The DNA-control elements in eukaryotic genomes to which transcription factors bind are often located much farther from the promoter they regulate than is the case in bacterial genomes. In some cases, transcription factors bind at regulatory sites tens of thousands of base pairs either **upstream** (opposite to the direction of transcription) or **downstream** (in the same direction as transcription) from the promoter. As a result, transcription of a gene may be regulated by the binding of multiple transcription factors to alternative control elements. This directs expression of the same

gene in different types of cells at different times during development.

For example, several transcription-control regions regulate expression of the mammalian gene encoding the transcription factor *Pax6*. As mentioned in Chapter 1, Pax6 protein is required for development of the eye. Pax6 is also required for development of certain regions of the brain and spinal cord and for cells in the pancreas that secrete insulin. Humans with only one functional *Pax6* gene are born with *aniridia*, a lack of irises in the eyes (Figure 1-31d). In mammals, the *Pax6* gene is expressed from at least three alternative promoters that function in different cell types and at different times during embryogenesis (Figure 8-4a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Part (c) Courtesy of Peter Gruss and Birgitta Kammandel.

FIGURE 8-4 Transcription-control regions of the mouse *Pax6* gene and the orthologous human *PAX6* gene. (a) Three alternative *Pax6* promoters are used at distinct times during embryogenesis in different tissues of the developing mouse embryo. Transcription-control regions regulating expression of *Pax6* in different tissues are indicated by colored rectangles. These control regions are some 200–500 bp in length. (b) Expression of a β-galactosidase reporter transgene fused to the 8 kb of mouse DNA upstream from exon 0. A transgenic mouse embryo 10.5 days after fertilization was stained with X-gal to reveal β-galactosidase. Lens pit (LP) is the tissue that will develop into the lens of the eye. Expression was also observed in tissue that will develop into the pancreas (P). (c) Expression in a mouse embryo at 13.5 days after fertilization of a β-galactosidase reporter gene linked to the sequence in part (a) between exons 4 and 5 marked *Retina*. Arrow points to nasal and temporal regions of the developing retina. (d) Human *PAX6* control regions identified in the 600-kb region between the upstream gene RCN1 and the promoter of the downstream ELP4 gene. RCN1 and ELP4 are transcribed to the left. Their exons are shown as black rectangles below the line representing the DNA sequence. *PAX6* exons are diagrammed as red rectangles above the line. The three *PAX6* promoters are shown by rightward arrowheads, and the control

regions shown in (a) are represented by gray rectangles. Regions flanking the gene where the sequence is partially conserved in most vertebrates (as in Figure 8-5a) are shown as ovals. Colored ovals represent sequences that cause expression of a transgene in specific neuroanatomical locations in the zebrafish where transcription control regions can be assayed much more rapidly than in transgenic mice.

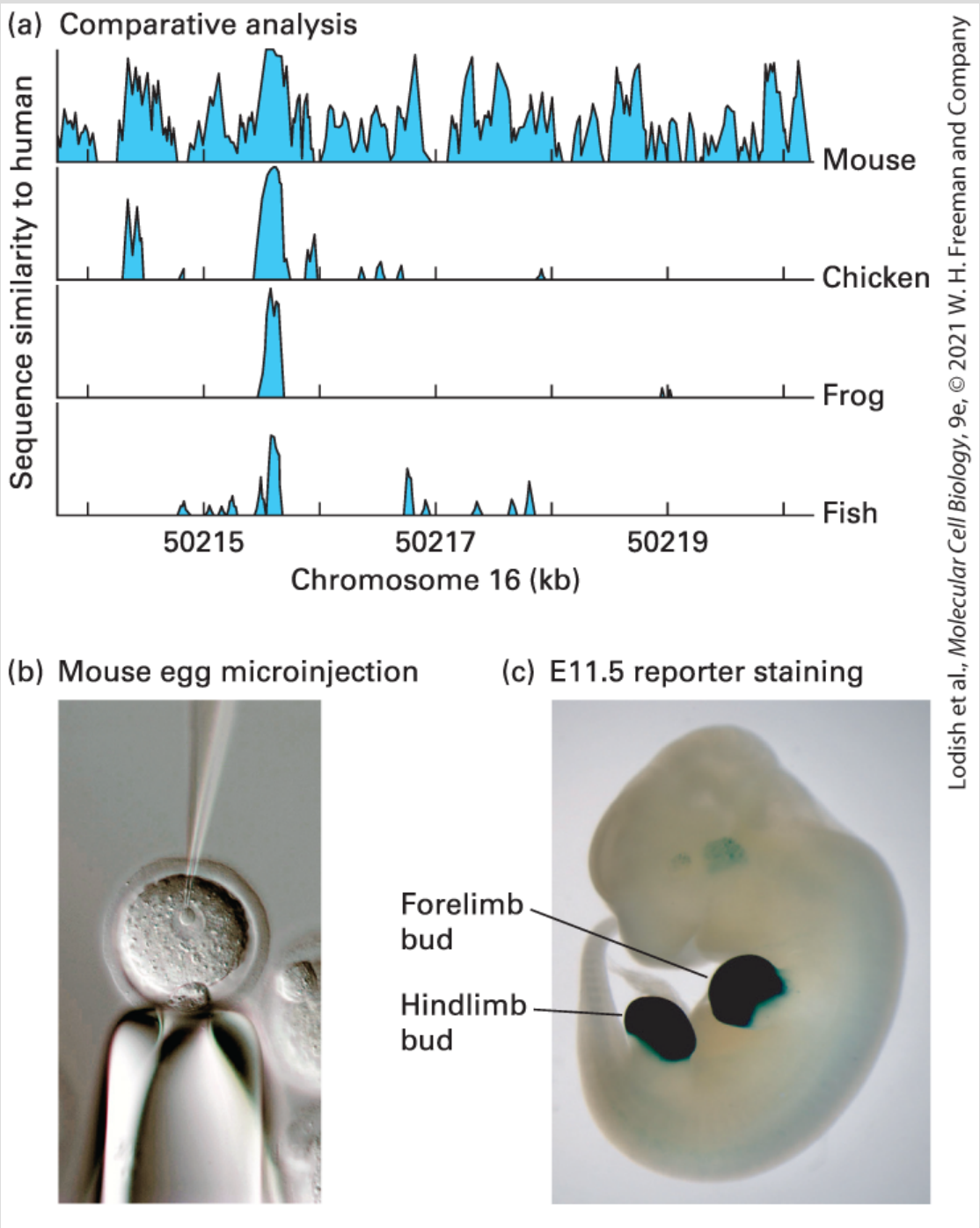
[Part (a) Data from B. Kammendal et al., 1999, *Devel. Biol.* **205**:79. Part (b) Republished with permission of Elsevier, B. Kammendal et al., 1999, "Distinct *cis*-Essential Modules Direct the Time–Space Pattern of the *Pax6* Gene Activity," *Devel. Biol.* **205**(1): 79–97; permission conveyed through Copyright Clearance Center, Inc. Part (d) Data from S. Bhatia et al., 2014, *Devel. Biol.* 387:214.]

Researchers often analyze transcription-control regions by preparing recombinant DNA molecules that combine a fragment of DNA to be tested for its ability to regulate transcription with the coding region for a **reporter gene** whose expression is easily assayed. Typical reporter genes include the gene that encodes luciferase, an enzyme that generates light that can be assayed with great sensitivity and over many orders of magnitude of intensity using a luminometer. Other frequently used reporter genes encode green fluorescent protein (GFP) and proteins that fluoresce other colors, which can be visualized by fluorescence microscopy (see Figures 4-9d and 4-16), and *Escherichia coli* β-galactosidase, which generates an intensely blue insoluble precipitate when incubated with the colorless soluble lactose analog X-gal. When transgenic mice are produced containing a β-galactosidase reporter gene fused to 8 kb of DNA upstream from *Pax6* exon 0 (see Figure 8-4a), β-galactosidase is observed in the developing lens, cornea, and pancreas of the embryo halfway through gestation (Figure 8-4b).

Analysis of transgenic mice with smaller fragments of DNA from this region allowed the mapping of the separate transcription-control regions regulating transcription in the pancreas and in both the lens and cornea. Transgenic mice with other reporter gene constructs revealed additional transcription-control regions (see Figure 8-4a). These regions control transcription in the developing retina and in different regions of the developing brain (encephalon). Some of these transcription-control regions are in introns between exons 4 and 5 and between exons 7 and 8. For example, a reporter gene under control of the region labeled *Retina* in Figure 8-4a between exons 4 and 5 led to reporter-gene expression specifically in the retina (Figure 8-4c).

Control regions for many genes are found hundreds of kilobases away from the coding exons of the gene. One method for identifying such distant control regions is to compare the sequences of distantly related organisms. Transcription-control regions for a conserved gene are also often conserved and can be recognized in the background of nonfunctional sequences that diverge during evolution. For example, there is a DNA sequence about 500 kb downstream of the *SALL1* gene that is highly conserved among humans, mice, chickens, frogs, and fish (Figure 8-5a). *SALL1* encodes a transcription factor required for normal development of the limbs. When transgenic mice were produced containing this conserved DNA sequence linked to a β-galactosidase reporter gene (Figure 8-5b), the transgenic embryos expressed a very high level of

β-galactosidase in the developing limb buds (Figure 8-5c). Human patients with deletions in this region of the genome develop with limb abnormalities. These results indicate that this conserved region directs transcription of the *SALL1* gene in the developing limb. Presumably, other transcription-control regions regulate expression of this gene in other types of cells, where it functions in the normal development of the ears, the lower intestine, and kidneys. After discussing the proteins that function with RNA polymerase to carry out transcription in eukaryotic cells, we will return to a discussion of how such distant transcription-control regions, called *enhancers*, function.



EXPERIMENTAL FIGURE 8-5 The human *SALL1* enhancer activates expression of a reporter gene in limb buds of the developing mouse embryo. (a) Graphic representation of the conservation of DNA sequence in a region of the human genome (in the interval of chromosome 16 from 50214 kb to 50220.5 kb) about 500 kb downstream from the *SALL1* gene, which encodes a zinc-finger transcription repressor. A region of roughly 500 bp of nonprotein-coding sequence is conserved

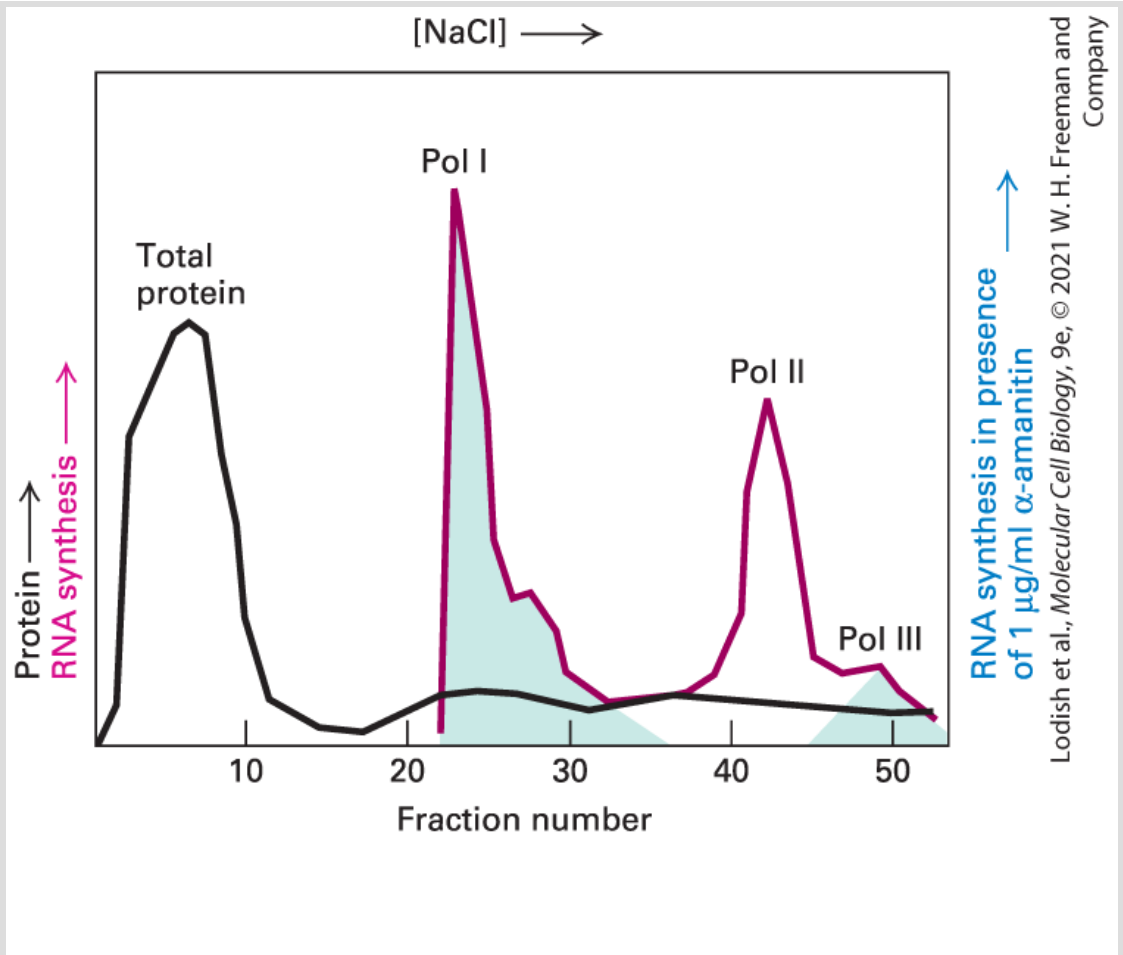
from zebra fish to human. Nine hundred base pairs of human DNA including this conserved region were inserted into a plasmid next to the coding region for *E. coli* β-galactosidase. (b) The plasmid was microinjected into a pronucleus of a fertilized mouse egg and implanted in the uterus of a pseudopregnant mouse to generate a transgenic mouse embryo with the reporter-gene-containing plasmid incorporated into its genome. (c) After 11.5 days of development, at the time when limb buds develop, the fixed and permeabilized embryo was incubated in X-gal, which is converted by β-galactosidase into an insoluble, intensely blue compound. The results showed that the conserved region contains an enhancer that stimulates strong transcription of the β-galactosidase reporter gene specifically in limb buds.

[Part (a) Data from the VISTA Enhancer Browser, http://enhancer.lbl.gov. Parts (b) Deco/Alamy and (c) Republished with permission of Nature Publishing Group, from L. A. Pennacchio et al., 2006, "In Vivo Enhancer Analysis of Human Conserved Non-Coding Sequences," *Nature* **444**:499–506; permission conveyed through Copyright Clearance Center, Inc.]

Three Eukaryotic Nuclear RNA Polymerases Catalyze Formation of Different RNAs

The nuclei of all eukaryotic cells examined so far contain three different RNA polymerases, designated I, II, and III. These enzymes elute at different salt concentrations during ion-exchange chromatography, reflecting the differences in their net charges. The three nuclear RNA polymerases also differ in their sensitivity to α-amanitin, a poisonous cyclic octapeptide produced by some mushrooms (Figure 8-6). RNA polymerase I is insensitive to α-amanitin, but RNA polymerase II is very sensitive — the drug binds

near the active site of the enzyme and inhibits translocation of the enzyme along the DNA template. RNA polymerase III has intermediate sensitivity. Each eukaryotic RNA polymerase catalyzes transcription of genes encoding different classes of RNA (Table 8-1). *RNA polymerase I* (Pol I), located in the nucleolus, transcribes genes encoding precursor rRNA (pre-rRNA), which is processed into 28S, 5.8S, and 18S rRNAs. *RNA polymerase III* (Pol III) transcribes genes encoding tRNAs, 5S rRNA, and an array of small stable RNAs, including one involved in RNA splicing (U6) and the RNA component of the signal recognition particle (SRP) involved in directing nascent proteins to the endoplasmic reticulum (see Chapter 13).



EXPERIMENTAL FIGURE 8-6 Liquid chromatography separates and identifies the three eukaryotic RNA polymerases, each with its own sensitivity to α-amanitin. A protein extract from the nuclei of cultured eukaryotic cells was passed through a DEAE Sephadex column and adsorbed protein eluted (black curve) with a solution of constantly increasing NaCl concentration. An aliquot of each fraction of eluate collected from the column was assayed for RNA polymerase activity without (red curve) and with (green shading) 1 μg/ml α-amanitin. This concentration of α-amanitin inhibits polymerase II activity but has no effect on polymerases I and III. Polymerase III is inhibited by 10 μg/ml of α-amanitin (not shown), whereas polymerase I is unaffected even at this higher concentration.

[Data from R. G. Roeder, 1974, *J. Biol. Chem.* **249**:241.]

TABLE 8-1 • Classes of RNA Transcribed by the Three Eukaryotic Nuclear RNA Polymerases and Their Functions

Polymerase	RNA Transcribed	RNA Function
RNA polymerase I	Pre-rRNA (28S, 18S, 5.8S rRNAs)	Ribosome components, protein synthesis
RNA polymerase II	mRNA	Encodes protein
	snRNAs	RNA splicing
	siRNAs	Chromatin-mediated repression, translation control
	miRNAs	Translation control
RNA	tRNAs	Protein synthesis

polymerase III		
	5S rRNA	Ribosome component, protein synthesis
	snRNA U6	RNA splicing
	7S RNA	Signal recognition particle for insertion of polypeptides into the endoplasmic reticulum
	Other small stable RNAs	Various functions, unknown for many

RNA polymerase II (Pol II) transcribes all protein-coding genes: that is, it functions in production of mRNAs. RNA polymerase II also produces four of the five small nuclear RNAs (snRNAs) that take part in RNA splicing and micro-RNAs (miRNAs) involved in translation control, as well as the closely related endogenous small interfering RNAs (siRNAs) and most of the long noncoding RNAs (lncRNAs) (see Chapter 9). We will be focusing on the regulation of Pol II in this chapter, but we will discuss the regulation of Pol I and Pol III at the end of the chapter.

The eukaryotic RNA polymerases are more complex than bacterial RNA polymerase, but all RNA polymerases have a similar overall design (Figure 8-7a, b). Each eukaryotic RNA polymerase contains two large subunits and 10–14 smaller subunits, some of which are common between two or all three of the polymerases. The best characterized eukaryotic RNA polymerases are from the yeast

Saccharomyces cerevisiae. Each of the yeast genes encoding the polymerase subunits has been subjected to gene-knockout mutations (see Chapter 6) and the resulting phenotypes characterized. In addition, the three-dimensional structure of yeast RNA polymerase II has been determined (Figure 8-7b, c). The three nuclear RNA polymerases from all eukaryotes so far examined are very similar to those of yeast. Plants contain two additional nuclear RNA polymerases (RNA polymerases IV and V), which are closely related to their RNA polymerase II but have a unique large subunit and some additional unique subunits. RNA polymerases IV and V function in transcriptional repression of plant transposons directed by nuclear siRNAs in plants. We discuss them toward the end of this chapter.

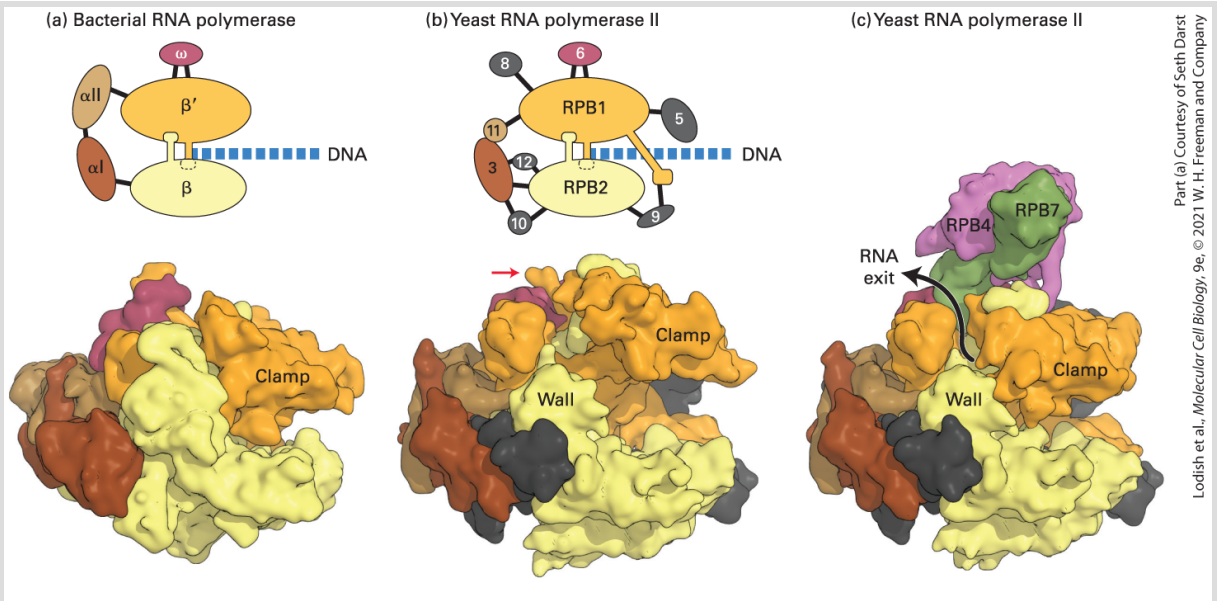


FIGURE 8-7 Comparison of three-dimensional structures of bacterial and eukaryotic RNA polymerases. These space-filling models are based on x-ray crystallographic analysis. (a) RNA polymerase from the bacterium *T. aquaticus*. The five subunits of the bacterial enzyme are distinguished by color. Only the N-terminal domains of the α subunits are included in

this model. (b) Core RNA polymerase II from *S. cerevisiae*. Ten of the 12 subunits constituting yeast RNA polymerase II are shown in this model. Subunits that are similar in conformation to those in the bacterial enzyme are shown in the same colors. The C-terminal domain of the large subunit RPB1 was not observed in the crystal structure, but it is known to extend from the position marked with a red arrow. (RPB is the abbreviation for "*R*NA *p*olymerase *B*," which is an alternative way of referring to RNA polymerase II.) DNA entering the polymerases as they transcribe to the right is diagrammed in the top row. (c) Space-filling model of yeast RNA polymerase II including subunits 4 and 7. These subunits extend from the core portion of the enzyme shown in (b) near the region of the C-terminal domain of the large subunit.

[Part (a) see N. Korzheva et al., 2000, *Science* **289**:619–625. Part (b) Data from P. Cramer et al., 2001, *Science* **292**:1863. Part (c) Data from K. J. Armache et al., 2003, *Proc. Nat'l. Acad. Sci. USA* **100**:6964, and D. A. Bushnell and R. D. Kornberg, 2003, *Proc. Nat'l. Acad. Sci. USA* **100**:6969.]

The two large subunits of all eukaryotic RNA polymerases are related in sequence and structure and are also similar to the large *E. coli* RNA polymerase subunits, β′ and β (Figure 8-8 and see Figure 8-7a). Each of the eukaryotic RNA polymerases also contains an ω-like and two nonidentical α-like subunits (see Figure 8-8). The extensive similarity in the structures of these core subunits in RNA polymerases from diverse organisms indicates that RNA polymerase arose early in evolution and has been largely conserved. This seems logical for an enzyme catalyzing a process as fundamental as the copying of RNA from DNA. In addition to the core subunits, all yeast RNA polymerases contain four additional small subunits, which are not in the bacterial RNA polymerase (see Figure 8-8). Finally, each eukaryotic nuclear RNA polymerase has several unique enzyme-specific subunits (see Figure 8-8). Three of these additional subunits of Pol I and Pol III are homologous to the three additional Pol II-

specific subunits. The other two Pol I-specific subunits are homologous to the Pol II general transcription factor TFIIF, discussed later, and the four additional subunits of Pol III are homologous to the Pol II general transcription factors TFIIF and TFIIE. The functions of these general transcription-factor-like subunits in Pol I and Pol III are like those of TFIIE and TFIIF. But during enzyme purification, these subunits remain associated with RNA polymerases I and III, whereas the functionally similar RNA polymerase II subunits separate from RNA polymerase II during purification.

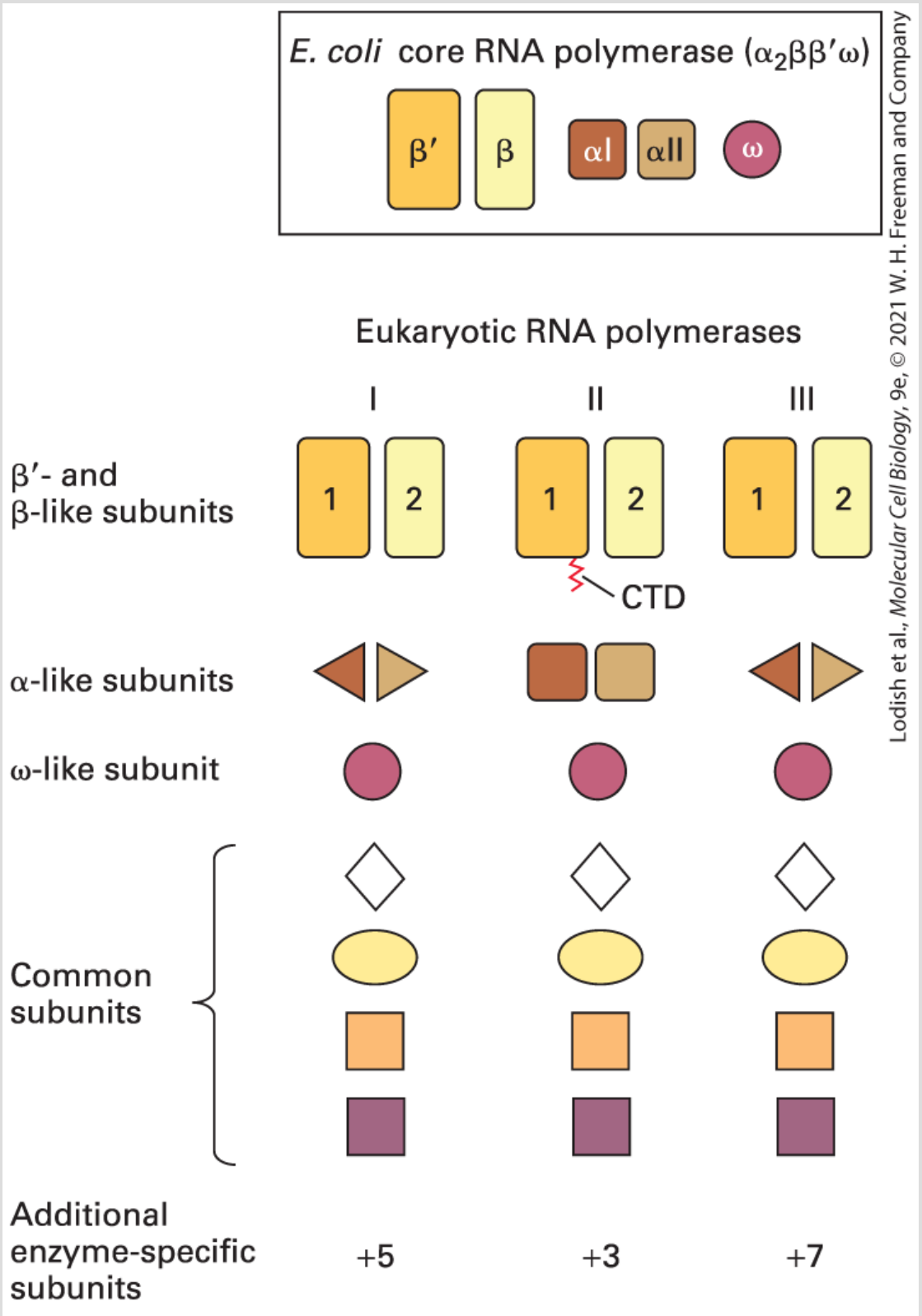


FIGURE 8-8 Schematic representation of the subunit structure of the *E. coli* RNA core polymerase and yeast nuclear RNA polymerases. All three yeast polymerases

have five core subunits homologous to the β, β′, two α, and ω subunits of *E. coli* RNA polymerase. The largest subunit (RPB1) of RNA polymerase II also contains an essential C-terminal domain (CTD). RNA polymerases I and III contain the same two nonidentical α-like subunits, whereas RNA polymerase II contains two other nonidentical α-like subunits. All three polymerases share the same ω-like subunit and four other common subunits. In addition, each yeast polymerase contains three to seven unique smaller subunits.

Archaea, like bacteria, have a single type of RNA polymerase involved in gene transcription, but archaeal RNA polymerases, like eukaryotic nuclear RNA polymerases, have on the order of a dozen subunits. Archaea also possess general transcription factors that are related to those of eukaryotes, which is consistent with the closer evolutionary relationship between archaea and eukaryotes than between bacteria and eukaryotes (see Figure 1-1).

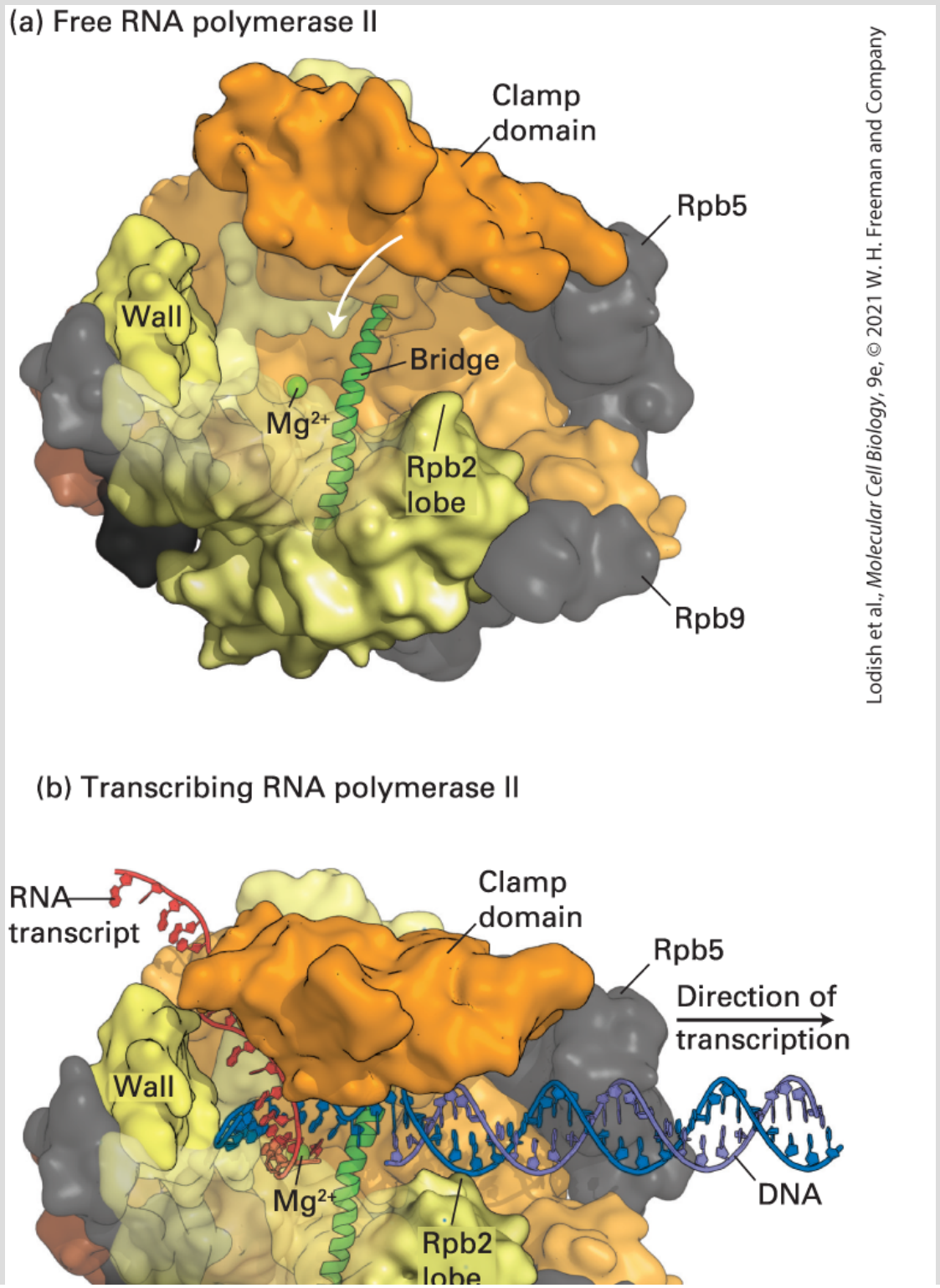
Gene-knockout experiments in yeast indicate that most of the subunits of the three nuclear RNA polymerases are essential for cell viability. Disruption of the genes encoding the few polymerase subunits that are not essential for viability (e.g., subunits 4 and 7 of RNA polymerase II) nevertheless results in cells that grow very poorly. Thus all of the subunits are necessary for eukaryotic RNA polymerases to function normally.

The Clamp Domain Enables RNA Polymerase II to Transcribe Long

Stretches of DNA

Subunit RPBI contains the *clamp domain*, so called because it has been observed to clamp down on DNA that enters the enzyme. In crystals of free Pol II (Figure 8-9a) and in a complex that mimics the elongating form of the enzyme (Figure 8-9b), the clamp domain is in different positions. The domain rotates on a hinge that is open when downstream DNA is inserted into the cleft between RPB1 and RPB2 subunits (see Figure 8-7b), and then swings shut when the enzyme is in its elongation mode (Figure 8-9b). It is postulated that when the 8–9-bp RNA-DNA hybrid region near the active site (where RNA is base-paired to the template strand; see Figure 8-9b) is bound between RBP1 and RBP2 and nascent RNA fills the exit channel, the clamp is locked in its closed position, anchoring the polymerase to the downstream double-stranded DNA. Furthermore, a transcription elongation factor called *DSIF*, discussed later, associates with the elongating polymerase, holding the clamp in its closed conformation. As a consequence, the polymerase is extraordinarily processive, which is to say that it continues to transcribe down the template DNA polymerizing ribonucleotides until transcription is terminated. After termination and release of RNA from the exit channel, the clamp can swing open, releasing the enzyme from the template DNA. This mechanism explains how human RNA polymerase II can transcribe the longest human gene, encoding dystrophin, which is some 2 million base pairs in length, without dissociating and terminating transcription. Since transcription

elongation proceeds at 1–2 kb per minute, transcription of the *DMD* gene requires approximately one day!



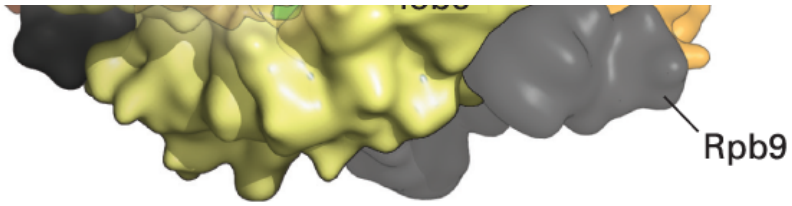


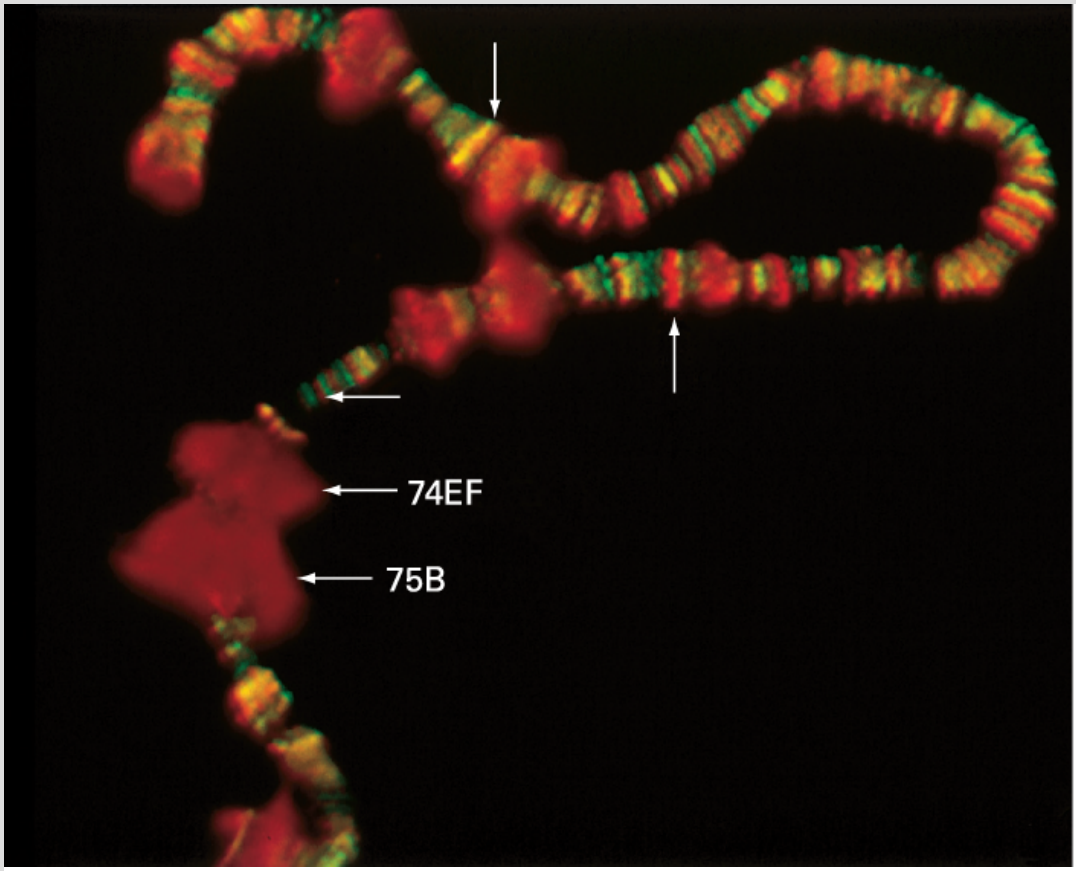
FIGURE 8-9 The clamp domain of RPBI. The structures of the free (a) and transcribing (b) RNA polymerase II differ mainly in the position of a clamp domain in the RPB1 subunit (orange), which swings over the cleft between the jaws of the polymerase during formation of the transcribing complex, trapping the template DNA strand and transcript. Binding of the clamp domain to the 8–9-bp RNA-DNA hybrid helps couple clamp closure to the presence of RNA, stabilizing the closed, elongating complex. RNA is shown in red, the template DNA strand in dark blue, and the downstream nontemplate DNA strand in cyan in this model of an elongating complex. The clamp closes over the incoming downstream DNA. This model is shown with portions of RBP2 that form one side of the cleft removed so that the nucleic acids can be better visualized. The Mg^{2+} ion that participates in catalysis of phosphodiester bond formation is shown in green. Wall is the domain of RPB2 that forces the template DNA entering the jaws of the polymerase to bend before it exits the polymerase. The bridge α helix, shown in green, extends across the cleft in the polymerase (see Figure 8-7b) and is postulated to bend and straighten as the polymerase translocates one base down the template strand. The nontemplate strand is thought to form a flexible, single-stranded region above the cleft (not shown), extending from three bases downstream of the template base-paired to the $3'$ base of the growing RNA to where the template strand exits the polymerase, where it hybridizes with the nontemplate strand to generate the transcription bubble.

[Data from A. L. Gnatt et al., 2001, *Science* **292**:1876, PDB ID 1i6h.]

The Largest Subunit in RNA Polymerase II Has an Essential Carboxy-Terminal Repeat

RNA polymerase II contains a unique domain at the carboxyl end of its RPB1 subunit. The domain is made of seven amino acids repeated multiple times. This heptapeptide repeat, with a consensus sequence of Tyr-Ser-Pro-Thr-Ser-Pro-Ser, is known as the **carboxy-terminal domain (CTD)** (see Figure 8-8, red squiggle). Yeast RNA polymerase II contains 26 or more of these repeats, vertebrate enzymes have 52 repeats, and an intermediate number of repeats occurs in RNA polymerase II from nearly all other eukaryotes. The CTD is critical for cell viability, and at least 10 copies of the repeat must be present for yeast to survive.

In vitro experiments with model promoters first showed that RNA polymerase II molecules that initiate transcription have a nonphosphorylated CTD. Once the polymerase initiates transcription and begins to move away from the promoter, many of the serine and some tyrosine residues in the CTD are phosphorylated. Analysis of polytene chromosomes from *Drosophila* salivary glands prepared just before molting of the larva, a time of active transcription, indicates that the CTD is also phosphorylated during in vivo transcription. The large chromosomal "puffs" induced at this time in development are regions where the genome is very actively transcribed. Staining with antibodies specific for either the phosphorylated or nonphosphorylated CTD demonstrated that RNA polymerase II associated with the highly transcribed puffed regions contains a phosphorylated CTD (Figure 8-10).



courtesy of J. R. Weeks and A. L. Greenleaf.

EXPERIMENTAL FIGURE 8-10 Antibody staining demonstrates that the carboxy-terminal domain of RNA polymerase II is phosphorylated during in vivo transcription. Salivary-gland polytene chromosomes were prepared from *Drosophila* larvae just before they molted. The preparation was treated with a rabbit antibody specific for phosphorylated CTD and with a goat antibody specific for nonphosphorylated CTD. The preparation was then stained with fluorescein-labeled anti-goat antibody (green) and rhodamine-labeled anti-rabbit antibody (red). Thus polymerase molecules with a nonphosphorylated CTD stained green, and those with a phosphorylated CTD stained red. The molting hormone ecdysone induces very high rates of transcription in the puffed regions labeled 74EF and 75B; note that only phosphorylated CTD is present in these regions. Smaller puffed regions transcribed at high rates are also visible. Nonpuffed sites that stained red (up arrow) or green (horizontal arrow) are also indicated, as is a site staining both red and green, producing a yellow color (down arrow).

[Republished with permission from Cold Spring Harbor Press, from: J. R. Weeks et al., 1993, “Locus-Specific Variation In Phosphorylation State of RNA Polymerase II In

Vivo: Correlations with Gene Activity and Transcript Processing," *Genes Dev.* **7**(12A):2329–2344;]

KEY CONCEPTS OF SECTION 8.1

Overview of Eukaryotic Transcription

- The primary purpose of gene control in multicellular organisms is the execution of precise developmental programs so that the proper genes are expressed in the proper cells at the proper times during embryologic development and cellular differentiation.
- Transcriptional control is the primary means of regulating gene expression in eukaryotes, as it is in bacteria.
- In eukaryotic genomes, DNA transcription-control elements may be located many kilobases away from the promoter they regulate. Different control elements can control transcription of the same gene in different cell types.
- Eukaryotes contain three types of nuclear RNA polymerases. All three contain two large and three smaller core subunits with homology to the β', β, α, and ω subunits of *E. coli* RNA polymerase, as well as several additional small subunits (see Figure 8-8).
- RNA polymerase I synthesizes only pre-rRNA. RNA polymerase II synthesizes mRNAs, some of the small nuclear RNAs that participate in mRNA splicing, micro- and small interfering RNAs (miRNAs and siRNAs) that regulate the translation and stability of mRNAs, and long noncoding (lnc) RNAs that regulate transcription. RNA polymerase III synthesizes tRNAs, 5S rRNA, and several other small stable RNAs (see Table 8-1).
- The carboxy-terminal domain (CTD) in the largest subunit of RNA polymerase II becomes phosphorylated during transcription initiation and remains phosphorylated as the enzyme transcribes the DNA template.

8.2 RNA Polymerase II Promoters and General Transcription Factors

The mechanisms that regulate transcription initiation and elongation by RNA polymerase II have been studied extensively because this polymerase is the one that transcribes mRNAs. Transcription initiation and elongation by RNA polymerase II are the initial biochemical processes required for the expression of protein-coding genes. These are the steps in gene expression that are most frequently regulated and thus determine when and in which cells specific proteins are synthesized. As we mentioned in the previous section, the expression of eukaryotic protein-coding genes is regulated by multiple protein-binding DNA sequences, generically referred to as *transcription-control regions*. These sequences include promoters, which determine where transcription of the DNA template begins, and other types of control elements located near transcription start sites. They also include sequences, called *enhancers*, located far from the genes they regulate, which regulate cell type-specific transcription and how frequently specific genes are transcribed. In this section, we take a closer look at the properties of various transcription-control elements found in eukaryotic protein-coding genes and some techniques used to identify them.

RNA Polymerase II Initiates Transcription at DNA Sequences Corresponding to the 5′ Cap of mRNAs

In vitro transcription experiments using purified RNA polymerase II, a protein extract prepared from the nuclei of cultured cells, and DNA templates containing sequences encoding the 5′ ends of mRNAs for a number of abundantly expressed genes revealed that the transcripts produced always contained a cap structure at their 5′ ends identical to that present at the 5′ end of the spliced mRNA normally expressed from the gene in vivo (see Figure 5-26). In these experiments, the 5′ cap was added to the 5′ end of the nascent RNA by enzymes in the nuclear extract, which can add a cap only to an RNA that has a 5′ tri- or diphosphate. Because a 5′ end generated by cleavage of a longer RNA would have a 5′ monophosphate, it would not be capped. Consequently, researchers concluded that the capped nucleotides generated in the in vitro transcription reactions must have been the nucleotides with which transcription was initiated. Sequence analysis revealed that, for any given gene, the sequence at the 5′ end of the RNA transcripts produced in vitro is the same as that at the 5′ end of the mRNAs isolated from cells, confirming that the capped nucleotide of eukaryotic mRNAs coincides with the transcription start site. Today the transcription start site for a newly characterized mRNA is generally determined simply by identifying

the DNA sequence encoding the 5′-capped nucleotide of the encoded mRNA.

The TATA Box, Initiators, and CpG Islands Function as Promoters in Eukaryotic DNA

Several different types of DNA sequences can function as promoters for RNA polymerase II, directing the polymerase where to initiate transcription of an RNA complementary to the template strand of a double-stranded DNA molecule. These sequences include TATA boxes, initiators, and CpG islands.

TATA Boxes

The first genes to be sequenced and studied through in vitro transcription systems were viral genes and cellular protein-coding genes that are very actively transcribed, either at particular times of the cell cycle or in specific differentiated cell types. In all these highly transcribed genes, a conserved sequence called the **TATA box** was found about 26–31 bp upstream of the transcription start site (Figure 8-11). Mutagenesis studies have shown that a single-base change in this nucleotide sequence drastically decreases in vitro transcription of the gene adjacent to it. If the base pairs between the TATA box and the normal transcription start site are deleted,

transcription of the altered, shortened template begins at a new site about 25 bp downstream from the TATA box.

Initiator Sequences, BREs and DPEs

Instead of a TATA box, some eukaryotic genes contain an alternative core promoter element called an **initiator**. Most naturally occurring initiator elements have a cytosine (C) at the −1 position and an adenine (A) residue at the transcription start site (+1). Directed mutagenesis of mammalian genes with an initiator-containing promoter revealed that the nucleotide sequence immediately surrounding the start site determines the strength of such promoters. In contrast to the conserved TATA box sequence, however, only an extremely degenerate initiator consensus sequence has been defined:

$$(5')\ \text{Y-Y-A}^{+1}\text{-N-T/A-Y-Y-Y}\ (3')$$

where A^{+1} is the base at which transcription starts, Y is a pyrimidine (C or T), N is any of the four bases, and T/A is T or A at position +3. As discussed below, transcription initiation by RNA polymerase II requires several *general transcription factors* (GTFs) that bind with the polymerase to template DNA. The largest of these, TFIID, has two subunits, which together bind the initiator sequence in promoters where it occurs.

Other DNA sequences within ~40 base pairs of the transcription start site can also influence the frequency of transcription initiation. Another GTF, TFIIB, binds to the major groove of DNA immediately upstream of the TATA box, and the strongest promoters contain the optimal sequence for this interaction, called the *TFIIB recognition element (BRE)* (Figure 8-11). Another promoter element designated *DPE* for *downstream promoter element* (Figure 8-11), can be bound by other subunits of TFIID (see below).

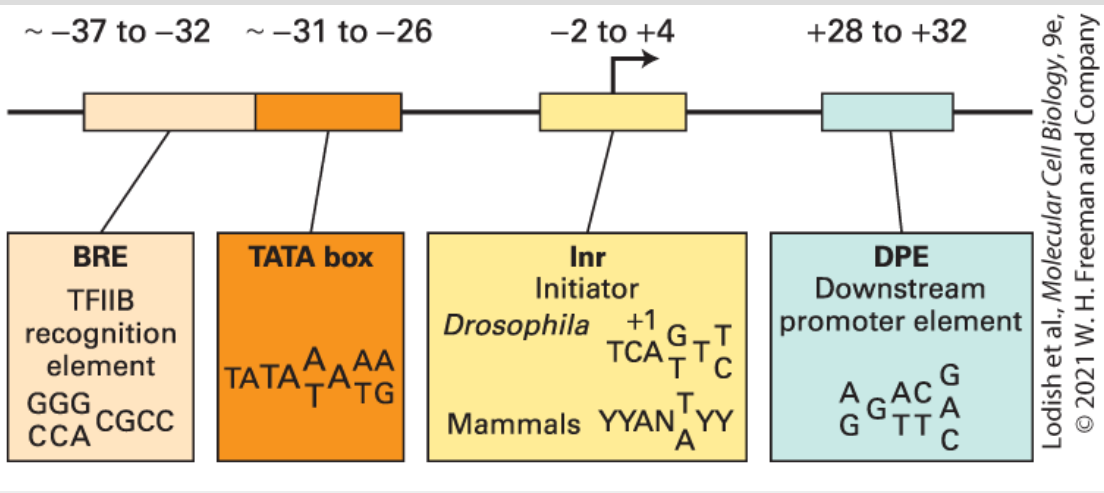


FIGURE 8-11 Core promoter elements in metazoans. The sequence of each element is shown with the 5′ end at the left and the 3′ end at the right.

CpG Islands

Transcription of genes with promoters containing a TATA box or initiator element begins at a well-defined initiation site. However, the transcription of ~70 percent of protein-coding genes in mammals occurs at a lower rate than at TATA box–containing and

initiator-containing promoters. For these genes, transcription begins at any of several alternative start sites within regions about 100–1000-bp long that have an unusually high frequency of CG sequences compared to the rest of the genome. Many such genes encode proteins that are not required in large amounts (e.g., genes encoding enzymes involved in basic metabolic processes required in all cells, often called "housekeeping genes"). These promoter regions are called **CpG islands** (where "p" represents the phosphate between the C and G nucleotides) because the sequence 5′ CG 3′ occurs relatively rarely in the genome sequences of mammals, for reasons explained shortly. Within CpG island promoters, the sequence 5′ CG 3′ occurs several times within a few tens of base pairs — a very unusual occurrence in mammalian genomes — leading to the image of rare islands of sequences containing multiple CGs in a sea of genomic sequence bereft of them.

The low frequency of the sequence CG in mammalian genomes is thought to have occurred by the following mechanism: In mammals, most Cs followed by a G are methylated at position 5 of the pyrimidine ring (5-methyl C, represented as C^{Me}; see Figure 2-17). This allows DNA repair mechanisms to distinguish the parental strand of newly replicated DNA containing C^{Me} from the newly synthesized daughter strand lacking C^{Me}, which is more likely to include a mutation introduced by an error in copying the parental strand. In mammalian genomes, CG sequences occur at only ~20 percent their expected frequency based on the percentages of GC

and AT base pairs in their genomes. This is thought to be because the slow, spontaneous deamination of 5-methyl C generates thymidine. Over the time scale of mammalian evolution, this is thought to have led to the conversion of most CGs to TG. As a consequence, the frequency of CG in the human genome is only 21 percent of that expected if Cs were randomly followed by a G. However, the Cs in active CpG island promoters are unmethylated. Consequently, when they deaminate spontaneously, they are converted to U, a base that is recognized by DNA-repair enzymes and converted back to C. As a result, the frequency of CG sequences within CpG island promoters (where they are not methylated) is close to that expected if C were followed by any of the other three nucleotides randomly.

CG-rich sequences are bound by histone octamers more weakly than CG-poor sequences because more energy is required to bend them into the small-diameter loops required to wrap around the histone octamer forming a nucleosome (see Figure 7-20). As a consequence, CpG islands coincide with linker regions of DNA between nucleosomes (see Figure 7-22). Much remains to be learned about the molecular mechanisms that control transcription from CpG island promoters, but a current hypothesis is that the general transcription factors discussed in the next section can bind to them because access to the DNA is not inhibited by DNA interactions with the surface of a histone octamer.

Divergent Transcription from CpG Island Promoters

Another remarkable feature of CpG islands is that transcription from these elements is initiated in both directions, even though only transcription of the sense strand yields an mRNA. By a mechanism(s) that remains to be fully elucidated, most RNA polymerase II molecules transcribing in the "wrong" direction — that is, transcribing the nonsense strand — terminate transcription about 1–3 kb from the transcription start site. This phenomenon was discovered by taking advantage of the stability conferred on the elongation complex by the RNA polymerase II clamp domain when an RNA-DNA hybrid is bound near the active site (see Figure 8-9b).

In one key experiment, nuclei were isolated from cultured human fibroblasts and incubated in a buffered solution containing salt and mild detergent, which removes RNA polymerases except for those in the process of elongation because of their stable association with template DNA. Nucleotide triphosphates were then added, with UTP replaced by bromo-UTP, containing uracil with a Br atom at position 5 on the pyrimidine ring (see Figure 2-17). The nuclei were then incubated at 30 °C long enough for about 100 nucleotides to be polymerized by the RNA polymerase II molecules that were in the process of elongation at the time the nuclei were isolated. RNA was isolated, and RNA containing bromo-U was immunoprecipitated with an antibody specific for BrU-labeled RNA. Thirty-three nucleotides at the 5′ ends of tens of millions of these isolated RNA

molecules were then sequenced by massively parallel DNA sequencing (see Chapter 6) of reverse transcripts, and the sequences were mapped on the human genome. The results of this experiment indicated that approximately equal numbers of RNA polymerase molecules transcribed most promoters (mostly CpG island promoters) in the sense direction *and* in the antisense direction, away from the gene. A peak of sense transcripts was observed at about +50 relative to the major transcription start site (TSS), indicating that at most genes in mammals, Pol II pauses in the +50 to +200 region before elongating further. A peak at −250 to −500 relative to the major transcription start site was also observed for Pol II transcribing in the antisense direction.

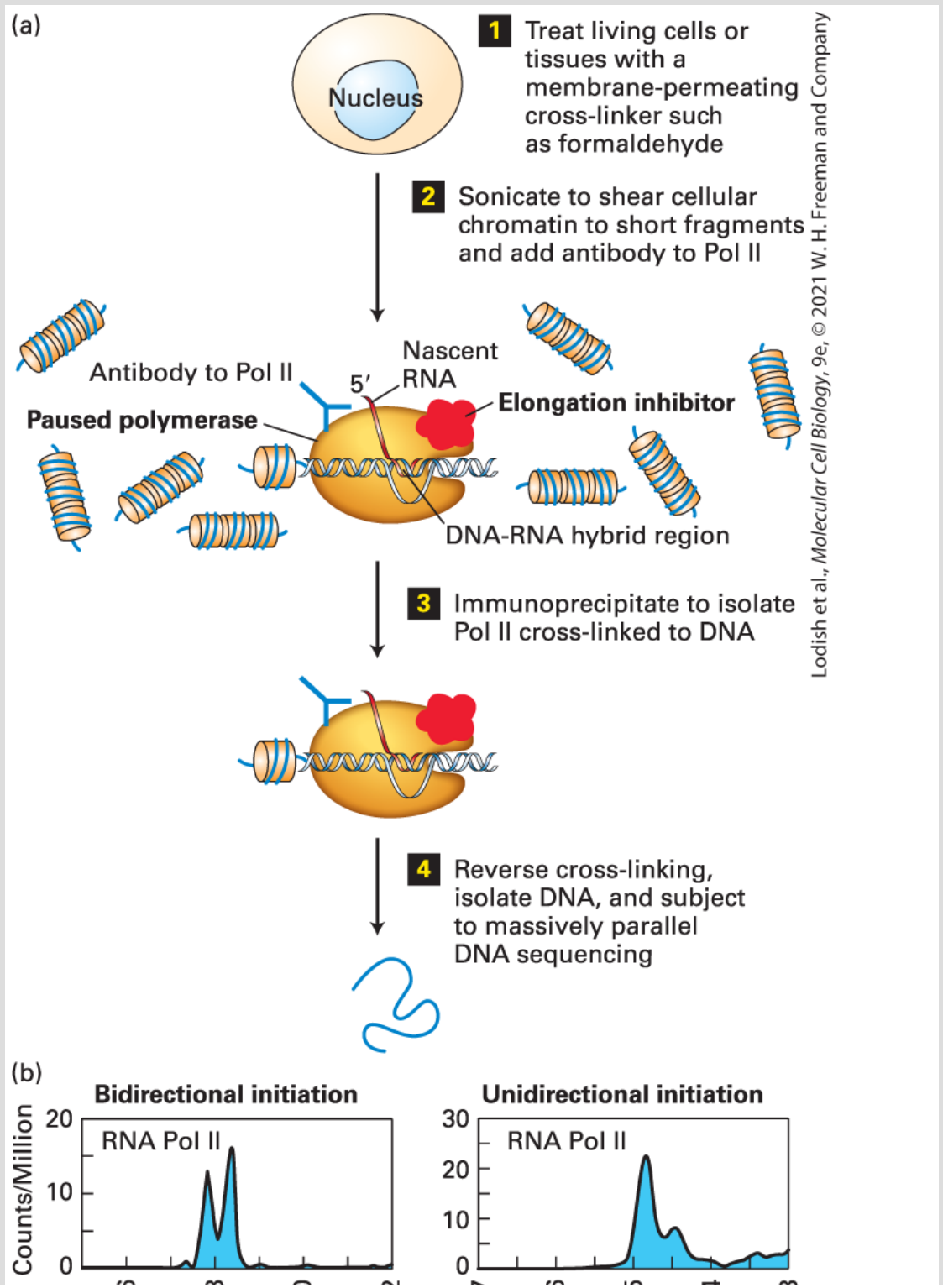
These observations, which greatly surprised most molecular biologists when they were first reported, are explained by understanding how transcription factors that bind to promoter sequences function, as we discuss in the following section. Briefly here, multiple processes that regulate the accessibility of DNA at promoters allow proteins known as *general transcription factors* to bind to promoters. At strong promoters with a TATA box or an initiator sequence, the direction of Pol II binding is determined by these promoter sequences and transcription occurs primarily in the sense direction. But at weaker promoters, the general transcription factors and Pol II associate with the promoter DNA in both possible directions randomly, resulting in transcription by half of the polymerases in one direction, and transcription by the other half of

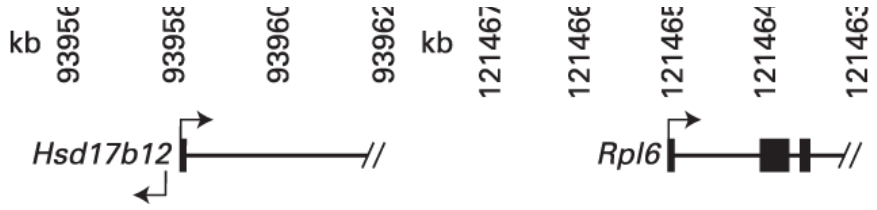
polymerases in the opposite direction. Then, RNA processing, discussed in Chapter 9, results in degradation of the RNAs that were transcribed in the wrong direction so that they are not processed into mRNAs. This apparent waste of energy from widespread antisense transcription followed by degradation of the antisense RNA apparently was not selected against during the evolution of mammals. This is probably because the most abundantly expressed genes are transcribed primarily in the sense direction, and because the amount of energy devoted to mRNA synthesis is only a small fraction of the energy required for other processes essential to mammals such as the movement of our muscles and the constant pumping of the ions required for neurological activity (Chapter 23).

The widely used technique of **chromatin immunoprecipitation** outlined in Figure 8-12a provided additional data supporting the occurrence of divergent transcription from most CpG island promoters in mammals. The method allows one to determine the multiple binding sites of a specific protein along nearly the entire length of the 3×10^9 bp human genome, within a resolution of $\sim$300 base pairs. First, starting with living cells, all proteins are rapidly cross-linked to nearby proteins and other macromolecules including DNA. This is usually done by adding formaldehyde to the media. The cross-linked chromatin is isolated and fragmented into lengths of two to three nucleosomes ($\sim$300 base pairs of DNA). The fragmented chromatin is subjected to immunoprecipitation using an antibody specific for the protein of interest. The data from this analysis are

reported as the number of times a specific sequence from a region of the genome was identified per million total bases of the immunoprecipitated DNA sequenced (Figure 8-12b). At divergently transcribed genes, such as the *Hsd17b12* gene encoding an enzyme involved in intermediary metabolism, two peaks of immunoprecipitated DNA were detected, corresponding to Pol II transcribing in the sense and antisense directions and then pausing. However, only Pol II transcribing in the sense direction was detected more than 1 kb from the start site (TSS). The number of counts per million from this region of the genome >1 kb from the TSS was very low because the gene is transcribed at low frequency. However, the number of counts per million at the transcription start site regions for both sense and antisense transcription was much higher, reflecting the fact that Pol II molecules had initiated transcription in both directions at this promoter, but paused before transcribing farther than 1 kb from the start sites in each direction. In contrast, the *Rpl6* gene, encoding a large ribosomal subunit protein that was abundantly transcribed in the proliferating mouse embryonic stem cells used in the study, was transcribed almost exclusively in the sense direction. The peak in counts per million less than 250 bp from the transcription start site again results from a long pause in transcription in the promoter-proximal region before the polymerase is released to transcribe into the gene. The number of sequence counts per million more than 1 kb downstream from the transcription start site in the sense direction was much higher than

for sense-direction transcription of the *Hsd17b12* gene, reflecting the high rate of transcription of the *Rpl6* gene.



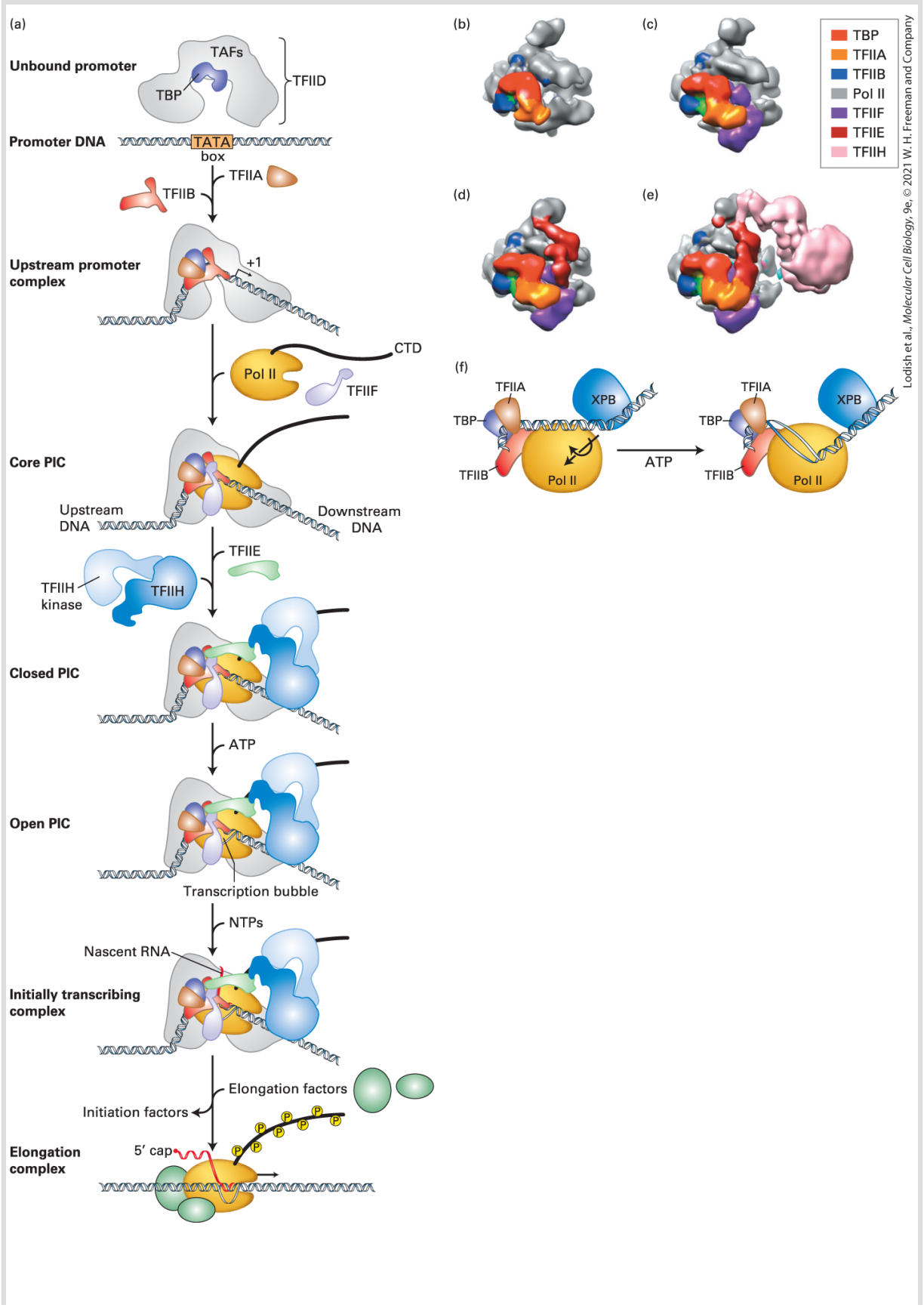


EXPERIMENTAL FIGURE 8-12 Chromatin immunoprecipitation technique localizes where a protein of interest associates with the genome. (a) Step **1**: Live cultured cells or tissues are incubated in 1 percent formaldehyde to covalently cross-link proteins to DNA and proteins to proteins. Step **2**: The preparation is then subjected to sonication to solubilize chromatin and shear it into fragments of 200–500 bp of DNA. Step **3**: An antibody to a protein of interest, here RNA polymerase II, is added, and DNA covalently linked to the protein of interest is immunoprecipitated. Step **4**: The covalent cross-linking is then reversed and the DNA is isolated. The isolated DNA can be analyzed by PCR with primers for a sequence of interest. Alternatively, total recovered DNA can be subjected to massively parallel DNA sequencing. See A. Hecht and M. Grunstein, 1999, *Method. Enzymol.* **304**:399. (b) Results from DNA sequencing of chromatin from mouse embryonic stem cells immunoprecipitated with antibody to RNA polymerase II are shown for a gene that is divergently transcribed (*left*) and a gene that is transcribed only in the sense direction (*right*). Data are plotted as the number of times a DNA sequence in a 50-bp interval was observed per million base pairs sequenced. The region encoding the $5'$ end of the gene is shown below, with exons shown as rectangles and introns as lines.

[Part (b) Data from P. B. Rahl et al., 2010, *Cell* **141**:432.]

General Transcription Factors Position RNA Polymerase II at Transcription Start Sites and Assist in Initiation

Initiation of transcription by RNA polymerase II requires several initiation factors. These initiation factors position Pol II molecules at transcription start sites and help to separate the DNA strands so that the template strand can enter the active site of the enzyme. They are called *general transcription factors* because they are required for transcription of most, if not all, genes transcribed by RNA polymerase II. These proteins are designated *TFIIA, TFIIB*, and so on, and most are multisubunit proteins. The largest is TFIID, which consists of a single 38-kDa **TATA box–binding protein (TBP)** and 13 TBP-associated factors (*TAFs*). General transcription factors with similar activities and homologous sequences are found in all eukaryotes. The complex of Pol II and its general transcription factors bound to a promoter and ready to initiate transcription is called a *preinitiation complex (PIC)*. Figure 8-13a summarizes the current model for the stepwise assembly of the Pol II transcription preinitiation complex on a promoter containing a TATA box. The model is based on cryoelectron microscopy structures of intermediates in the assembly (Figure 8-13b–e).



Lodish et al., *Molecular Cell Biology*, 9e. © 2021 W. H. Freeman and Company

FIGURE 8-13 Model for the sequential assembly of an RNA polymerase II preinitiation complex. (a) The indicated general transcription factors and purified RNA polymerase II (Pol II) bind sequentially to promoter DNA with a TATA box to form a preinitiation complex (PIC). ATP hydrolysis then provides the energy for the unwinding of DNA at the transcription start site by a TFIIH helicase subunit that pushes upstream DNA into the polymerase. The DNA is held in position in the PIC by binding of the TATA box by the TBP subunit of TFIID, and the resulting strain on the structure of the duplex DNA assists the N-terminal region of TFIIB and Pol II to melt the DNA at the transcription start site, forming the transcription bubble. As Pol II initiates transcription in the resulting open complex, the polymerase transcribes away from the promoter, its CTD becomes phosphorylated by the TFIIH kinase domain, and the general transcription factors dissociate from the promoter. Data from S. Sainsbury, C. Berrnecky, and P. Cramer, 2015, *Nat. Rev. Mol. Cell Biol.* **16**:129. (b–e) Cryoelectron microscopy structures of (b) TBP (red), TFIIA (orange), and TFIIB (blue) bound to a strong TATA box promoter. The DNA is not visible because it is obscured by TFIIA and TFIIB. (c) The core PIC composed of TBP-TFIIA-TFIIB-TFIIF-Pol2 bound to promoter DNA. (d) TFIIE bound to the core PIC; (e) TFIIH bound to the core PIC-TFIIE complex, generating the complete, closed preinitiation complex. (f) Diagram of how the TFIIH helicase subunit XPB pushes downstream DNA into the polymerase active site. This distorts the DNA structure because the TATA box is held by TBP, TFIIA, and TFIIB at a fixed distance. This assists the N-terminus of TFIIB and the polymerase in melting the DNA at the transcription start site and inserting the coding strand into the channel between RPB1 and RPB2 where it associates with the polymerase active site, resulting in melting of the transcription start site region, generating the “open complex.”

[Republished with permission of Nature Publishing Group, from Y. He et al., 2013, “Structural Visualization of Key Steps in Human Transcription Initiation.” *Nature* **495**:481–486; permission conveyed through Copyright Clearance Center, Inc.]

The TBP subunit of TFIID is the first protein to bind to a TATA box promoter. All eukaryotic TBPs analyzed to date have very similar C-terminal domains of 180 residues. This domain of TBP folds into a saddle-shaped structure; the two halves of the molecule exhibit an overall dyad symmetry but are not identical. TBP interacts with the

minor groove in DNA, bending the helix considerably (see Figure 5-5). The DNA-binding surface of TBP is conserved in all eukaryotes, explaining the high conservation of the TATA box promoter element (see Figure 8-11).

Once TBP has bound to the TATA box, TFIIA and TFIIB can bind (see Figure 8-13a). TFIIA is a heterotrimer larger than TBP, and TFIIB is a monomeric protein, slightly smaller than TBP. TFIIA associates with TBP and DNA on the upstream side of the TBP–TATA box complex. The C-terminal domain of TFIIB clamps onto the C-terminal "stirrup" of the saddle-shaped TBP molecule (see Figure 5-5) and contacts the major groove of DNA on either side of the TATA box. Next, a complex of Pol II and TFIIF, a heterodimer, associates with the promoter DNA-TFIIA-TFIIB complex forming a core *p*romoter *i*nitiation *c*omplex or *core PIC* (see Figure 8-13a). The extended N-terminal domain of TFIIB is inserted into the RNA exit channel of RNA polymerase II (see Figure 8-7c) where it interacts with the double-stranded DNA, stabilizing the complex and helping to hold the DNA in the region of the transcription start site (TSS) over the cleft between RPB1 and RPB2 when the clamp is open (see Figures 8-9a and 8-13a, c).

Next TFIIE, a heterodimer of two different subunits binds next to TIIF, completely enclosing the template DNA in a protein channel over the TSS region, and further stabilizing the complex with promoter DNA (see Figure 8-13d). TFIIE also contains a docking site

for TFIIH, another multisubunit factor nearly as large as the polymerase itself, containing 10 different subunits. Binding of TFIIH completes assembly of the transcription preinitiation complex (see Figures 8-13a closed PIC, and 8-13e).

The **helicase** activity of one of the core TFIIH subunits (XPB in humans, Ssl2 in yeast; see Figure 8-13f) uses energy from ATP hydrolysis to help unwind the DNA duplex at the start site, allowing Pol II to form an *open* complex in which the DNA duplex surrounding the start site is melted and the template strand is bound at the polymerase active site (see Figure 8-13a, Open PIC). As the polymerase transcribes away from the promoter region, the N-terminal domain of TFIIB is released from the RNA exit channel as the 5′ end of the nascent RNA enters it. Three TFIIH subunits form a kinase module (TFIIH in Figure 8-13a) that phosphorylates the Pol II CTD multiple times on serine 5 (underlined) of the Tyr-Ser-Pro-Thr-Ser-Pro-Ser repeat that constitutes the CTD. As we will discuss further in Chapter 9, when multiphosphorylated at Ser5, the CTD is a docking site for the enzymes that form the cap structure (see Figure 5-26) on the 5′ end of an RNA transcribed by RNA polymerase II. In the minimal in vitro transcription assay with TBP substituting for the full TFIID complex, TBP remains bound to the TATA box as the polymerase transcribes away from the promoter region, but the other general transcription factors, and probably the TFIID-TAF subunits, dissociate.

Remarkably, the first subunits of TFIIH to be cloned from humans were identified because mutations in them cause defects in the repair of DNA damaged by alkylation, such as a base with a covalently linked mutagen, or a UV-induced thymidine dimer (see Figure 5-17). In normal individuals, when a transcribing RNA polymerase becomes stalled at a region of damaged template DNA, the core TFIIH complex, lacking the three subunits of the kinase domain (see Figure 8-13a) but including the helicase subunit mentioned above plus an additional helicase subunit whose activity is not required for transcription initiation, recognizes the stalled polymerase and then associates with other proteins that function with TFIIH in repairing the damaged DNA region. In patients with mutant forms of these TFIIH subunits, such repair of damaged DNA in transcriptionally active genes is impaired. As a result, affected individuals have extreme skin sensitivity to sunlight (a common cause of DNA damage through the generation of thymidine dimers) and exhibit a high incidence of cancer. Consequently, these subunits of TFIIH serve two functions in the cell, one in the process of transcription initiation (where one of the helicase subunits helps to open the DNA strands at the transcription start site) and a second in the repair of DNA (where the activities of both helicase subunits are required; see Figure 5-18). Depending on the severity of the defect in TFIIH function, these individuals may suffer from diseases such as xeroderma pigmentosum (see Chapter 25) and Cockayne syndrome (see Chapter 5). ■

The TAF subunits of TFIID function in initiating transcription from promoters that lack a TATA box (>70 percent of human promoters, usually transcribed at lower level compared to TATA box promoters). For instance, some TAF subunits contact the initiator, and or DPE promoter elements (see Figure 8-11) in promoters where they occur (Figure 8-14). This probably explains how such sequences can replace a TATA box (see Figure 8-13) to allow TFIID binding to a TATA-minus promoter.

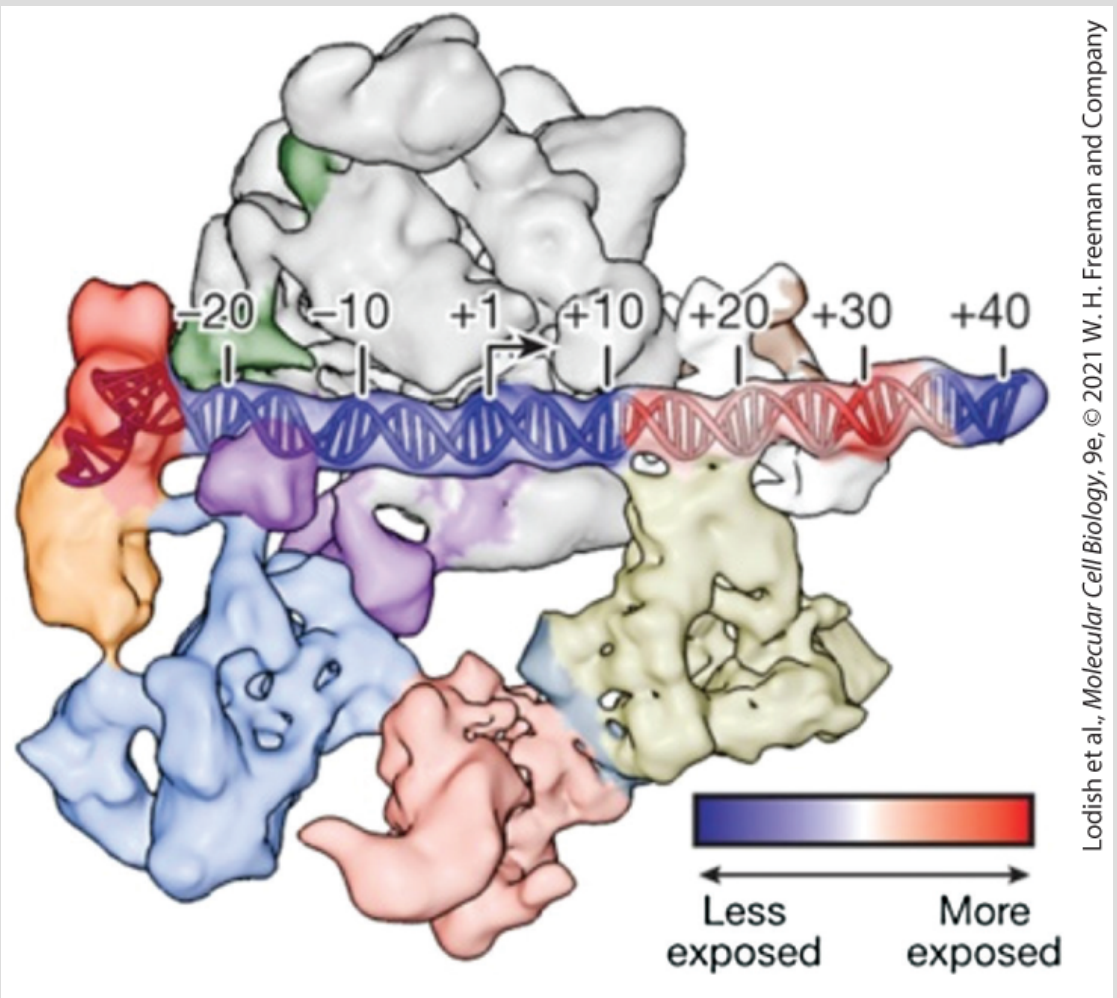


FIGURE 8-14 Cryoelectron microscopy structure of a closed complex of TFIID, TFIIA, TFIIB, and Pol II bound to promoter DNA. TBP is shown in red, TFIIA in

orange, TAF complexes in light blue, pink, and tan, and Pol II in gray. The color scale at the bottom indicates the sensitivity of the DNA in the complex to digestions by DNase I.

[Republished with permission of Nature Publishing Group, from R. Louder et al., 2016, "Structure of Promoter-Bound TFIID and Model of Human Pre-Initiation Complex Assembly." *Nature* **531**:604–609; https://doi.org/10.1038/nature17394; permission conveyed through Copyright Clearance Center, Inc.]

Chromatin immunoprecipitation assays (see Figure 8-12) using antibodies to TBP show that TBP binds in the region between the divergent transcription start sites in CpG island promoters. Consequently, the general transcription factors required for initiation from promoters containing a TATA box are probably also required for initiation from the weaker CpG island promoters. The absence of any of the promoter elements summarized in Figure 8-11 may account for the divergent transcription from multiple start sites observed from CpG island promoters, since cues from the DNA sequence are not present to correctly orient the preinitiation complex. TFIID and the other general transcription factors may instead choose among nearly equivalent, weak binding sites in CpG island promoters, which may contribute to the low frequency of transcription initiation.

Elongation Factors Regulate the Initial Stages of Transcription in the Promoter-Proximal Region

In metazoans, at most promoters, Pol II pauses after transcribing fewer than 100 nucleotides, due to the binding of a protein called *NELF* (*n*egative *el*ongation *f*actor). NELF binds to Pol II along with an elongation factor called *DSIF* (*D*RB *s*ensitivity-*i*nducing *f*actor, so named because an ATP analog called *DRB* inhibits further transcription elongation in its presence) (Figure 8-15a). The surfaces of Pol II bound by DSIF and NELF are covered by TFIIF and TFIIE in the preinitiation complex, but become exposed when polymerase initiates and separates from the general transcription factors. NELF appears to inhibit transcription elongation by blocking the principal channel through which NTPs reach the active site of the enzyme and by inhibiting conformational changes in the RPB1 and RPB2 subunits of Pol II required for translocation of the enzyme down the template.

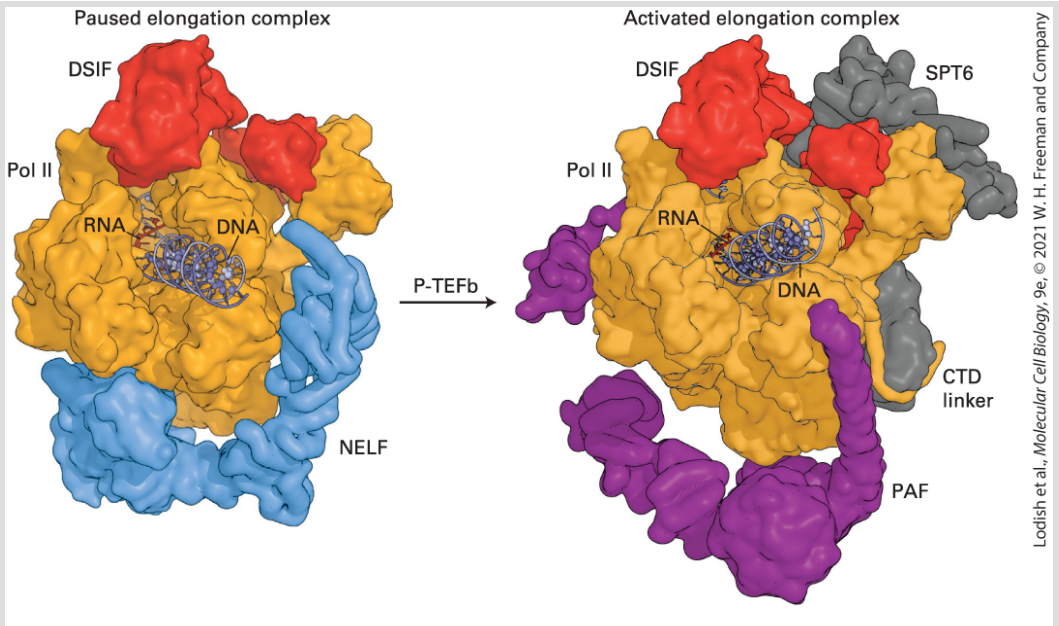


FIGURE 8-15 Switch from Pol II pausing to active elongation. The structures of the paused transcription elongation complex Pol II–DSIF–NELF and the

activated elongation complex Pol II–DSIF–PAF–SPT6. The view is along downstream DNA and looking into the polymerase cleft. The general elongation factor DSIF (red) is present in the paused and activated elongation complex, but changes its conformation between the two complexes. The negative elongation factor NELF (blue) stabilizes the paused state. In the activated elongation complex, the positive elongation factor SPT6 (gray) binds the phosphorylated linker to the Pol II CTD, and the PAF complex (purple) binds the polymerase funnel and competes with NELF.

[Republished with permission of Nature Publishing Group, from P. Cramer, 2019, "Organization and Regulation of Gene Transcription." *Nature* **573**:45–54; permission conveyed through Copyright Clearance Center, Inc.]

The inhibition of elongation that results from NELF/DSIF binding to Pol II is relieved when DSIF, NELF, and serine 2 of the Pol II CTD (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) are phosphorylated by the heterodimeric protein kinase CDK9–cyclin T, also called *P-TEFb*, which associates with the Pol II-NELF-DSIF complex, leading to phosphorylation of NELF and DSIF. Phosphorylated NELF dissociates from the complex, allowing two additional elongation factors to bind. These factors, *PAF* and *SPT6* (Figure 8-15b), block further binding of NELF, which allows the polymerase to resume elongation. Phosphorylated DSIF binds to the RPB1 clamp (see Figure 8-9) and RPB2 on the other side of the cleft between RPB1 and RPB2, helping to hold the clamp closed, which allows the polymerase to transcribe long distances without terminating and dissociating from the template. These same elongation factors also regulate transcription from CpG island promoters. So we see that factors that regulate elongation in the promoter-proximal region

provide another mechanism for controlling gene transcription in addition to the many factors that regulate transcription initiation.

Transcription of HIV (human immunodeficiency virus), the cause of AIDS, is dependent on the activation of CDK9–cyclin T by a small viral protein called **Tat**. Cells experimentally infected with *tat*$^-$ mutants produce short viral transcripts about 50 nucleotides long. In contrast, cells infected with wild-type HIV synthesize many more long viral transcripts that encompass the entire integrated proviral genome (see Figure 8-16; also Figure 5-43). Thus Tat functions as an *anti-pausing factor*, permitting RNA polymerase II to read through the transcriptional block caused by NELF binding. (Tat is initially made by rare transcripts that fail to be arrested by NELF when the HIV promoter is transcribed at a high rate in activated T-lymphocytes; see Chapter 24.) Tat is a sequence-specific RNA-binding protein. It binds to the RNA copy of a sequence called *TAR*, which forms a stem-loop structure near the 5′ end of the HIV transcript (Figure 8-16). TAR also binds cyclin T, holding the cyclin T-CDK9 complex close to the polymerase, where it efficiently phosphorylates its substrates, resulting in release of NELF and transcription elongation. Chromatin immunoprecipitation assays done after treating cells with specific inhibitors of CDK9 indicate that transcription of some 30 percent of mammalian genes is regulated by controlling the activity of cyclin T-CDK9 (P-TEFb). This regulation is probably due to interactions between P-TEFb and

sequence-specific DNA-binding transcription factors, rather than by an RNA-binding protein, as in the case of HIV Tat. ■

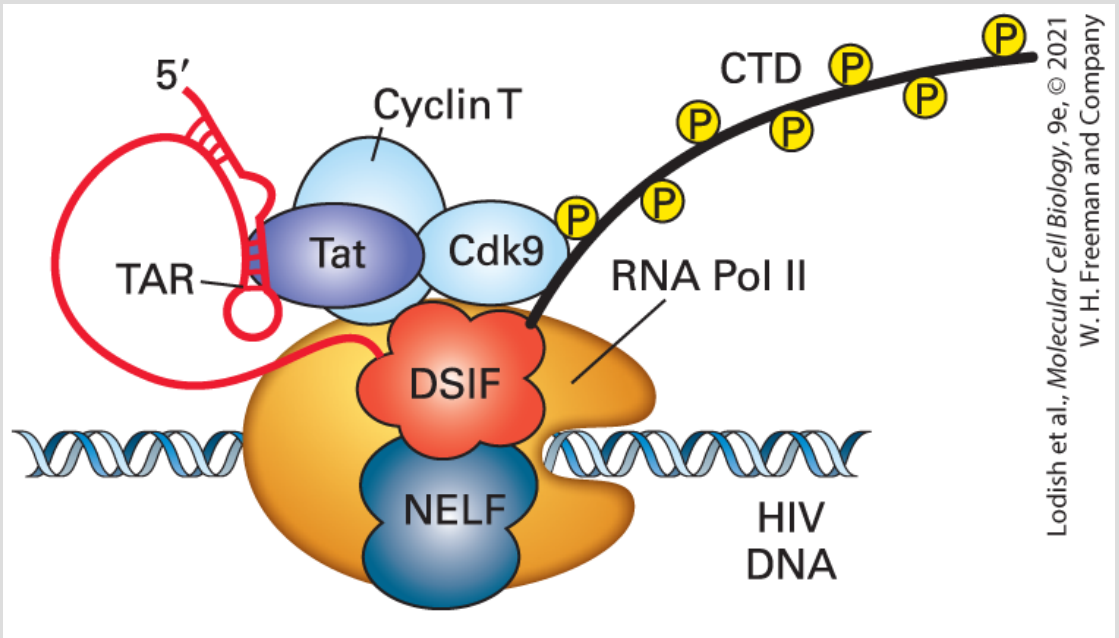


FIGURE 8-16 Model of antitermination complex composed of HIV Tat protein and several cellular proteins. The TAR element in the HIV transcript contains sequences recognized by Tat and the cellular protein cyclin T. Cyclin T activates and helps position the protein kinase CDK9 near its substrate, the CTD of RNA polymerase II. CTD phosphorylation at serine 2 of the Pol II CTD heptad repeat is required for transcription elongation. Cellular proteins DSIF and the NELF complex are also involved in regulating Pol II elongation, as discussed in the text. See T. Wada et al., 1998, *Genes Dev.* **12**:343; Y. Yamaguchi et al., 1999, *Cell* **97**:451; and T. Yamada et al., 2006, *Mol. Cell* **21**:227.

KEY CONCEPTS OF SECTION 8.2

RNA Polymerase II Promoters and General Transcription Factors

- RNA polymerase II initiates transcription of genes at the nucleotide in the DNA template that corresponds to the $5'$ nucleotide that is capped in the encoded mRNA.
- Four principal types of promoter sequences have been identified in eukaryotic DNA. The TATA box is prevalent in highly transcribed genes. Initiator promoters are found in some genes, and many strong promoters contain a downstream promoter element (DPE) near +30 and a BRE near −35 required for the highest rates of transcription (see Fig 8-11). CpG islands, the promoters for about 70 percent of protein-coding genes in vertebrates, are characteristic of genes transcribed at a low rate.
- Transcription by Pol II of genes with a TATA box promoter is initiated by sequential binding of the following in the indicated order: TFIID, which contains the TBP subunit that binds to TATA box DNA; TFIIA and TFIIB; a complex of Pol II and TFIIF; TFIIE; and finally, TFIIH (see Figure 8-13).
- The helicase activity of a TFIIH subunit helps to separate the DNA strands at the transcription start site in most promoters, a process that requires hydrolysis of ATP. As Pol II begins transcribing away from the start site, its CTD is phosphorylated on serine 5 by the TFIIH kinase domain.
- In metazoans, NELF and DSIF associate with Pol II after initiation, inhibiting elongation to fewer than 100 bp from the transcription start site. Inhibition of elongation is relieved when cyclin T-CDK9 (also called *P-TEFb*) associates with the paused elongation complex and CDK9 phosphorylates subunits of NELF, DSIF, and serine 2 of the Pol II CTD, causing dissociation of NELF from the elongating complex.
- Elongation factors PAF and SPT6 bind to the Pol II surfaces bound by NELF, preventing further NELF binding, and preventing opening of the cleft between RPB1 and RBP2, assisting in long distance processive elongation.

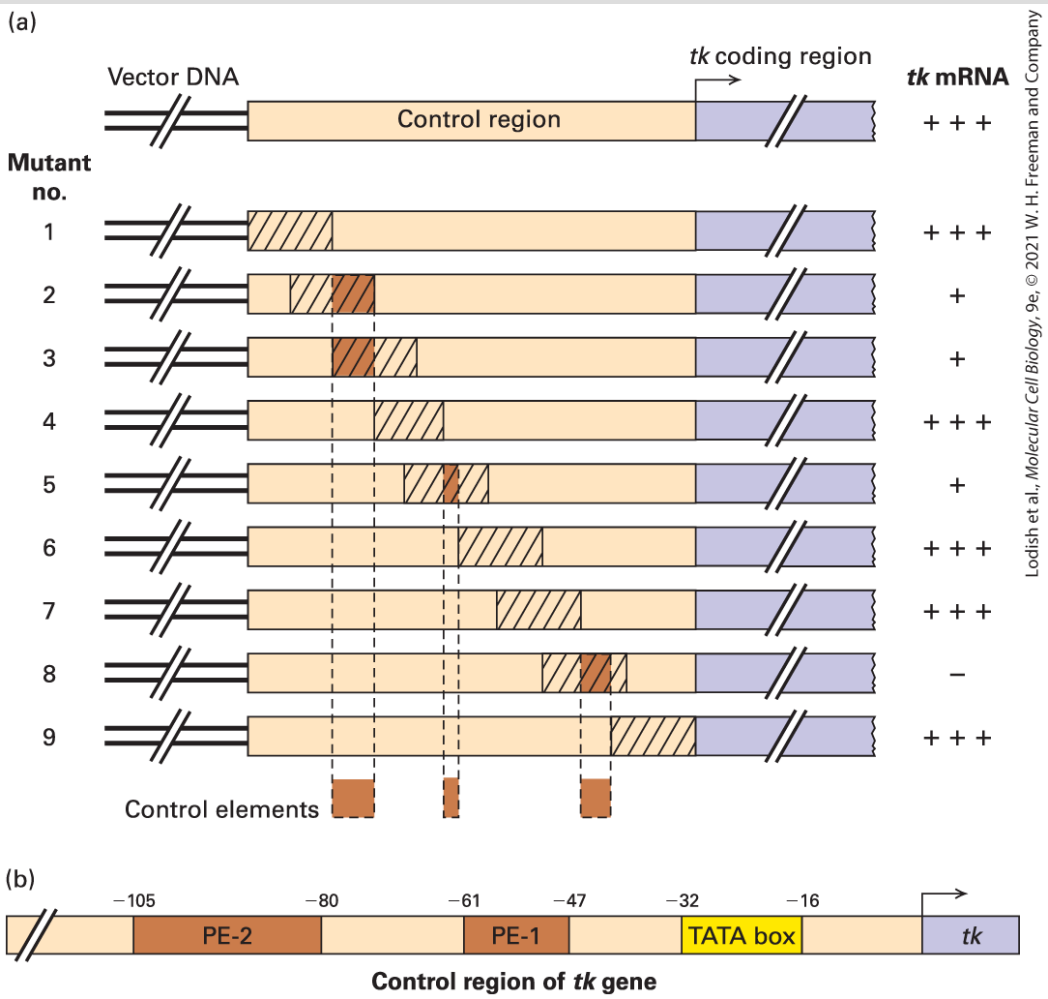
8.3 Regulatory Sequences for Protein-Coding Genes and the Proteins Through Which They Function

As we have seen, expression of eukaryotic protein-coding genes is regulated by multiple DNA sequences, generically referred to as *transcription-control regions*. These regions include promoters and other types of control elements located near transcription start sites, as well as sequences located far from the genes they regulate. In this section, we take a closer look at the properties of various control elements found in eukaryotic protein-coding genes and the proteins that bind to them.

Promoter-Proximal Elements Help Regulate Eukaryotic Genes

Recombinant DNA techniques have been used to systematically mutate the nucleotide sequences of various eukaryotic genes in order to identify transcription-control regions. The use of *linker scanning mutagenesis*, for example, can pinpoint the sequences within a regulatory region that function to control transcription. In this approach, a set of constructs with contiguous overlapping mutations are assayed for their effect on expression of a reporter

gene or production of a specific mRNA (Figure 8-17a). This type of analysis identified **promoter-proximal elements** of the thymidine kinase (*tk*) gene from herpes simplex type I virus (HSV-I). The results demonstrated that the DNA region upstream of the HSV-I *tk* gene contains three separate transcription-control sequences: a TATA box in the interval from −32 to −16 and two other control elements farther upstream (Figure 8-17b). Experiments using mutants containing single-base-pair changes in promoter-proximal control elements revealed that these elements are generally about 6–12 bp long. In human genes, promoter-proximal elements are usually defined as being within ∼200 bp of a transcription start site (TSS). In vertebrates, promoter-proximal elements are found both upstream and downstream of the TSS at equal frequency. While, strictly speaking, the term *promoter* refers to the DNA sequence that determines where a polymerase initiates transcription, the term is often used to refer to both a promoter and its associated promoter-proximal control elements.



EXPERIMENTAL FIGURE 8-17 Linker scanning mutations identify transcription-control elements. (a) In linker scanning mutagenesis, a region of eukaryotic DNA (tan) that supports high-level expression of a reporter gene (light purple) is cloned in a plasmid vector as diagrammed at the top. Overlapping linker scanning (LS) mutations (crosshatched areas) are introduced from one end of the region being analyzed to the other. These mutations are created by scrambling the nucleotide sequence in a short stretch of the DNA. After the mutant plasmids are transfected separately into cultured cells, the activity of the reporter-gene product is assayed. In the example shown here, the sequence from −120 to +1 of the herpes simplex virus thymidine kinase gene, LS mutations 1, 4, 6, 7, and 9 have little or no effect on expression of the reporter gene, indicating that the regions altered in these mutants contain no control elements. Reporter-gene expression is significantly reduced in mutants 2, 3, 5, and 8, indicating that control elements (brown) lie in the

intervals shown at the bottom. (b) Analysis of these LS mutations identified a TATA box and two promoter-proximal elements (PE-1 and PE-2). See S. L. McKnight and R. Kingsbury, 1982, *Science* **217**:316.

In bacteria, regulation of transcription by a control element close to a promoter generally requires its precise localization relative to the transcription start site (TSS). For example, stimulation of *E. coli lac* operon transcription by CAP protein (cyclic AMP-regulated activator protein) is completely lost when two base pairs are inserted between the CAP protein binding site and the TSS. Extensive experimental analyses in *E. coli* have shown that this is because transcriptional activation by CAP requires highly specific direct interactions between CAP bound to its specific binding site in DNA and *E. coli* RNA polymerase bound to the *lac* promoter. Insertion of two base pairs between their binding sites rotates the CAP site 72° relative to the polymerase and moves it ~70 Å away, preventing the interactions between CAP and RNA polymerase required to stimulate transcription.

To test the spacing constraints on control elements in the HSV-I *tk* promoter region identified by analysis of linker scanning mutations, researchers prepared and assayed constructs containing small deletions and insertions between the elements. Changes in spacing between the promoter and promoter-proximal control elements of 20 bp or fewer had little effect. However, insertions of 30–50 bp between a HSV-I *tk* promoter-proximal element and the TATA box

was equivalent to deleting the element. Similar analyses of other eukaryotic promoters have also indicated that considerable flexibility in the spacing between promoter-proximal elements is generally tolerated, but that separations of several tens of base pairs may decrease transcription.

Distant Enhancers Often Stimulate Transcription by RNA Polymerase II

Transcription from many eukaryotic promoters can also be stimulated by control elements located thousands of base pairs away from the transcription start site. Such long-distance transcription-control elements, referred to as **enhancers**, are common in eukaryotic genomes but fairly rare in bacterial genomes. Procedures such as linker scanning mutagenesis have indicated that enhancers, usually on the order of 200 bp long, are, like promoter-proximal elements, composed of several functional sequence elements of about 6–12 bp each. As we will see, each of these regulatory elements is a binding site for a sequence-specific DNA-binding transcription factor.

Analyses of many different metazoan enhancers have shown that they can occur with equal probability upstream from a promoter or downstream from a promoter, where they are within an intron, or even downstream from the final exon of a gene, as in the case of the

SALL1 gene (see Figure 8-5a). Many enhancers are cell-type-specific. For example, an enhancer controlling *Pax6* expression in the retina is found in the intron between exons 4 and 5 (see Figure 8-4a), whereas an enhancer controlling *Pax6* expression in the hormone-secreting cells of the pancreas is located in a roughly 200-bp region upstream of exon 0 (so named because it was discovered after the exon called "exon 1").

Most Eukaryotic Genes Are Regulated by Multiple Transcription-Control Elements

Initially, enhancers and promoter-proximal elements were thought to be distinct types of transcription-control elements. However, as more enhancers and promoter-proximal elements were analyzed, the distinctions between them became less clear. For example, both types of elements can generally stimulate transcription even when inverted, and both types are often cell-type-specific. The consensus now is that a spectrum of control elements regulates transcription by RNA polymerase II. At one extreme are enhancers, which can stimulate transcription from a promoter tens of thousands of base pairs away. At the other extreme are promoter-proximal elements, such as the upstream elements controlling the HSV-I *tk* gene, which lose their influence when moved 30–50 bp farther from the promoter. Researchers have identified a large number of

transcription-control elements that can stimulate transcription from distances between these two extremes.

Figure 8-18a summarizes the locations of transcription-control sequences for a hypothetical mammalian gene with a promoter containing a TATA box. The transcription start site encodes the first (5′) nucleotide of the first exon of an mRNA, the nucleotide that is capped. In addition to the TATA box at about −31 to −26, promoter-proximal elements, which are relatively short (∼6–10 bp), are located within the first 200 bp either upstream or downstream of the start site. Enhancers, in contrast, are usually about 50–200 bp long and are composed of multiple elements of about 6–10 bp. Enhancers may be located up to 50 kb or more upstream or downstream from the start site or within an intron. Like the *Pax6* gene, many mammalian genes are controlled by multiple enhancer regions that function in different types of cells.

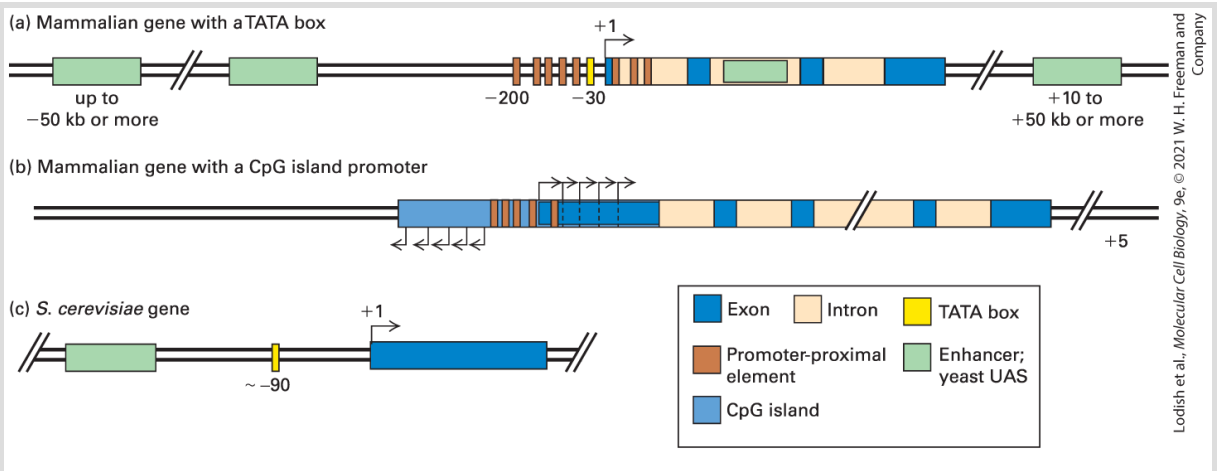


FIGURE 8-18 General organization of control elements that regulate gene expression in multicellular eukaryotes and yeast. (a) Mammalian genes with a TATA box promoter are

regulated by promoter-proximal elements and enhancers. The promoter elements shown in Figure 8-11 position RNA polymerase II to initiate transcription at the start site and influence the rate of transcription. Enhancers may be either upstream or downstream and as far away as hundreds of kilobases from the transcription start site. In some cases, enhancers lie within introns. Promoter-proximal elements are found upstream and downstream of transcription start sites at equal frequency in mammalian genes. (b) For mammalian genes with a CpG island promoter, transcription initiates at several sites in both the sense and antisense directions from the ends of the CpG-rich region. Transcripts in the sense direction are elongated and are processed into mRNAs by RNA splicing. These genes express mRNAs with alternative 5′ exons determined by the transcription start site. Genes with CpG island promoters contain promoter-proximal control elements. Currently, it is not clear whether they are also regulated by distant enhancers. (c) Most *S. cerevisiae* genes contain only one regulatory region, called an *upstream activating sequence (UAS)*, and a TATA box, which is about 90 bp upstream from the transcription start site.

Figure 8-18b summarizes the promoter region of a mammalian gene with a CpG island promoter. About 70 percent of mammalian genes are expressed from CpG island promoters, usually at much lower levels than genes with TATA box promoters. Multiple alternative transcription start sites are used, generating mRNAs with alternative 5′ ends for the first exon derived from each start site. Transcription occurs in both directions, but Pol II molecules transcribing in the sense direction are elongated to 3 kb or more much more efficiently than transcripts in the antisense direction.

In the model organism *S. cerevisiae* (budding yeast), genes are closely spaced (see Figure 7-4c), and few genes contain introns. In this organism, enhancers, which are referred to as **upstream activating sequences (UASs)**, usually lie within 200 bp upstream of the

promoters of the genes they regulate. Most yeast genes contain only one UAS. In addition, *S. cerevisiae* genes contain a TATA box about 90 bp upstream from the transcription start site (Figure 8-18c).

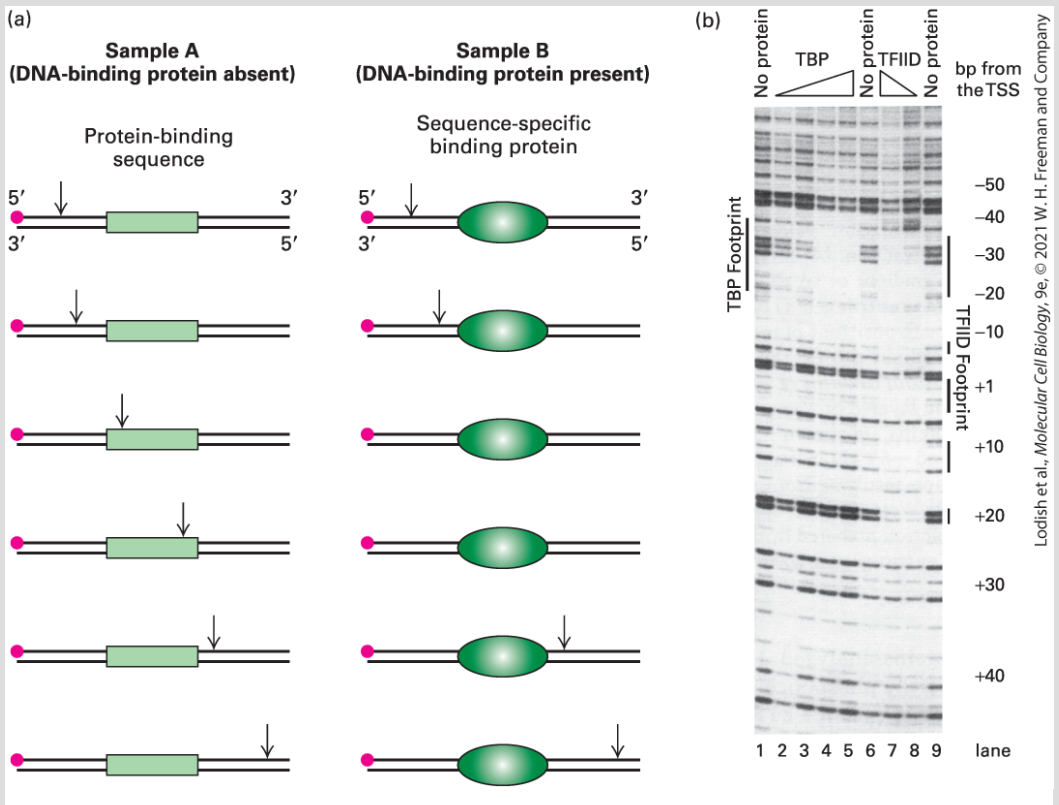
DNase I Footprinting and EMSA Detect Protein-DNA Interactions

The various transcription-control elements found in eukaryotic DNA are binding sites for regulatory proteins called *transcription factors*. The simplest eukaryotic cells encode hundreds of transcription factors, and the human genome encodes at least 1600. The transcription of each gene in the genome is independently regulated by *combinations of specific transcription factors* that bind to the gene's transcription-control regions. The number of possible combinations of this many transcription factors is astronomical, sufficient to generate unique controls for every gene encoded in the genome.

In yeast, *Drosophila*, and other genetically tractable eukaryotes, numerous genes encoding transcription activators and repressors have been identified by classical genetic analyses like those described in Chapter 6. However, in mammals and other vertebrates, which are less amenable to such genetic analysis, most transcription factors have been detected initially and subsequently purified by biochemical techniques. In this approach, a DNA-regulatory element that has been identified by the kinds of

mutational analyses described above is used to identify *cognate* proteins — those proteins that bind specifically to it. Two common techniques for detecting such cognate proteins are DNase I footprinting and the electrophoretic mobility shift assay.

DNase I footprinting takes advantage of the fact that when a protein is bound to a region of DNA, it protects that DNA sequence from digestion by nucleases. As illustrated in Figure 8-19, samples of a DNA fragment that has been labeled with a radioactive molecule at one end of one strand are digested under carefully controlled conditions in the presence and absence of a DNA-binding protein. Then the DNA is purified away from all of the protein, denatured and electrophoresed, and the resulting gel is subjected to autoradiography. The region protected by the bound protein appears as a gap, or *footprint*, in the array of bands resulting from digestion in the absence of the protein. When footprinting is performed with a DNA fragment containing a known transcription-control element, the appearance of a footprint indicates the presence of a transcription factor that binds that control element in the protein sample being assayed. Footprinting also identifies the specific DNA sequence to which the transcription factor binds.



EXPERIMENTAL FIGURE 8-19 DNase I footprinting reveals the region of a DNA sequence where a transcription factor binds. (a) A DNA fragment known to contain a transcription-control element is labeled at one end with $^{32}\mathrm{P}$ (red dot). Portions of the labeled DNA sample are then digested with DNase I in the presence and in the absence of protein samples containing a sequence-specific DNA-binding protein. DNase I hydrolyzes the phosphodiester bonds of DNA between the $3'$ oxygen on the deoxyribose of one nucleotide and the $5'$ phosphate of the next nucleotide. A low concentration of DNase I is used so that, on average, each DNA molecule is cleaved just once (vertical arrows). If the protein sample does not contain a protein that binds to a specific sequence in the labeled DNA, the DNA fragment is cleaved at multiple positions between the labeled and unlabeled ends of the original fragment, as in sample A (*left*). If the protein sample does contain such a protein, as in sample B (*right*), the protein binds to its cognate sequence in the DNA, thereby protecting a portion of the fragment from digestion. Following DNase treatment, the DNA is separated from protein, denatured to separate the strands, and electrophoresed. Autoradiography of the resulting gel detects only labeled strands and reveals fragments extending from the labeled end to the site of

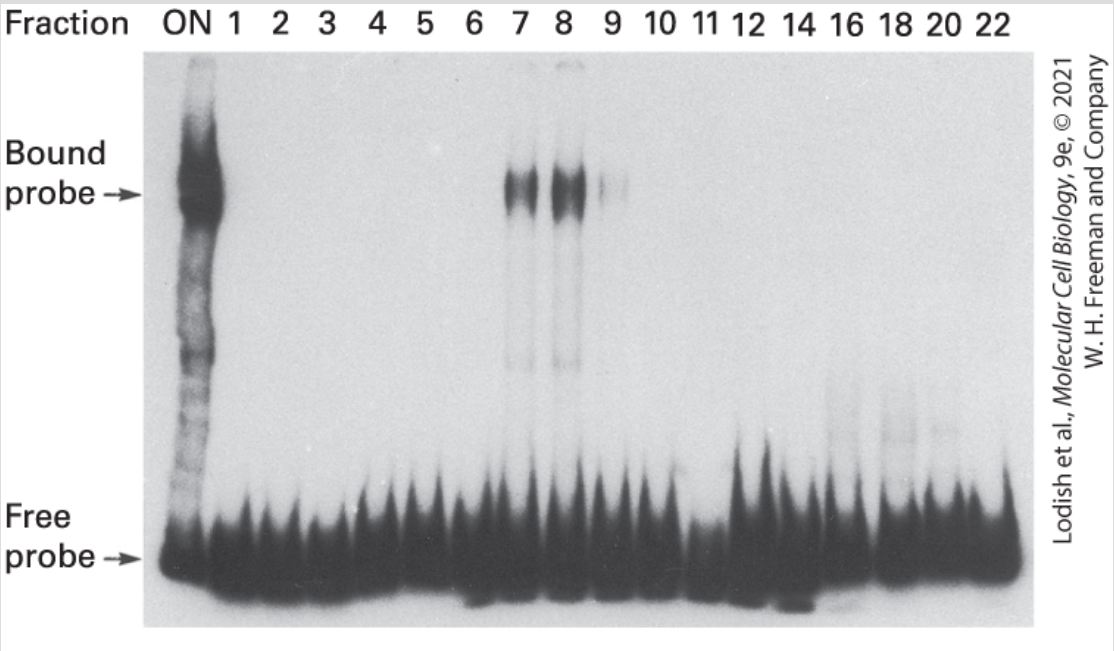
cleavage by DNase I. Cleavage fragments containing the transcription-control element show up on the gel for sample A but are missing in sample B because the bound cognate protein has blocked cleavages within that sequence and thus production of the corresponding fragments. The missing bands on the gel constitute the footprint. (b) Footprints produced by increasing amounts of TBP (indicated by the triangle) and of TFIID on the strong adenovirus major late promoter.

[Part (b) Republished with permission from Cold Spring Harbor Laboratory Press, from Q. Zhou et al., 1992, "Holo-TFIID Supports Transcriptional Stimulation by Diverse Activators and from a TATA-Less Promoter," *Genes Dev.* **6**(10):1964–1974.]

For example, DNase I footprinting of the strong adenovirus late promoter shows a protected region over the TATA box when TBP is added to the labeled DNA before DNase I digestion (Figure 8-19b). DNase I does not digest all phosphodiester bonds in a duplex DNA at equal rate. Consequently, in the absence of added protein (lanes 1, 6, and 9), a particular pattern of bands is observed that depends on the DNA sequence and results from cleavage at some phosphodiester bonds and not others. However, when increasing amounts of TBP are incubated with the end-labeled DNA before digestion with DNase I, TBP binds to the TATA box, and when sufficient TBP is added to bind all the labeled DNA molecules, it protects the region between about −35 and −20 from digestion (lanes 2–5). In contrast, increasing amounts of TFIID (lanes 7 and 8) protect not only the TATA box region but also regions near −7 to +5, +10 to +15, and +20 from digestion, producing a footprint different from TBP's. Results such as this tell us that other subunits of TFIID (the TBP-associated

factors, or TAFs) also bind to the DNA in the region downstream from the TATA box.

The **electrophoretic mobility shift assay (EMSA)**, also called the *gel-shift* or *band-shift assay*, is more useful than the footprinting assay for quantitative analysis of DNA-binding proteins. In general, the electrophoretic mobility of a DNA fragment is reduced when it is complexed with protein, causing a shift in the location of the fragment band. EMSA can be used to detect a transcription factor in protein fractions incubated with a radiolabeled DNA fragment (the probe) containing a known control element (Figure 8-20). As the amount of transcription factor added to the binding reaction is increased, the more labeled probe is shifted to the position of the DNA-protein complex. The results indicate that the concentration of this specific DNA-binding protein was highest in fraction 8, lower in fraction 7, still lower in fraction 9, and undetectable by this assay in the other fractions eluted from the column.



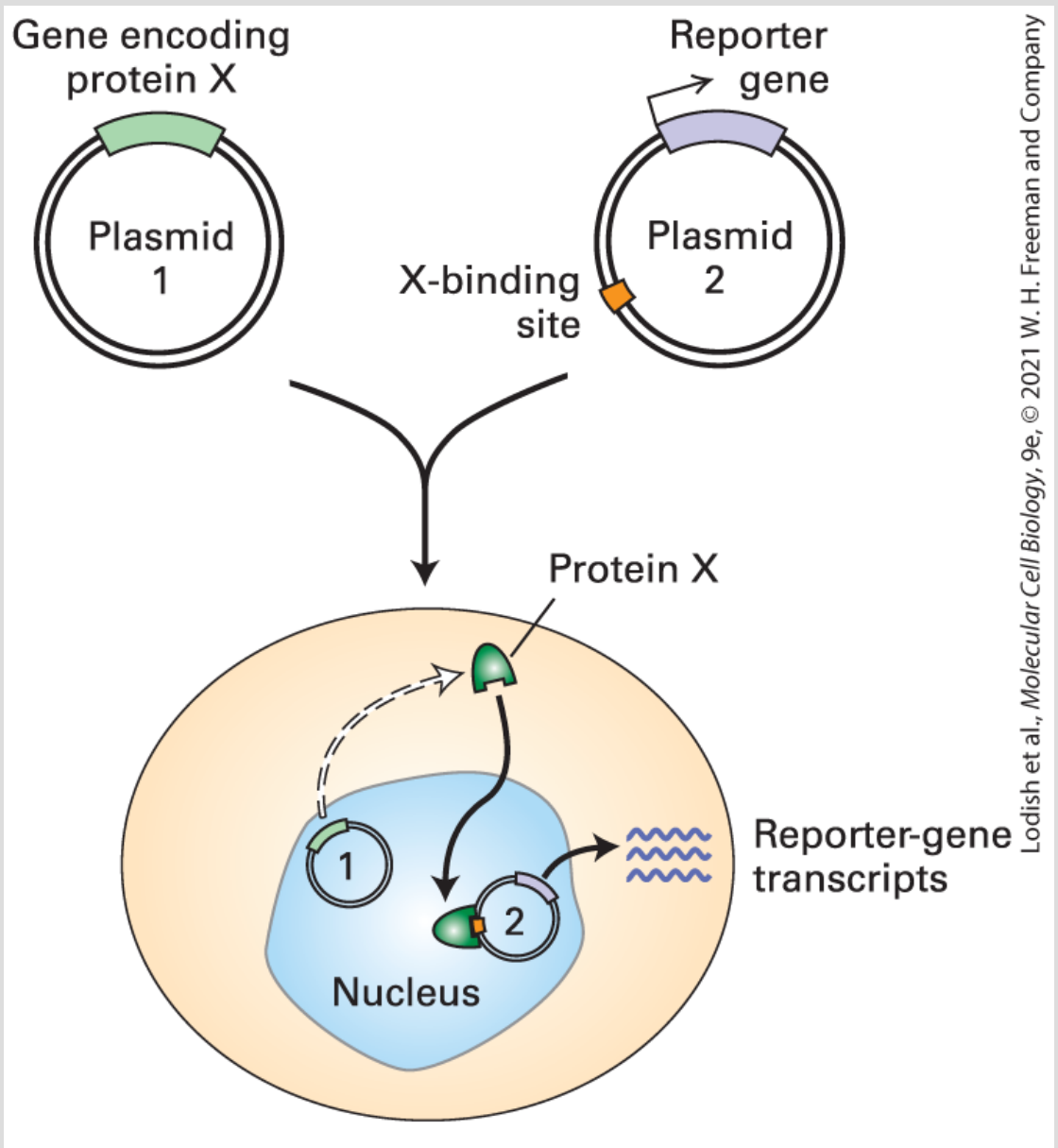
EXPERIMENTAL FIGURE 8-20 The electrophoretic mobility shift assay can be used to detect transcription factors during purification. In this example, protein fractions separated by column chromatography were assayed for their ability to bind to a radiolabeled DNA-fragment probe containing a known regulatory element. An aliquot of the protein sample that had been loaded onto the column (ON) and successive column fractions (numbers) were incubated with the labeled probe. The samples were then electrophoresed under conditions that do not disrupt protein-DNA interactions. The free probe not bound to protein migrated to the bottom of the gel. A protein in fractions 7 and 8 bound to the probe (as did protein in the unfractionated sample in column ON), forming a DNA-protein complex that migrated more slowly than the free probe. These fractions are therefore likely to contain the regulatory protein being sought.

[Republished with permission of *American Society for Biochemistry and Molecular Biology*, from S. Yoshinaga et al., 1989, "Purification and Characterization of Transcription Factor IIIC2," *J. Biol. Chem*. **264**:10726; permission conveyed through Copyright Clearance Center, Inc.]

In the biochemical isolation of a transcription factor, an extract of cell nuclei is commonly subjected sequentially to several types of

liquid chromatography (see Chapter 3). Fractions eluted from the columns are assayed by DNase I footprinting or EMSA using DNA fragments containing an identified regulatory element (see Figure 8-17). Fractions containing a protein that binds to the regulatory element in these assays contain a putative transcription factor. A powerful technique that is commonly used for the final step in purifying transcription factors is *sequence-specific DNA affinity chromatography*, a particular type of affinity chromatography in which long DNA strands containing multiple copies of the transcription-factor-binding site are coupled to a column matrix.

Once a transcription factor has been isolated and purified, its partial amino acid sequence can be determined and used to clone the gene or cDNA encoding it, as outlined in Chapter 6. The isolated gene can then be used to test the ability of the encoded protein to activate or repress transcription in an in vivo transfection assay (Figure 8-21).



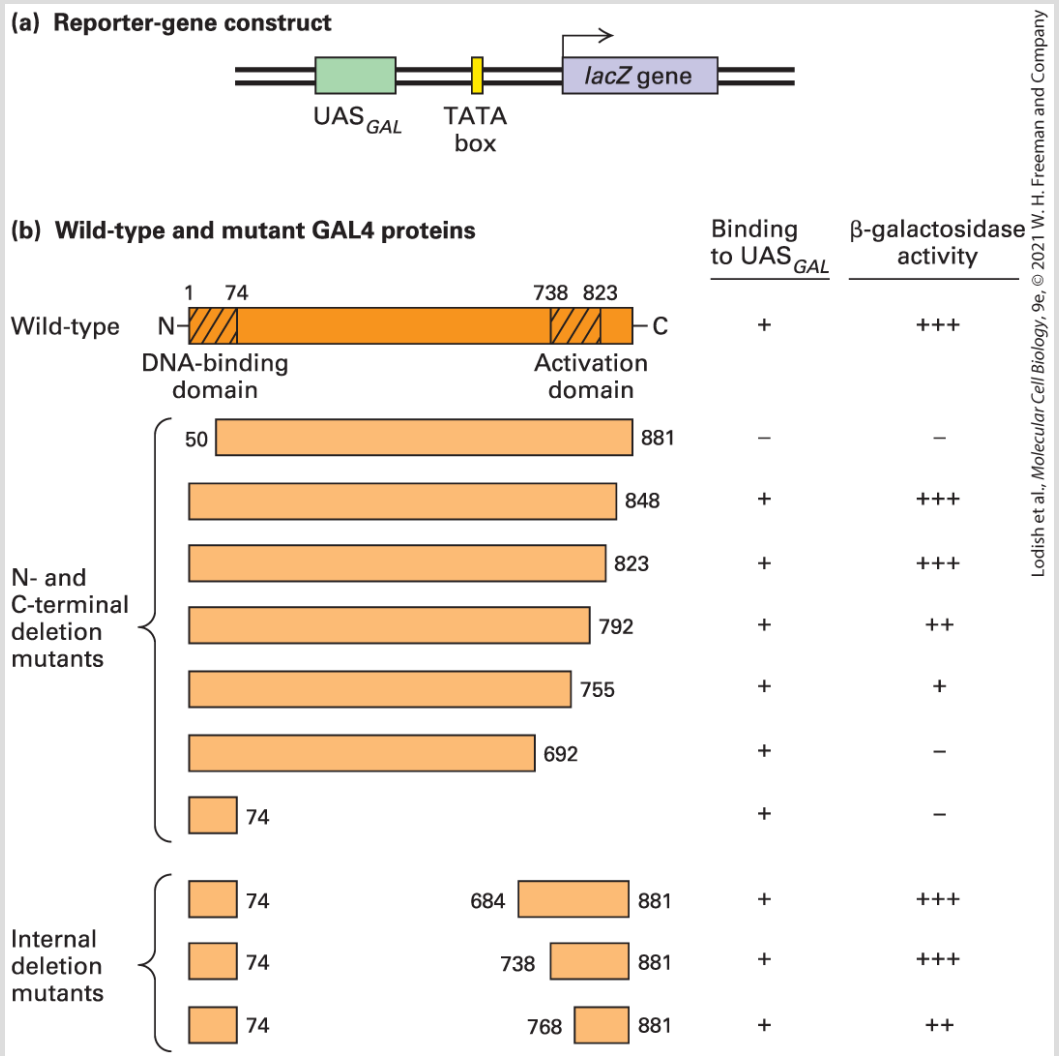
EXPERIMENTAL FIGURE 8-21 An in vivo transfection assay measures transcription activity to evaluate proteins believed to be transcription factors. The assay system requires two plasmids. One plasmid contains the gene encoding the putative transcription factor (protein X). The second plasmid contains a reporter gene (e.g., luciferase) and one or more binding sites for protein X. Both plasmids are simultaneously introduced into cells that lack the gene-encoding protein X. The production of reporter-gene RNA transcripts is measured; alternatively, the activity of the encoded protein can be assayed. If reporter-gene transcription is greater in the presence of the X-encoding plasmid than in its absence, then the protein is an activator; if transcription is less, then it is a repressor. By use of plasmids encoding a

mutated or rearranged transcription factor, important domains of the protein can be identified.

Activators Are Composed of Distinct Functional Domains

Studies with a yeast transcription activator called *Gal4* provided early insight into the domain structure of transcription factors. The gene encoding Gal4, which promotes expression of enzymes needed to metabolize galactose, was identified by complementation analysis of *gal4* mutants that cannot form colonies on an agar medium in which galactose is the only source of carbon and energy (see Chapter 6). Directed mutagenesis studies, like those described previously, identified UASs for the genes activated by Gal4 (see Figures 8-17). Each of these UASs was found to contain one or more copies of a 17-bp sequence called UAS_{GAL}. DNase I footprinting assays with recombinant Gal4 protein produced in *E. coli* from the yeast *GAL4* gene showed that Gal4 binds to UAS_{GAL} sequences. When a copy of UAS_{GAL} was cloned upstream of a TATA box followed by a β-galactosidase reporter gene, and that construct was introduced into yeast cells, expression of β-galactosidase was activated in galactose media in wild-type cells, but not in *gal4* mutants. These results showed that UAS_{GAL} is a transcription-control element activated by the Gal4 transcription factor in galactose media.

A remarkable set of experiments demonstrated that the Gal4 transcription factor is composed of separable functional domains: an N-terminal **DNA-binding domain**, which binds to specific DNA sequences, and a C-terminal **activation domain**, which interacts with other proteins to stimulate transcription from a nearby promoter (Figure 8-22). When the N-terminal DNA-binding domain of Gal4 was fused directly to various portions of its own C-terminal region, deleting internal sequences, the resulting truncated proteins retained the ability to stimulate expression of a reporter gene in an in vivo assay like that depicted in Figure 8-21. Thus the internal portion of the protein is not required for the functioning of Gal4 as a transcription factor. Similar experiments with another yeast activator, Gcn4, which regulates genes required for the synthesis of many amino acids, indicated that it contains a roughly 50–amino acid DNA-binding domain at its C-terminus and a roughly 20–amino acid activation domain near the middle of its sequence.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 8-22 Deletion mutants of the *GAL4* gene in yeast with a UAS_{GAL} reporter-gene construct demonstrate the separate functional domains in a transcription activator. (a) Diagram of DNA construct containing a *lacZ* reporter gene (encoding β-galactosidase) and TATA box ligated to UAS_{GAL}, a regulatory element that contains several Gal4-binding sites. The reporter-gene construct and DNA encoding wild-type or mutant (deleted) Gal4 were simultaneously introduced into mutant (*gal4*) yeast cells, and the activity of β-galactosidase expressed from *lacZ* was assayed. Activity should be high if the introduced *GAL4* DNA encodes a functional protein. (b) Schematic diagrams of wild-type Gal4 and various mutant forms. Small numbers refer to positions in the wild-type sequence. Deletion of 50 amino acids from the N-terminal end destroyed the ability of Gal4 to bind to UAS_{GAL} and to stimulate expression of β-galactosidase from the reporter gene. Proteins with extensive deletions from

the C-terminal end still bound to UAS_{GAL}. These results localize the DNA-binding domain to the N-terminal end of Gal4. The ability to activate β-galactosidase expression was not entirely eliminated unless somewhere between 126 and 189 or more amino acids were deleted from the C-terminal end. Thus the activation domain lies in the C-terminal region of Gal4. Proteins with internal deletions (*bottom*) were also able to stimulate expression of β-galactosidase, indicating that the central region of Gal4 is not crucial for its function in this assay. See J. Ma and M. Ptashne, 1987, *Cell* **48**:847; I. A. Hope and K. Struhl, 1986, *Cell* **46**:885; and R. Brent and M. Ptashne, 1985, *Cell* **43**:729.

Further evidence for the existence of distinct activation domains in Gal4 and Gcn4 came from experiments in which their activation domains were fused to a DNA-binding domain from an entirely unrelated *E. coli* DNA-binding protein. When these fusion proteins were assayed in vivo, they activated transcription of a reporter gene containing the cognate site for the *E. coli* protein. Thus functional transcription factors can be constructed from novel combinations of bacterial and eukaryotic protein domains.

Studies such as these have now been carried out with many eukaryotic transcription factors. The structural model of eukaryotic activators that has emerged from these studies is a modular one. Specifically, one or more activation domains are connected to a sequence-specific DNA-binding domain by flexible, intrinsically disordered protein domains (Figure 8-23). In some cases, amino acids included in the DNA-binding domain also contribute to transcriptional activation. As discussed in a later section, activation domains function by binding other proteins involved in both the

control of chromatin condensation and transcription. The presence of flexible, intrinsically disordered protein domains (see Figure 3-13) connecting the DNA-binding domain to the activation domains may explain why alterations in the spacing between control elements are so well tolerated in eukaryotic control regions. Thus even when the positions of transcription factors bound to DNA are shifted relative to each other, their activation domains may still be able to interact because they are attached to their DNA-binding domains through flexible protein regions. Also, as discussed in Section 8.4, some intrinsically disordered regions in transcription factors and other proteins involved in transcription interact through multiple low-affinity contacts to form *condensates* within the nucleus. These condensates concentrate proteins involved in transcription, increasing the rates with which they interact, thus greatly stimulating the rate of transcription.

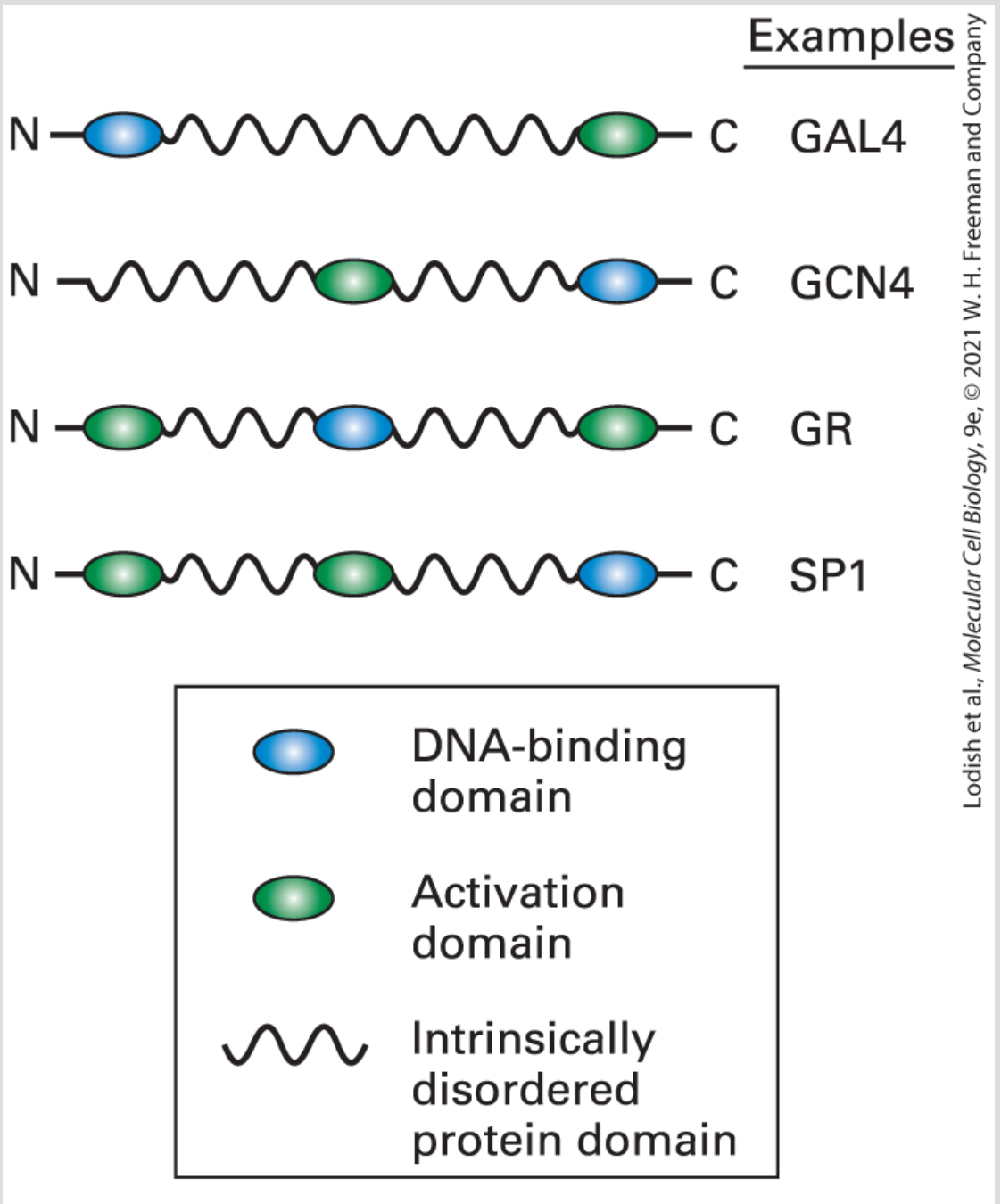


FIGURE 8-23 Schematic diagrams illustrating the modular structure of eukaryotic transcription activators. Transcription factors may contain more than one activation domain but rarely contain more than one DNA-binding domain. Gal4 and Gcn4 are yeast transcription activators. The glucocorticoid receptor (GR) promotes transcription of target genes when certain hormones are bound to the C-terminal activation domain. SP1 binds to GC-rich promoter elements (e.g., GCGGCGCGGC) in a large number of mammalian genes.

Repressors Are the Functional Converse of Activators

Eukaryotic transcription is regulated by repressors as well as activators. For example, geneticists have identified mutations in yeast that result in continuously high expression of certain genes. This type of unregulated, abnormally high expression, called **constitutive** expression, results from the inactivation of a repressor that normally inhibits the transcription of these genes. Similarly, mutants of *Drosophila melanogaster* and *Caenorhabditis elegans* have been isolated that are defective in embryonic development because they express genes in embryonic cells where those genes are normally repressed. The mutations in these mutants inactivate repressors, leading to abnormal development.

Repressor-binding sites in DNA have been identified by systematic linker scanning mutation analyses similar to the one depicted in Figure 8-17. In this type of analysis, whereas mutation of an activator-binding site leads to decreased expression of the linked reporter gene, mutation of a repressor-binding site leads to increased expression of a reporter gene. The repressor proteins that bind such sites can be purified and assayed for their ability to bind to the repressing DNA-control element using the same biochemical techniques described earlier for activator proteins.

Eukaryotic transcription repressors are the functional converse of activators. They can inhibit transcription of a gene they do not normally regulate when their cognate binding sites are placed within tens of base pairs to many kilobases of the gene's transcription start site. Like activators, most eukaryotic repressors are modular proteins that have two functional domains: a DNA-binding domain and a **repression domain**. Like activation domains, repression domains continue to function when fused to another type of DNA-binding domain. If binding sites for this second DNA-binding domain are inserted within a few hundred base pairs of a promoter, expression of the fusion protein inhibits transcription from the promoter. Also like activation domains, repression domains function by interacting with other proteins, as discussed later in this chapter.

DNA-Binding Domains Can Be Classified into Numerous Structural Types

To activate or repress transcription, transcription factors must bind to the specific DNA sequences of transcription control regions via their DNA-binding domains. The DNA-binding domains of eukaryotic transcription factors are comprised of a variety of structural motifs that bind specific DNA sequences. The ability of DNA-binding proteins to bind to specific DNA sequences commonly

results from noncovalent interactions between atoms in an α helix in the DNA-binding domain and atoms on the edges of the bases within the major groove in the DNA. Ionic interactions between positively charged residues arginine and lysine and negatively charged phosphates in the sugar-phosphate backbone, and in some cases, interactions with atoms in the DNA minor groove, also contribute to binding.

The principles of specific protein-DNA interactions were first discovered during the study of bacterial repressors. Many bacterial repressors are dimeric proteins in which an α helix from each monomer inserts into the major groove in the DNA helix and makes multiple, specific interactions with the atoms there (Figure 8-24). This α helix is referred to as the *recognition helix* or *sequence-reading helix* because most of the amino acid side chains that contact bases in the DNA extend from this helix. The recognition helix, which protrudes from the surface of a bacterial repressor, is usually supported in the protein structure in part by hydrophobic interactions with a second α helix located just to the amino-terminal side of it. This entire structural element, which is present in many bacterial repressors, is called a *helix-turn-helix* motif (see Figure 8-24).

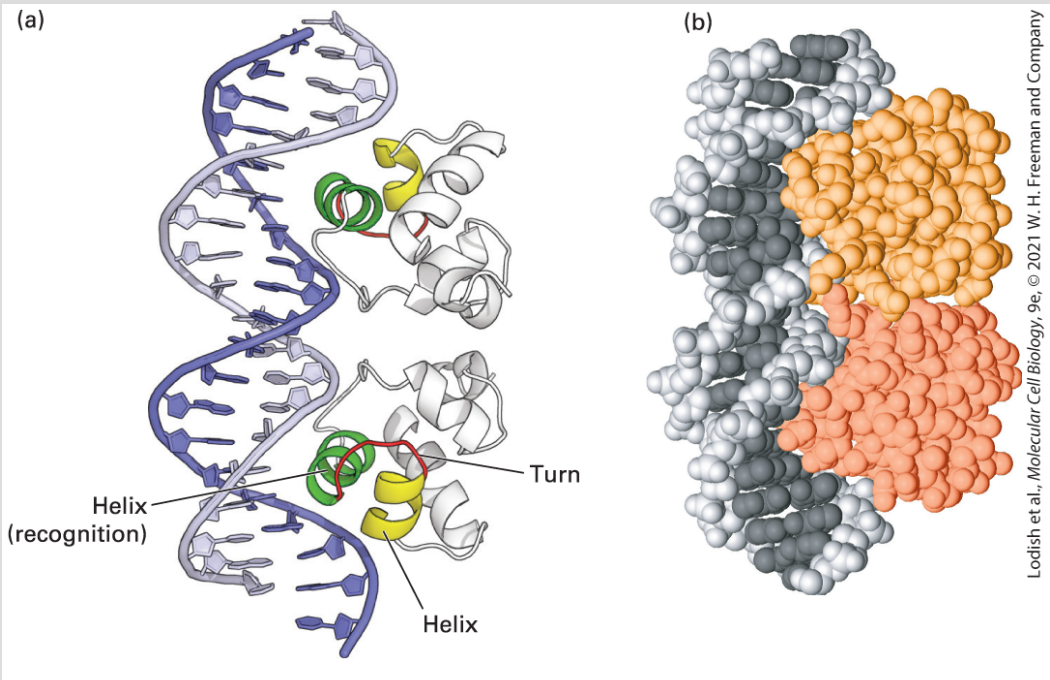


FIGURE 8-24 Interaction of bacteriophage 434 repressor with DNA. (a) Ribbon diagram of 434 repressor bound to its specific operator DNA. The recognition helices are green. The beginning of the α helix N-terminal to the recognition helix is yellow and the turn in the polypeptide backbone between the helices in the helix-turn-helix structural motif is red. (b) A space-filling model of the repressor-operator complex shows how the protein interacts intimately with one side of the DNA molecule over a length of 1.5 turns.

[Data from A. K. Aggarwal et al., 1988, *Science* **242**:899, PDB ID 2ori.]

There are a number of structural motifs in eukaryotic transcription factors that present an α helix to the major groove of DNA. Because most of these motifs have characteristic consensus–amino acid sequences, potential transcription factors can be recognized among the cDNA sequences from various tissues that have been characterized in humans and other species. Here we introduce several common classes of DNA-binding proteins whose three-dimensional structures have been determined. In all of these

examples, and in many other transcription factors, at least one α helix is inserted into the major groove of DNA. However, note that some transcription factors contain alternative structural motifs that interact with DNA (e.g., β strands and loops; see NFAT in Figure 8-28a as an example).

Homeodomain Proteins

Many eukaryotic transcription factors that function during development contain a conserved 60-residue DNA-binding motif, called a **homeodomain**, that is similar to the helix-turn-helix motif of bacterial repressors. These transcription factors were first identified in *Drosophila* mutants in which one body part was transformed into another during development (see Figure 8-2b). The conserved homeodomain sequence has also been found in vertebrate transcription factors, including those that have similar master-control functions required for the development of anatomical structures.

Zinc-Finger Proteins

The DNA-binding domain found most often in vertebrate genomes is called a **zinc finger**. The polypeptide chain in this structural domain is folded around a central Zn^{2+} ion, producing a compact domain from a relatively short length of polypeptide (∼25 amino acid residues). This structural motif was first recognized in DNA-binding

domains, but is now known to occur in other proteins that do not bind to DNA. Here we describe two of the several classes of zinc-finger motifs that have been identified in eukaryotic transcription factors.

The *C_2H_2 zinc finger* is the most common DNA-binding motif encoded in the human genome and the genomes of other vertebrates and is also common in multicellular plants. This motif has a 23–26-residue consensus sequence containing two conserved cysteine (C) and two conserved histidine (H) residues, whose side chains bind one Zn^{2+} ion (see Figure 3-7c). The name "zinc finger" was coined because a two-dimensional diagram of the structure resembles a finger. When the three-dimensional structure was solved, it became clear that the binding of the Zn^{2+} ion by the two cysteine and two histidine residues folds the relatively short polypeptide sequence into a compact domain, which can insert its α helix into the major groove of DNA. C_2H_2 zinc finger proteins contain at least two of these Zn-fingers, and many contain multiple C_2H_2 zinc fingers, which interact with successive groups of base pairs, within the major groove, as the protein wraps around the DNA double helix (Figure 8-25a).

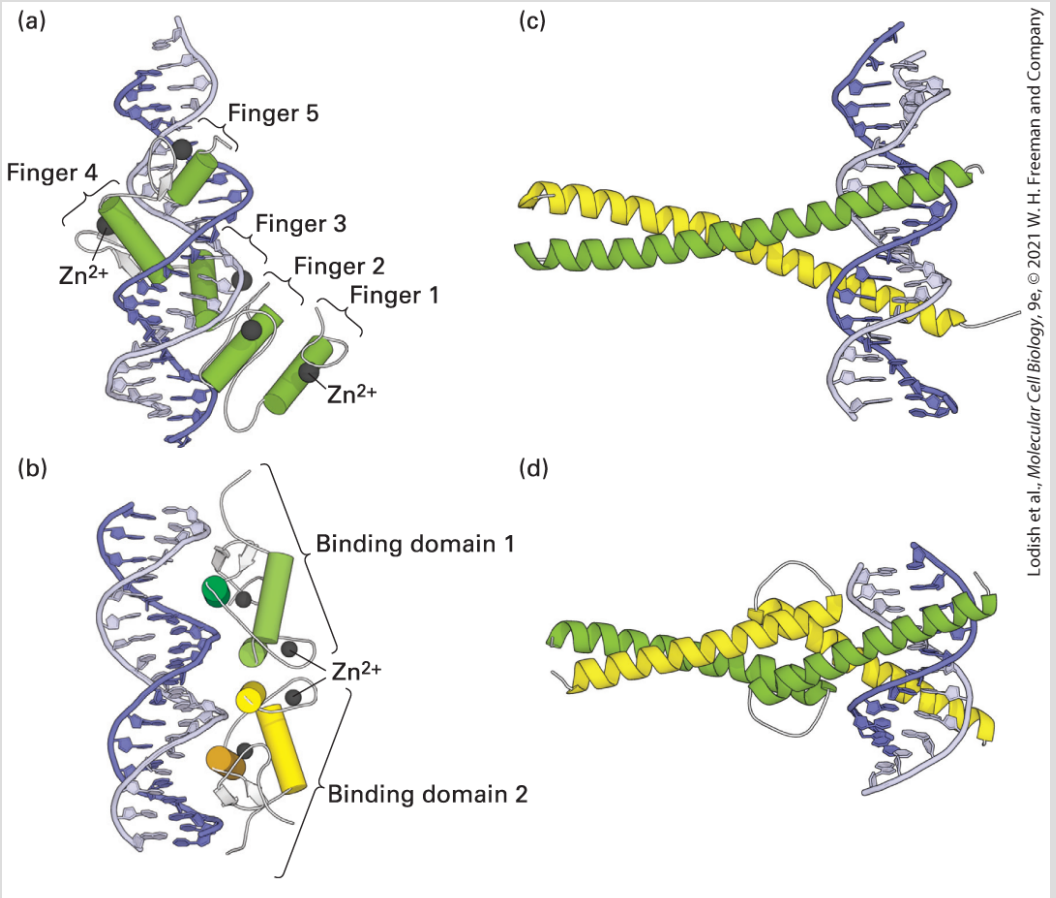


FIGURE 8-25 Eukaryotic DNA-binding domains that use an α helix to interact with the major groove of specific DNA sequences. (a) The GL1 DNA-binding domain is monomeric and contains five C_2H_2 zinc fingers. The α helices are shown as cylinders, the Zn^{2+} ions as spheres. Finger 1 does not interact with DNA, whereas the other four fingers do. (b) The glucocorticoid receptor is a homodimeric C_4 zinc-finger protein. The α helices are shown as cylinders, the β strands as white arrows, and the Zn^{2+} ions as black spheres. Two α helices, one in each monomer, interact with the DNA. Like all C_4 zinc-finger homodimers, this transcription factor has twofold rotational symmetry. (c) In leucine-zipper proteins, basic residues in the extended α-helical regions of the monomers interact with the DNA backbone at adjacent sites in the major groove. The coiled-coil dimerization domain is stabilized by hydrophobic interactions between the monomers. (d) In bHLH proteins, the DNA-binding helices at the bottom (N-termini of the monomers) are separated by nonhelical loops from a leucine zipper–like region containing a coiled-coil dimerization domain.

[Part (a) Data from N. P. Pavletich and C. O. Pabo, 1993, *Science* **261**:1701. Part (b) Data from B. F. Luisi et al., 1991, *Nature* **352**:497. Part (c) Data from T. E. Ellenberger et al., 1992, *Cell* **71**:1223, PDB ID 1ysa. Part (d) Data from A. R. Ferre-D'Amare et al., 1993, *Nature* **363**:38, PDB ID 1an2.]

A second type of zinc-finger structure, designated the *C_4 zinc finger* (because it has four conserved cysteines in contact with the Zn^{2+}), is found in some 50 human transcription factors. The first members of this class were identified as specific intracellular high-affinity binding proteins, or "receptors," for steroid hormones, which led to the name *steroid receptor superfamily.* Because similar intracellular receptors for nonsteroid hormones were subsequently found, these transcription factors are now commonly called **nuclear receptors**. The characteristic feature of C_4 zinc fingers is the presence of two groups of four critical cysteines, one toward each end of the 55–56-residue domain.

Although the C_4 zinc finger was initially named by analogy with the C_2H_2 zinc finger, the three-dimensional structures of proteins containing these DNA-binding motifs were later found to be quite distinct. A particularly important difference between the two is that C_2H_2 zinc-finger proteins generally contain three or more repeating finger units and bind as monomers, whereas C_4 zinc-finger proteins generally contain only two finger units and generally bind to DNA as homodimers or heterodimers. Homodimers of C_4 zinc-finger DNA-binding domains have twofold rotational symmetry (Figure 8-25b).

Consequently, homodimeric nuclear receptors bind to consensus DNA sequences that are inverted repeats.

Leucine-Zipper Proteins

Another structural motif present in the DNA-binding domains of a large class of transcription factors contains the hydrophobic amino acid leucine at every seventh position in the sequence. These proteins bind to DNA as dimers, and mutagenesis of the leucines has shown that they are required for dimerization. Consequently, the name **leucine zipper** was coined to denote this structural motif of a coiled coil of two α helices.

The DNA-binding domain of the yeast Gcn4 transcription factor mentioned earlier is a leucine-zipper domain. X-ray crystallographic analysis of complexes between DNA and the Gcn4 DNA-binding domain has shown that the dimeric protein contains two extended α helices that grip the DNA molecule, much like a pair of scissors, at two adjacent sites in the major groove separated by about half a turn of the double helix (Figure 8-25c). The portions of the α helices contacting the DNA include positively charged (basic) residues that interact with phosphates in the DNA backbone and additional residues that interact with specific bases in the major groove.

Gcn4 forms dimers via hydrophobic interactions between the C-terminal regions of the α helices, forming a coiled-coil structure. This structure is common in proteins containing amphipathic α

helices in which hydrophobic amino acid residues are regularly spaced alternately three or four positions apart in the sequence, forming a stripe down one side of the α helix. These hydrophobic stripes make up the interacting surfaces between the α-helical monomers in a coiled-coil dimer (see Figure 3-7a).

Although the first leucine-zipper transcription factors to be analyzed contained leucine residues at every seventh position in the dimerization region, additional DNA-binding proteins containing other hydrophobic amino acids in these positions were subsequently identified. Like leucine-zipper proteins, they form dimers containing a C-terminal coiled-coil dimerization region and an N-terminal DNA-binding domain. The term *basic zipper (bZIP)* is now frequently used to refer to all proteins with these common structural features. Many basic-zipper transcription factors are heterodimers of two different polypeptide chains, each containing one basic-zipper motif.

Basic Helix-Loop-Helix (bHLH) Proteins

The DNA-binding domain of another class of dimeric transcription factors contains a structural motif that is very similar to the basic-zipper motif except that a nonhelical loop of the polypeptide chain separates two α-helical regions in each monomer (Figure 8-25d). Termed a **basic helix-loop-helix (bHLH)**, this motif was predicted from the amino acid sequences of these proteins, which contain an

N-terminal α helix with basic residues that interact with DNA, a middle loop region, and a C-terminal region, with hydrophobic amino acids spaced at intervals characteristic of an amphipathic α helix, that dimerizes into a coiled coil. As with basic-zipper proteins, different bHLH proteins can form heterodimers.

Structurally Diverse Activation and Repression Domains Regulate Transcription

Experiments with fusion proteins composed of the Gal4 DNA-binding domain and random segments of *E. coli* proteins demonstrated that a diverse group of amino acid sequences (∼1 percent of all *E. coli* sequences) can function as activation domains, even though they evolved to perform other functions. Many transcription factors contain activation domains marked by an unusually high percentage of particular amino acids. Gal4, Gcn4, and most other yeast transcription factors, for instance, have activation domains that are rich in acidic amino acids (aspartic and glutamic acids). These so-called *acidic activation domains* are generally capable of stimulating transcription in nearly all types of eukaryotic cells — fungal, animal, and plant cells. Activation domains from some *Drosophila* and mammalian transcription factors are glutamine-rich, and some are proline-rich; still others are rich in the closely related amino acids serine and threonine,

both of which have hydroxyl groups. However, some strong activation domains are not particularly rich in any specific amino acid.

Biophysical studies indicate that acidic activation domains have an unstructured, random-coil, intrinsically disordered conformation. These domains stimulate transcription when they are bound to a protein *co-activator*. The interaction with a co-activator causes the activation domain to assume a more structured α-helical conformation in the activation domain–co-activator complex. A well-studied example of a transcription factor with an acidic activation domain is the mammalian CREB protein, which is phosphorylated in response to increased levels of cAMP. This regulated phosphorylation is required for CREB to bind to its co-activator CBP (*C*REB *b*inding *p*rotein), resulting in the transcription of genes whose control regions contain a CREB-binding site (see Figure 15-24). When the phosphorylated random-coil activation domain of CREB interacts with CBP, it undergoes a conformational change to form two α helices linked by a short loop, which wrap around the interacting domain of CBP (Figure 8-26a).

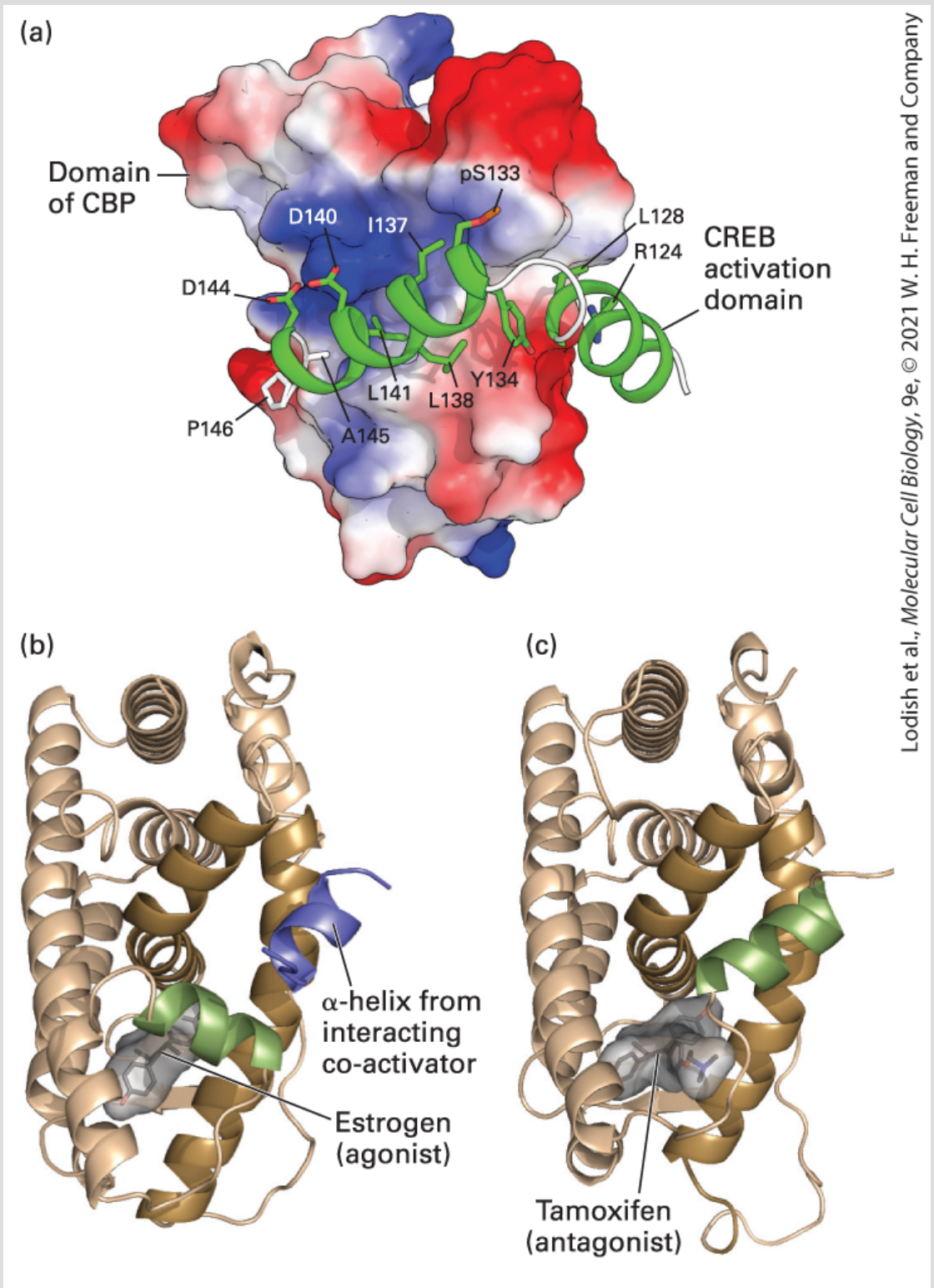


FIGURE 8-26 Activation domains may be random coils until they interact with co-activator proteins or folded protein domains. (a) The activation domain of CREB (*c*yclic AMP *r*esponse *e*lement-*b*inding protein) is activated by phosphorylation at

serine 133. It is a random coil until it interacts with a domain of its co-activator, CBP (shown as a space-filling surface model with negatively charged regions in red and positively charged regions in blue). When the CREB activation domain binds to CBP, it folds into two amphipathic α helices. Side chains in the activation domain that interact with the surface of the CBP domain are labeled. (b) The ligand-binding activation domain of the estrogen receptor is a folded-protein domain. When estrogen is bound to the domain, the green α helix interacts with the ligand, generating a hydrophobic groove in the ligand-binding domain (dark brown helices), which binds an amphipathic α helix in a co-activator subunit (blue). (c) The conformation of the estrogen receptor in the absence of hormone is stabilized by binding of the estrogen antagonist tamoxifen. In this conformation, the green helix of the receptor folds into a conformation that interacts with the co-activator–binding groove of the active receptor, sterically blocking binding of co-activators.

[Part (a) Data from I. Radhakrishnan et al., 1997, *Cell* **91**:741, PDB ID 1kdx. Parts (b) and (c) Data from A. K. Shiau et al., 1998, *Cell* **95**:927, PDB ID 3erd and 3ert.]

Some activation domains are larger and more highly structured than acidic activation domains. For example, the ligand-binding domains of nuclear receptors function as activation domains when they bind their specific hormone ligand (Figure 8-26b, c). Binding of ligand induces a large conformational change in the nuclear receptor that allows the ligand-binding domain with bound hormone to interact with a short α helix in a co-activator; the resulting complex can then activate transcription of genes whose control regions bind the nuclear receptor.

Thus the acidic activation domain in CREB and the ligand-binding activation domains in nuclear receptors represent two structural extremes. The CREB acidic activation domain is an intrinsically

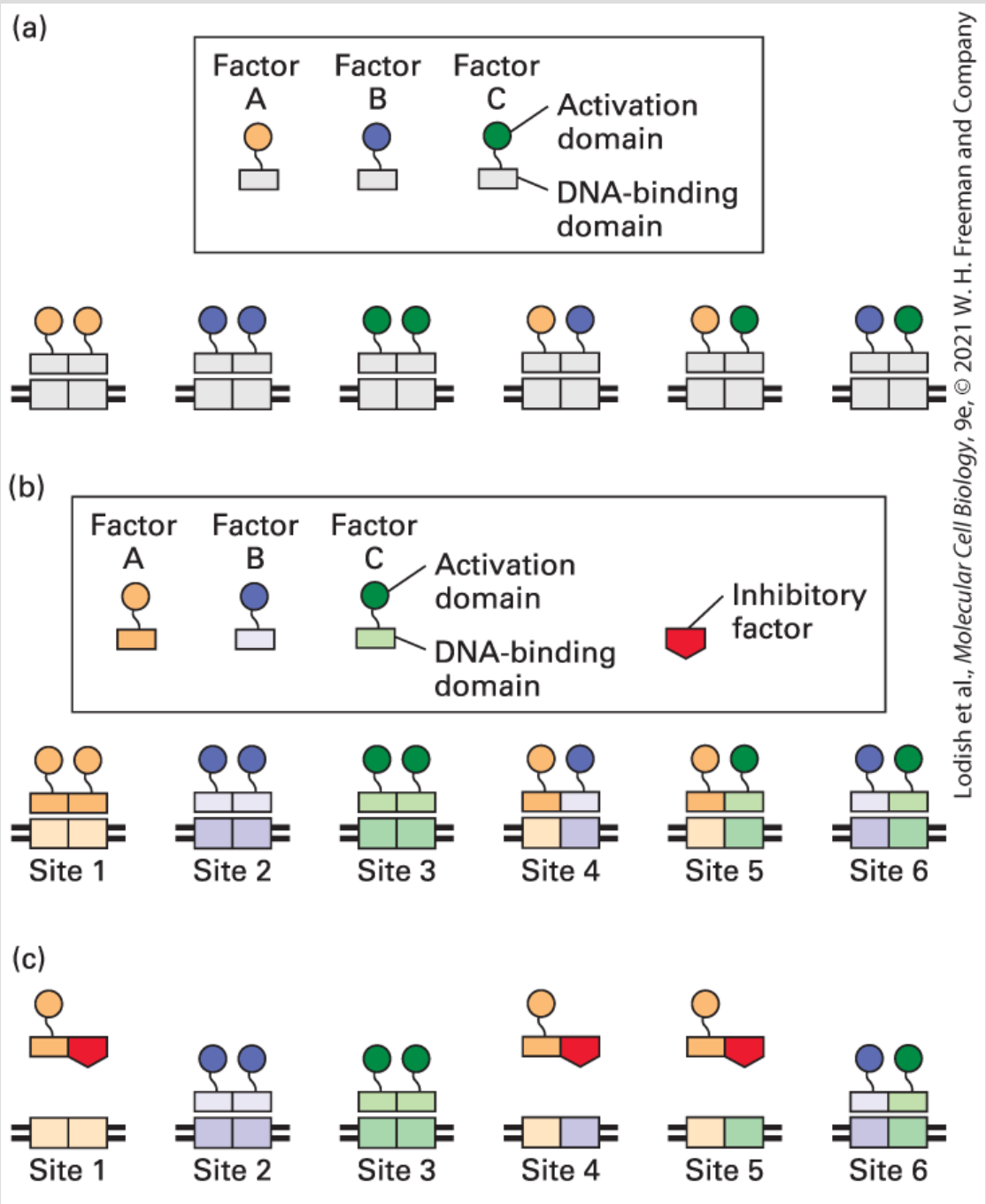
disordered random coil that folds into two α helices when it binds to the surface of a globular domain in a co-activator. In contrast, the nuclear-receptor ligand-binding activation domain is a structured globular domain that interacts with a short α helix in a co-activator, which probably is a random coil before it is bound. In both cases, however, specific protein-protein interactions between a co-activator and the activation domain permit the transcription factor to stimulate gene expression.

Currently, less is known about the structure of repression domains. The globular ligand-binding domains of some nuclear receptors function as repression domains in the absence of their specific hormone ligand. Like activation domains, repression domains may be relatively short, comprising 15 or fewer amino acids. Biochemical and genetic studies indicate that repression domains also mediate protein-protein interactions and bind to *co-repressor* proteins, forming a complex that inhibits transcription initiation by mechanisms that are discussed later in the chapter.

Transcription Factor Interactions Increase Gene-Control Options

Two types of DNA-binding proteins discussed previously — bZIP and bHLH proteins — often exist in alternative heterodimeric combinations of monomers. Other classes of transcription factors

not discussed here also form heterodimeric proteins. In some heterodimeric transcription factors, each monomer recognizes the same sequence. In these cases, the formation of alternative heterodimers does not increase the number of different sites on which the monomers can act, but rather allows the activation domains associated with each monomer to be brought together in alternative combinations that bind to the same site (Figure 8-27a). As we will see later, and in subsequent chapters, the activities of individual transcription factors can be regulated by multiple mechanisms. Consequently, a single bZIP- or bHLH-binding DNA-regulatory element in the transcription-control region of a gene may elicit different transcriptional responses depending on which bZIP or bHLH monomers are expressed in the cell and how their activities are regulated.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

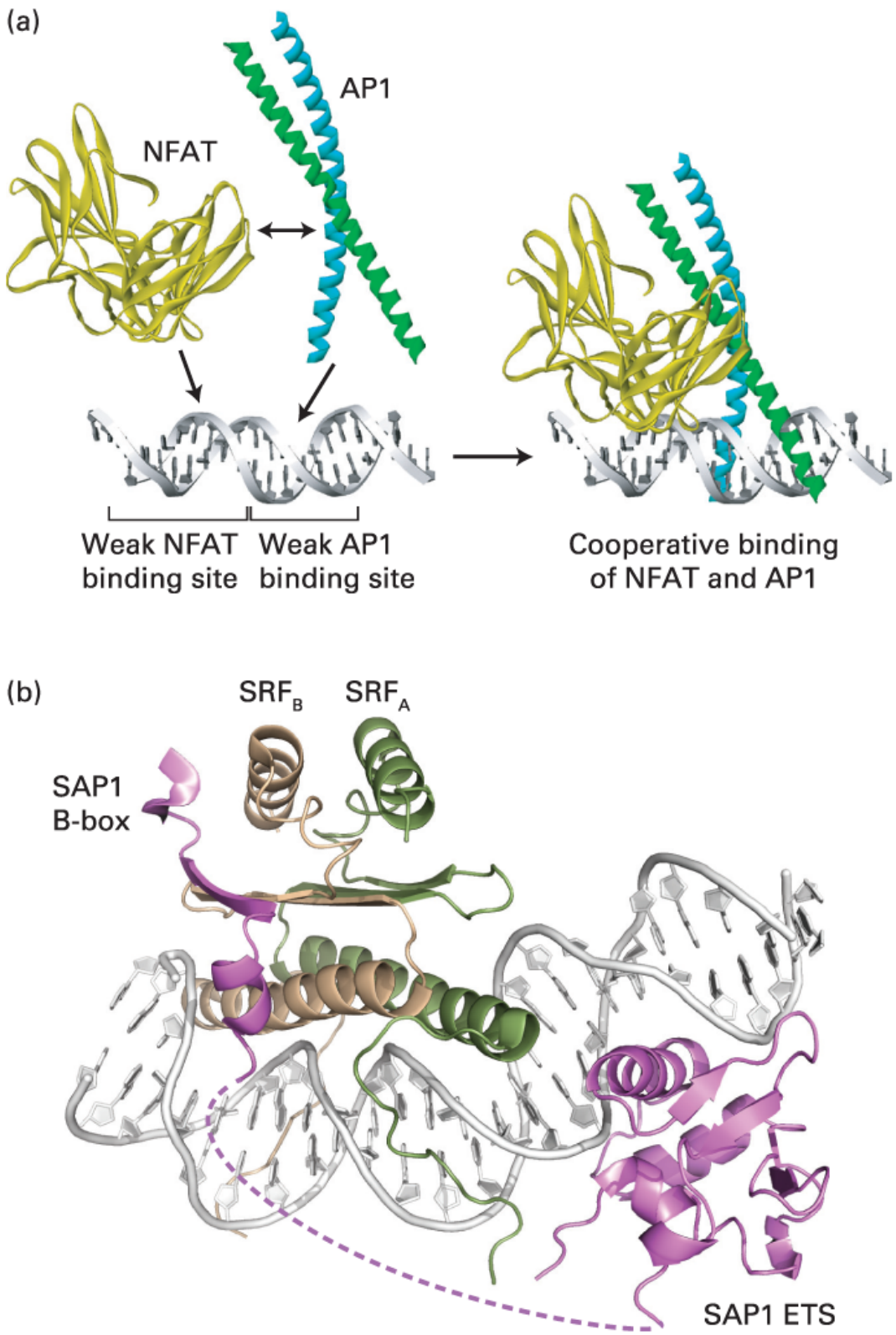
FIGURE 8-27 Combinatorial possibilities due to formation of heterodimeric transcription factors. (a) In some heterodimeric transcription factors, the activation domain of each monomer recognizes the same DNA sequence. In the hypothetical example shown, transcription-factor monomers A, B, and C can all interact with one another, creating six different alternative combinations of activation domains that can all bind at the same site. Each composite binding site is divided into two half-sites, and each heterodimeric factor contains the activation domains of its two

constituent monomers. (b) When transcription-factor monomers recognize different DNA sequences, six alternative combinations of the transcription-factor monomers A, B, and C, each with a unique pair of activation domains, can bind to six different DNA sequences (sites 1–6). (c) Expression of an inhibitory factor (red) that interacts only with dimerization domain of A inhibits binding; hence transcriptional activation at sites 1, 4, and 5 is inhibited, but activation at sites 2, 3, and 6 is unaffected.

In some heterodimeric transcription factors, however, each monomer has a different DNA-binding specificity. The resulting combinatorial possibilities increase the number of potential DNA sequences that a family of transcription factors can bind. Three different transcription-factor monomers could theoretically combine to form six different homo- and heterodimeric transcription factors, as illustrated in Figure 8-27b. Four different monomers could form a total of ten dimeric factors; five monomers, sixteen dimeric factors; and so forth. In addition, inhibitory factors are known that bind to some bZIP and bHLH monomers, thereby blocking their binding to DNA. When these inhibitory factors are expressed, they repress transcriptional activation by the factors with which they interact (Figure 8-27c). Thus the rules governing the interactions of members of a heterodimeric transcription factor family are complex. This combinatorial complexity expands both the number of DNA sites from which these factors can activate transcription and the ways in which they can be regulated.

Similar combinatorial transcription regulation is achieved through the interaction of structurally unrelated transcription factors bound to closely spaced binding sites in DNA. An example is the interaction

of two transcription factors, NFAT and AP1, that bind to neighboring sites in a composite promoter-proximal element regulating the gene encoding interleukin-2 (IL-2). Expression of the *IL-2* gene is critical to the immune response, but abnormal expression of IL-2 can lead to autoimmune diseases such as rheumatoid arthritis (see Chapter 24). Neither NFAT nor AP1 binds to its site in the *IL-2* control region in the absence of the other protein. The affinities of these factors for these particular DNA sequences are too low for the individual factors to form a stable complex with DNA. However, when both NFAT and AP1 are present, protein-protein interactions between them stabilize the ternary complex composed of NFAT, AP1, and DNA (Figure 8-28a). Such *cooperative DNA binding* by various transcription factors results in considerable combinatorial complexity of transcriptional control. As a result, the 1600 or so transcription factors encoded in the human genome can bind to DNA through a much larger number of cooperative interactions, resulting in unique transcriptional control for each of the roughly 20,000 human genes. In the case of *IL-2*, transcription occurs only when NFAT is activated, which results in its transport from the cytoplasm to the nucleus, and when the genes encoding the two subunits of AP1 are activated at the same time. These two events are controlled by distinct signal transduction pathways (see Chapters 15 and 16), allowing stringent control of IL-2 expression.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 8-28 Cooperative binding of two unrelated transcription factors to neighboring sites in a composite control element. (a) By themselves, both monomeric NFAT and heterodimeric AP1 transcription factors have low affinity for their respective binding sites in the *IL-2* promoter-proximal region. Protein-protein interactions between NFAT and AP1 add to the overall stability of the NFAT-AP1-DNA complex, so that the two proteins bind to the composite site cooperatively. (b) Cooperative DNA binding by dimeric SRF and monomeric SAP1 can occur when their binding sites are separated by 5–30 bp and when the SAP1 binding site is inverted because the domain of SAP1 that interacts with SRF is connected to the DNA-binding domain of SAP1 by a flexible linker region of the SAP1 polypeptide chain (dotted line).

[Part (a) Data from L. Chen et al., 1998, *Nature* **392**:42, PDB ID 1a02. Part (b) Data from M. Hassler and T. J. Richmond, 2001, *EMBO J*. **20**:3018, PDB ID 1hbx.]

Cooperative binding by NFAT and AP1 occurs only when their weak binding sites are positioned quite close to each other in DNA. The sites must be located at a precise distance from each other for effective binding. The requirements for cooperative binding are not so stringent in the case of some other transcription factors and transcription-control regions. For example, the *EGR-1* control region contains a composite binding site to which the SRF and SAP1 transcription factors bind cooperatively (Figure 8-28b). Because SAP1 has a long, flexible domain that interacts with SRF, the two proteins can bind cooperatively when their individual sites in DNA are separated by any distance up to about 30 bp or when the sites are inverted relative to each other.

Multiprotein Complexes Form on Enhancers

Recall that enhancers generally range in length from about 50 to 200 bp and include binding sites for multiple transcription factors. Let's look at the roughly 50-bp enhancer that regulates expression of β-interferon, an important protein in the defense against viral infections in vertebrates. The β-interferon enhancer binds several dimeric and single-subunit transcription factors. These can bind to specific transcription-factor-binding sites that constitute the enhancer (Figure 8-29). The term **enhanceosome** has been coined to describe large DNA-protein complexes, such as these, that assemble from transcription factors as they bind to the multiple binding sites in an enhancer.

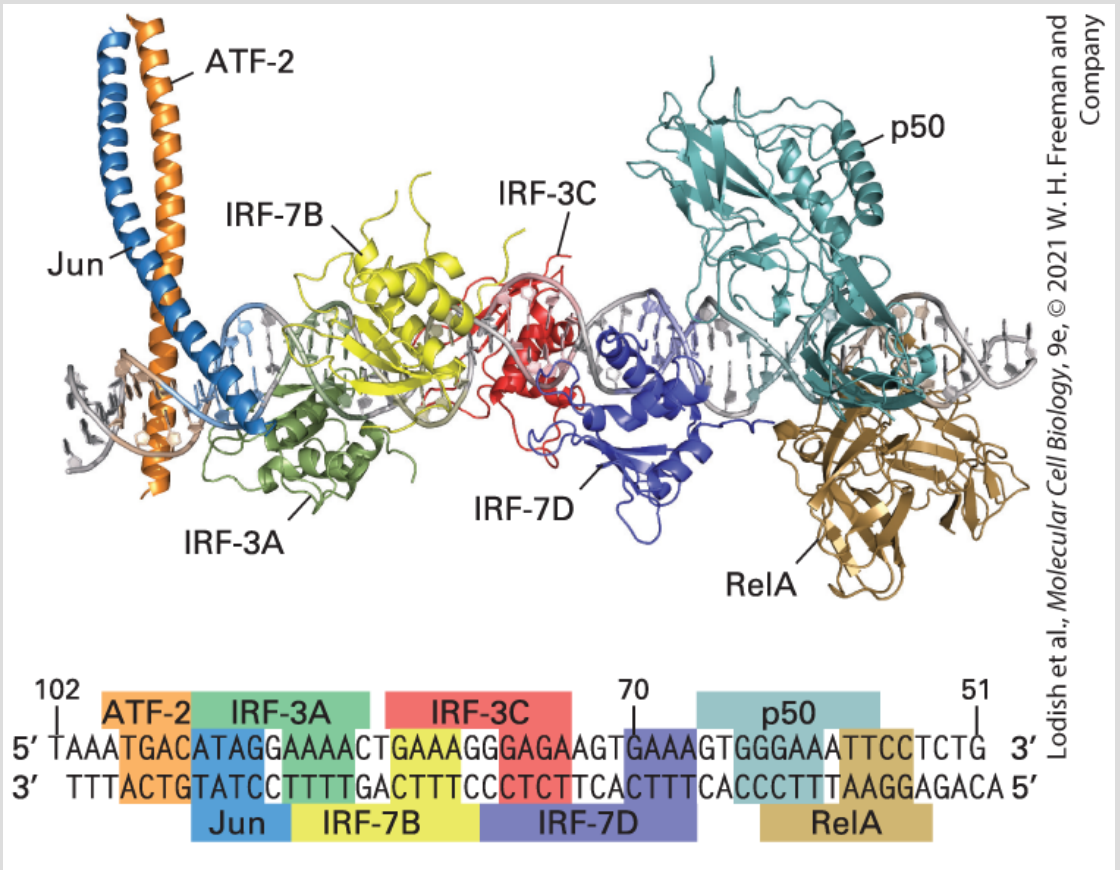
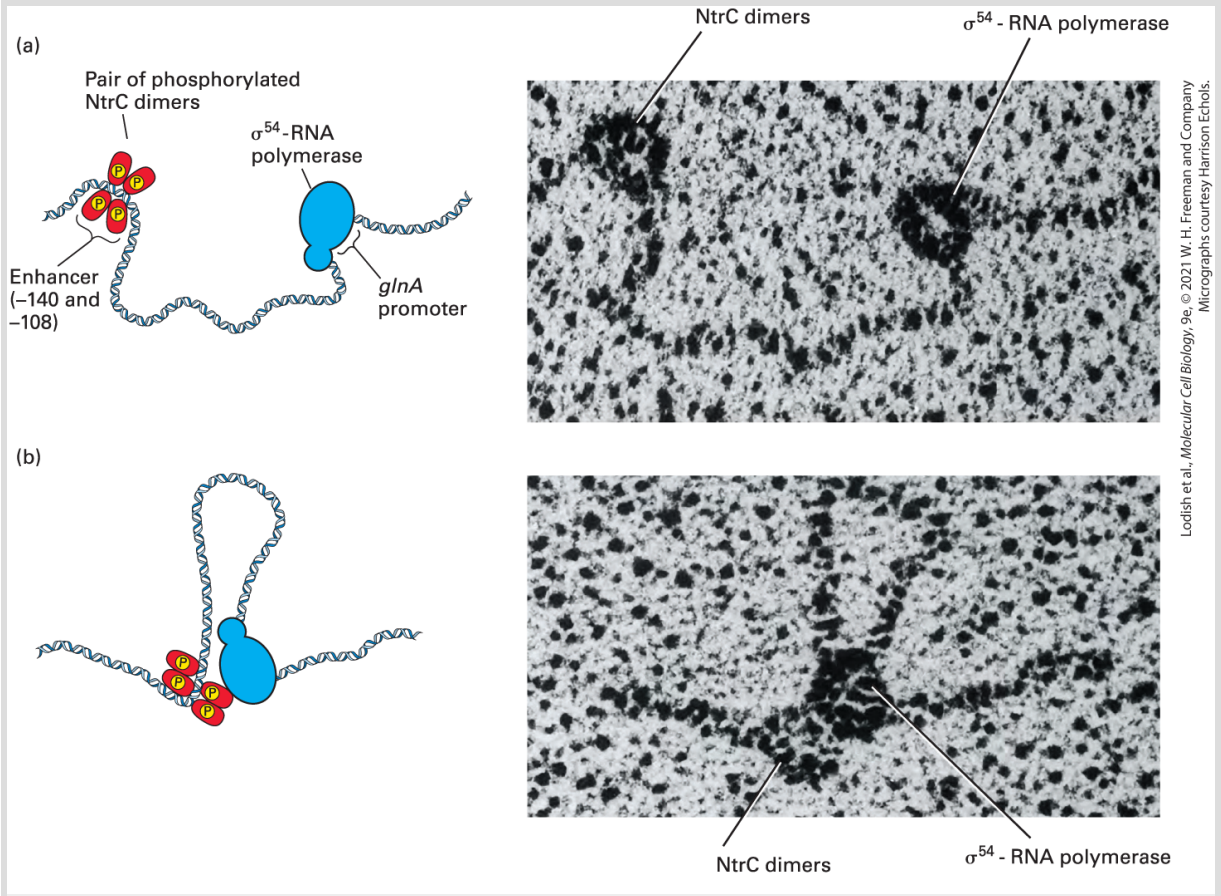


FIGURE 8-29 Model of the enhanceosome that forms on the β-interferon enhancer. Two heterodimeric factors, Jun/ATF-2 and p50/RelA (NF-κB), and two copies each of the monomeric transcription factors IRF-3 and IRF-7, bind to the six overlapping binding sites in this enhancer. See D. Penne, T. Manniatis, and S. Harrison, 2007, *Cell* **129**:1111.

Because of the presence of flexible regions connecting the DNA-binding domains and activation or repression domains in transcription factors (see Figure 8-23), and because of the ability of interacting proteins bound to distant sites to produce loops in the DNA between their binding sites (Figure 8-30), considerable leeway in the spacing between regulatory elements in transcription-control regions is permissible. This tolerance for variable spacing between

binding sites for specific transcription factors, and between promoter binding sites for the general transcription factors and for Pol II, probably contributed to rapid evolution of gene control in eukaryotes. Transposition of DNA sequences and recombination between repeated sequences over evolutionary time probably created new combinations of control elements that were subjected to natural selection and retained if they proved beneficial. The latitude in spacing between regulatory elements probably allowed many more functional combinations to be subjected to this evolutionary experimentation than would be the case if constraints on the spacing between regulatory elements were strict, as is true for most genes in bacteria.



EXPERIMENTAL FIGURE 8-30 Example of DNA looping between RNA polymerase and an activator protein from *E. coli.* DNA looping permits interaction of bound NtrC and σ^{54}–RNA polymerase. (a) Drawing (*left*) and electron micrograph (*right*) of DNA restriction fragment with phosphorylated, activated NtrC dimers bound to the enhancer region near one end of the restriction fragment and σ^{54}–RNA polymerase bound to the *glnA* promoter near the other end. (b) Drawing (*left*) and electron micrograph (*right*) of the same fragment preparation, showing NtrC dimers and σ^{54}– RNA polymerase bound to each other, with the intervening DNA forming a loop between them. See W. Su et al., 1990, *Proc. Nat'l. Acad. Sci. USA* **87**:5505.

KEY CONCEPTS OF SECTION 8.3

Regulatory Sequences in Protein-Coding Genes and the Proteins Through Which They Function

- Expression of eukaryotic protein-coding genes is generally regulated through multiple protein-binding transcription-control regions that are located close to or distant from the transcription start site (see Figure 8-18).
- Promoters direct binding of RNA polymerase II to DNA, determine the site of transcription initiation, and influence the frequency of transcription initiation.
- Promoter-proximal elements occur within about 200 bp of a start site. Several such elements, containing 6–10 bp, may help regulate a particular gene.
- Enhancers, which contain multiple short control elements, may be located from 200 bp to tens of kilobases upstream or downstream from a promoter, within an intron, or downstream from the final exon of a gene.
- Promoter-proximal elements and enhancers are often cell-type-specific, functioning only in specific differentiated cell types.
- Transcription factors, which activate or repress transcription, bind to promoter-proximal regulatory elements and enhancers in eukaryotic DNA.
- Transcription activators and repressors are generally modular proteins containing a single DNA-binding domain and one or a few activation domains (for activators) or repression domains (for repressors). The different domains are frequently linked by flexible intrinsically disordered polypeptide regions (see Figure 8-23).
- Among the most common structural motifs found in the DNA-binding domains of eukaryotic transcription factors are the homeodomain, C_2H_2 zinc finger, basic zipper (leucine zipper), and basic helix-loop-helix (bHLH). All these and many other DNA-

binding motifs contain one or more α helices that interact with the major groove in their cognate site in DNA.

- Activation and repression domains in transcription factors exhibit a variety of amino acid sequences and three-dimensional structures. In general, these functional domains interact with co-activators or co-repressors, which are critical to the ability of transcription factors to modulate gene expression.
- The transcription-control regions of most genes contain binding sites for multiple transcription factors. Transcription of such genes varies depending on the particular repertoire of transcription factors that are expressed and activated in a particular cell at a particular time.
- Combinatorial complexity in transcriptional control results from alternative combinations of monomers that form heterodimeric transcription factors (see Figure 8-27) and from cooperative binding of transcription factors to composite control sites (see Figure 8-28).
- Binding of multiple transcription factors to multiple sites in an enhancer forms a large DNA-protein complex called an *enhanceosome* (see Figure 8-29).

8.4 Molecular Mechanisms of Transcription Repression and Activation

The repressors and activators that bind to specific sites in DNA and regulate expression of the associated protein-coding genes do so by three general mechanisms. First, these regulatory proteins act in concert with other proteins to modulate chromatin structure, inhibiting or stimulating the ability of general transcription factors to bind to promoters. Recall from Chapter 7 that the DNA in eukaryotic cells is not free, but is associated with a roughly equal mass of protein in the form of **chromatin**. The basic structural unit of chromatin is the **nucleosome**, which is composed of about 147 bp of DNA wrapped tightly around a disk-shaped core of *histone* proteins. Residues within the N-terminal region of each histone, and the C-terminal regions of histones H2A and H2B, called *histone tails,* extend from the surface of the nucleosome and can be reversibly modified (see Figure 7-26). Such modifications influence the relative condensation of chromatin and thus its accessibility to proteins required for transcription initiation. Second, activators and repressors interact with a large multiprotein complex called the *mediator of transcription complex,* or simply **Mediator**. This complex, in turn, binds to Pol II and directly regulates assembly of the preinitiation complex. In addition, some activation domains interact with TFIID-TAF subunits or other components of the preinitiation

complex, and these interactions contribute to preinitiation complex assembly. Finally, activation domains may also interact with the elongation factor P-TEFb (cyclin T-CDK9) and other as yet unknown factors to stimulate elongation by Pol II away from the promoter region.

In this section, we review the current understanding of how repressors and activators control chromatin structure and preinitiation complex assembly. In the next section of the chapter, we discuss how the concentrations and activities of activators and repressors themselves are controlled, so that gene expression is precisely attuned to the needs of the cell and organism.

Formation of Heterochromatin Silences Gene Expression at Telomeres, near Centromeres, and in Other Regions

Inactive genes in eukaryotic cells are often associated with **heterochromatin**, regions of chromatin that are highly condensed and stain darkly with DNA dyes (see Figure 7-28a). In contrast, most transcribed genes are located in lightly staining **euchromatin** (see Figure 7-28a). Regions of chromosomes near the centromeres and telomeres, as well as additional specific regions that vary in different cell types, are organized into heterochromatin. The DNA in

heterochromatin is less accessible to externally added proteins than is DNA in euchromatin and consequently is often referred to as "closed" chromatin. For instance, in an experiment described in Chapter 7, the DNA of inactive genes was found to be far more resistant to digestion by DNase I than the DNA of transcribed genes (see Figure 7-27).

Study of DNA regions in *S. cerevisiae* that behave like the heterochromatin of higher eukaryotes provided early insight into the **chromatin-mediated repression** of transcription. This yeast can grow either as haploid or diploid cells. Haploid cells exhibit one of two possible mating types, called **a** and α. Cells of different mating type can mate, or fuse, to generate a diploid cell (see Figure 1-24b). When a haploid cell divides by budding, the larger mother cell switches its mating type. Genetic and molecular analyses have revealed that three genetic loci on yeast chromosome III control the mating type of yeast cells (Figure 8-31). The central mating-type locus, termed *MAT* — the only one of the three that is actively transcribed — encodes transcription factors (**a**1, or α1 and α2) that regulate genes that determine the mating type. In any one cell, either an **a** or α DNA sequence is located at the *MAT* locus. The two additional loci, termed *HML* and *HMR*, near the left and right telomere, respectively, contain "silent" (nontranscribed) copies of the **a** or α genes. These sequences are transferred alternately from *HML*α or *HMR***a** into the *MAT* locus by a type of nonreciprocal recombination between homologous sequences during cell division.

When the *MAT* locus contains the DNA sequence from *HML*α, the cells behave as α cells. When the *MAT* locus contains the DNA sequence from *HMR***a**, the cells behave like **a** cells.

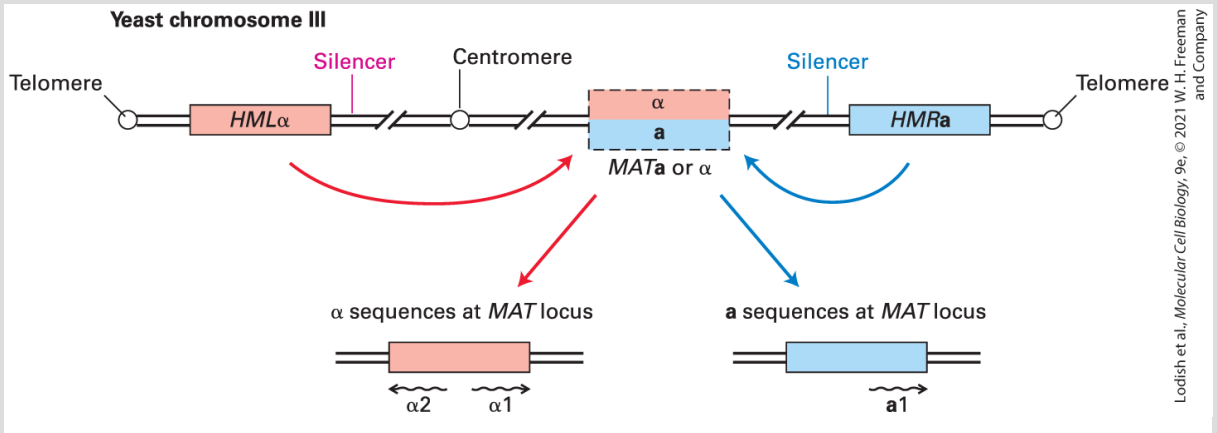


FIGURE 8-31 Arrangement of mating-type loci on chromosome III in the yeast *S. cerevisiae*. Silent (unexpressed) mating-type genes (either **a** or α) are located at the *HML* locus. The opposite mating-type gene is present at the silent *HMR* locus. When the α or **a** sequences are present at the *MAT* locus, they can be transcribed into mRNAs whose encoded proteins specify the mating-type phenotype of the cell. The silencer sequences near *HML* and *HMR* bind proteins that are critical for repression of these silent loci. Haploid cells can switch mating types in a process that transfers the DNA sequence from *HML* or *HMR* to the transcriptionally active *MAT* locus.

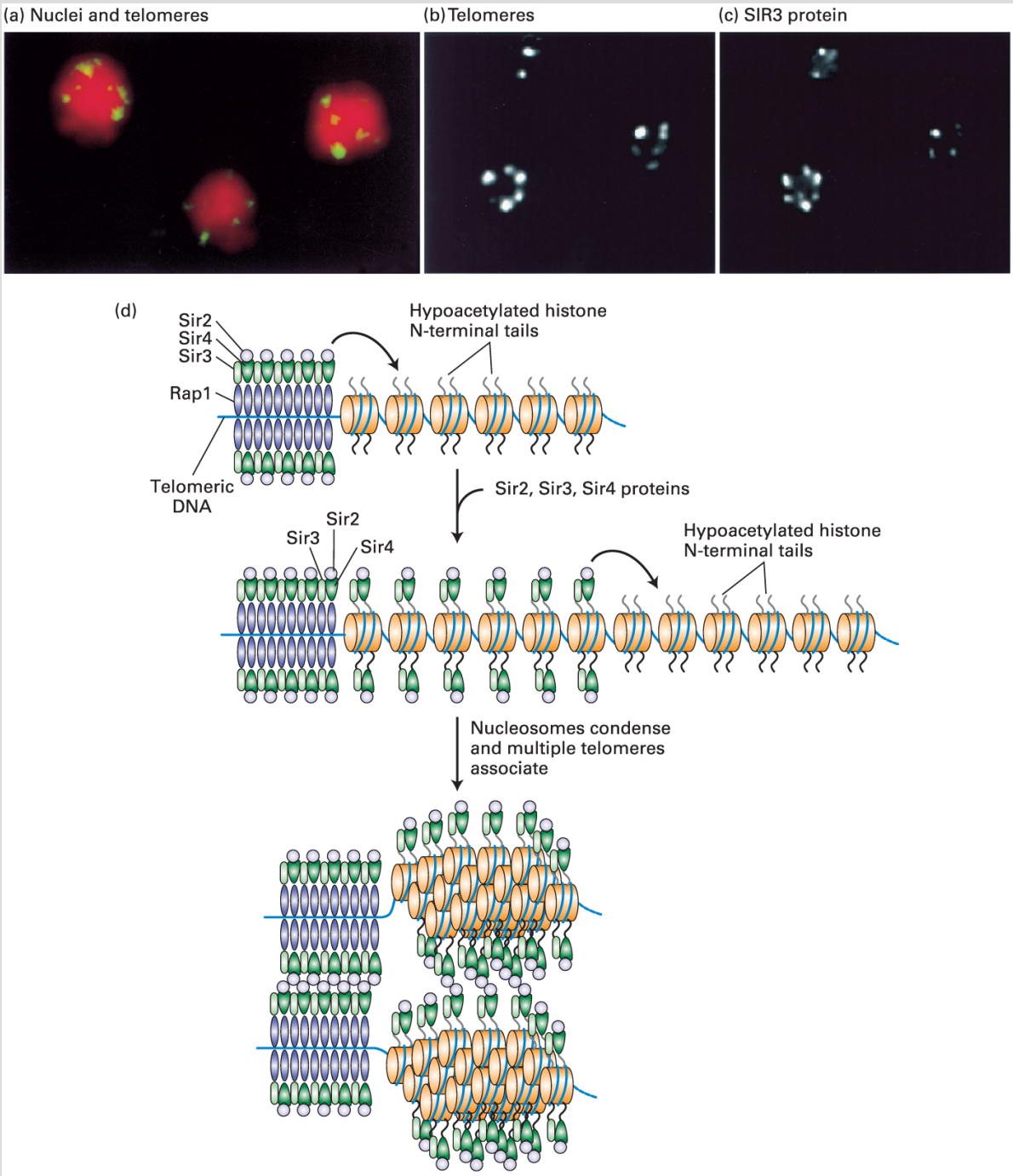
Our interest here is in how transcription of the silent mating-type genes at *HML* and *HMR* is repressed. If these genes are expressed, as they are in yeast mutants with defects in the repressing mechanism, both **a** and α proteins are expressed, causing the cells to behave like diploid cells, which cannot mate. The promoters and UASs controlling transcription of the **a** and α genes lie near the center of the DNA sequence that is transferred and are identical whether the sequences are at the *MAT* locus or at one of the silent

loci. This arrangement indicates that the function of the transcription factors that interact with these sequences must somehow be blocked at *HML* and *HMR*, but not at the *MAT* locus. This repression of the silent loci depends on **silencer sequences** located next to the region of transferred DNA at *HML* and *HMR* (see Figure 8-31). If the silencer is deleted, the adjacent locus is transcribed. Remarkably, any gene placed near the yeast mating-type silencer sequence by recombinant DNA techniques is repressed, or "silenced," even a tRNA gene transcribed by RNA polymerase III, which uses a different set of general transcription factors than RNA polymerase II uses, as discussed later.

Several lines of evidence indicate that repression of the *HML* and *HMR* loci results from a condensed chromatin structure that sterically blocks transcription factors from interacting with the DNA. In one telling experiment, the gene encoding an *E. coli* enzyme that methylates adenine residues in the sequence GATC was introduced into yeast cells under the control of a yeast promoter so that the enzyme was expressed. Researchers found that GATC sequences within the *MAT* locus and most other regions of the genome in these cells were methylated but not those within the *HML* and *HMR* loci. These results indicate that the DNA of the silent loci is inaccessible to the *E. coli* methylase, and presumably to proteins in general, including transcription factors and RNA polymerase. Similar experiments conducted with various yeast histone mutants indicated that specific interactions involving the

histone tails of H3 and H4 are required for formation of a fully repressed chromatin structure. Other studies have shown that the telomeres of every yeast chromosome also behave like silencer sequences. For instance, when a gene is placed within a few kilobases of any yeast telomere, its expression is repressed. In addition, this repression is relieved by the same mutations in the H3 and H4 histone tails that interfere with repression at the silent mating-type loci.

Genetic studies led to identification of several proteins, RAP1 and three SIR proteins, that are required for repression of the silent mating-type loci and the telomeres in yeast. RAP1 was found to bind within the DNA-silencer sequences associated with *HML* and *HMR* and to a sequence that is repeated multiple times at each yeast-chromosome telomere. Further biochemical studies showed that the SIR2 protein is a *histone deacetylase*; it removes acetyl groups on lysines of the histone tails. Furthermore, the RAP1 and SIR2, 3, and 4 proteins bind to one another, and SIR3 and SIR4 bind to the N-terminal tails of histones H3 and H4, which are maintained in a largely nonacetylated state by the deacetylase activity of SIR2. A series of experiments using fluorescence confocal microscopy on yeast cells either stained with fluorescent-labeled antibody to any one of the SIR proteins or RAP1 or hybridized to a labeled telomere-specific DNA probe revealed that these proteins form large, condensed telomeric nucleoprotein structures resembling the heterochromatin found in higher eukaryotes (Figure 8-32a, b, c).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 8-32 Antibody and DNA probes colocalize SIR3 protein with telomeric heterochromatin in yeast nuclei. (a) Confocal micrograph 0.3 μm thick through three diploid yeast cells, each containing 68 telomeres. Telomeres were labeled by hybridization to a fluorescent telomere-specific probe (yellow). DNA was stained red to reveal the nuclei. The 68 telomeres coalesce into a much smaller number of regions near the nuclear periphery. (b, c) Confocal micrographs of yeast cells labeled with a telomere-specific

hybridization probe (b) and a fluorescent-labeled antibody specific for SIR3 (c). Note that SIR3 is localized in the repressed telomeric heterochromatin. Similar experiments with RAP1, SIR2, and SIR4 have shown that these proteins also colocalize with the repressed telomeric heterochromatin. (d) Schematic model of the silencing mechanism at yeast telomeres. (*Top left*) Multiple copies of RAP1 bind to a simple repeated sequence at each telomere region that lacks nucleosomes. SIR3 and SIR4 bind to RAP1, and SIR2 binds to SIR4. SIR2 is a histone deacetylase that deacetylates the tails on the histones neighboring the repeated RAP1-binding site. (*Middle*) The hypoacetylated histone tails are also binding sites for SIR3 and SIR4, which in turn bind additional SIR2, deacetylating neighboring histones. Repetition of this process results in spreading of the region of hypoacetylated histones with associated SIR2, SIR3, and SIR4. (*Bottom*) Interactions between complexes of SIR2, SIR3, and SIR4 cause the chromatin to condense and several telomeres to associate, as shown in (a)–(c). The higher order chromatin structure generated sterically blocks other proteins from interacting with the underlying DNA. See M. Grunstein, 1997, *Curr. Opin. Cell Biol*. **9**:383.

[Parts (a)–(c) Data from M. Gotta et al., *1996, J. Cell Biol.* **134**:1349–1363; https://doi.org/10.1083/jcb.134.6.1349.]

Figure 8-32d depicts a model for the chromatin-mediated silencing at yeast telomeres based on these and other studies. Formation of heterochromatin at telomeres is nucleated by multiple RAP1 protein molecules bound to repeated sequences in a nucleosome-free region at the extreme end of a telomere. A network of protein-protein interactions involving telomere-bound RAP1, three SIR proteins (2, 3, and 4), and hypoacetylated histones H3 and H4 creates a nucleoprotein complex that includes several telomeres and in which the DNA is largely inaccessible to external proteins. Recall that nucleosomes in regions of heterochromatin have low levels of acetylation on lysines of the histone tails, while nucleosomes in euchromatin have histone tails that are acetylated on several lysines

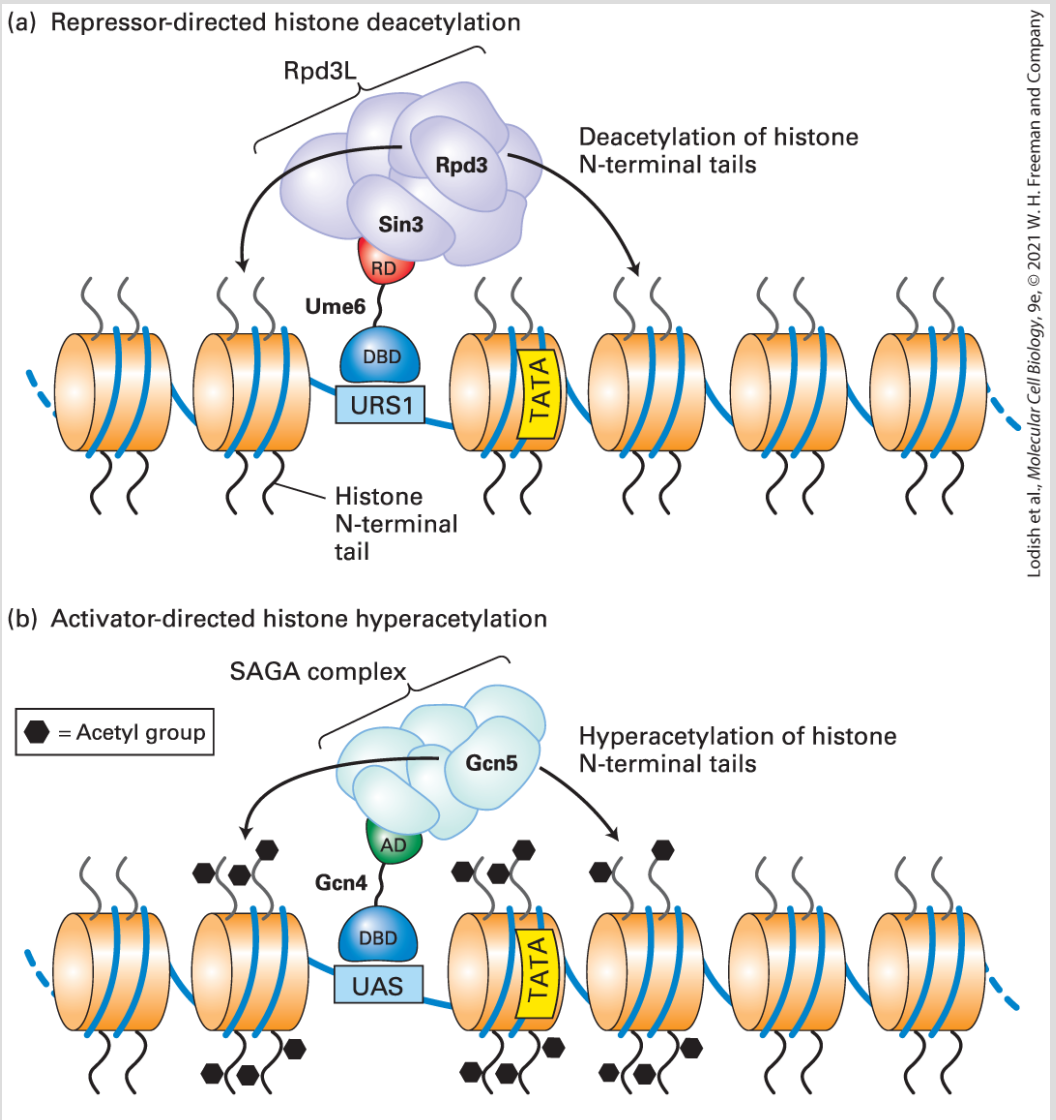
(see Figure 7-28b). One additional protein, SIR1, is also required for silencing of the mating-type loci. It binds to the silencer regions associated with *HML* and *HMR* together with RAP1 and other proteins to initiate assembly of a similar multiprotein-silencing complex that encompasses *HML* and *HMR*.

An important feature of this model is the dependence of repression on *hypoacetylation* of the histone tails. This was demonstrated in experiments with yeast mutants in which lysines in histone N-termini were replaced with arginines, glutamines, or glycines. Arginine is positively charged, like lysine, but cannot be acetylated. Glutamine and glycine are neutral and simulate the neutral charge of acetylated lysine. Repression at telomeres and at the silent mating-type loci was defective in mutants with glutamine or glycine substitutions for lysine in the H3 and H4 histone tails. Repression was normal in mutants with arginine substitutions due to the positive charge that arginine provides to the histone tails. Further, acetylation of H3 and H4 lysines interferes with binding by SIR3 and SIR4 and consequently prevents repression at the silent loci and telomeres. Finally, chromatin immunoprecipitation experiments (see Figure 8-12a) using antibodies against acetylated lysines at specific positions in the histone N-terminal tails (see Figure 8-26a) confirmed that histones in repressed regions near telomeres and at the silent mating loci are hypoacetylated but become hyperacetylated in *sir* mutants when genes in these regions are derepressed.

Repressors Can Direct Histone Deacetylation at Specific Genes

The importance of **histone deacetylation** in chromatin-mediated gene repression was further supported by studies of eukaryotic repressors that regulate genes within chromosome arms, distant from telomeres and centromeres. These repressor proteins act in part by causing deacetylation of histone tails in nucleosomes that encompass the TATA box and promoter-proximal region of the genes they repress. In vitro studies have shown that when promoter DNA is part of a nucleosome with nonacetylated histones, the general transcription factors cannot bind to the TATA box and promoter-proximal region. In nonacetylated histones, the N-terminal lysines are positively charged and may interact with phosphates in the DNA backbone. Nonacetylated histone tails also interact with neighboring histone octamers and other chromatin-associated proteins, favoring the folding of chromatin into condensed regions (see Figures 7-24b and 7-27a). The net effect is that general transcription factors cannot assemble into a preinitiation complex on a promoter associated with hypoacetylated histones. In contrast, binding of general transcription factors is repressed much less by histones with hyperacetylated tails, in which the positively charged lysines are neutralized and electrostatic interactions within chromatin are eliminated.

The connection between histone deacetylation and repression of transcription at specific yeast promoters became clearer when the cDNA encoding a human histone deacetylase was found to have high homology to the yeast *RPD3* gene, known to be required for the normal repression of a number of yeast genes. Further work showed that the yeast Rpd3 protein has histone deacetylase activity. The ability of Rpd3 to deacetylate histones at a number of promoters depends on two other proteins: Ume6, a repressor that binds to a specific upstream regulatory sequence (URS1), and Sin3, which is part of a large multiprotein complex called *Rpd3L* that also contains Rpd3 (Figure 8-33a). Sin3 also binds to the repression domain of Ume6, thus positioning the Rpd3 histone deacetylase in the complex so that it can interact with nearby promoter-associated nucleosomes and remove acetyl groups from histone-tail lysines. Additional experiments, using the chromatin immunoprecipitation technique outlined in Figure 8-12a and antibodies to specific histone acetylated lysines, demonstrated that in wild-type yeast, one or two nucleosomes in the immediate vicinity of Ume6-binding sites are hypoacetylated. These sites include the promoters of genes repressed by Ume6. In *sin3* and *rpd3* deletion mutants, not only were these promoters derepressed, but the nucleosomes near the Ume6-binding sites were hyperacetylated. All of these findings provide considerable support for the model of repressor-directed deacetylation shown in Figure 8-33a.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 8-33 Proposed mechanism of histone deacetylation and hyperacetylation in yeast transcriptional control. (a) Repressor-directed deacetylation of histone N-terminal tails. The DNA-binding domain (DBD) of the repressor Ume6 interacts with a specific upstream control element of the genes it regulates, called *URS1*. The Ume6 repression domain (RD) binds Sin3, a subunit of a multiprotein complex that includes Rpd3, a histone deacetylase. Deacetylation of histone N-terminal tails on nucleosomes in the region of the Ume6-binding site inhibits binding of general transcription factors at the TATA box, thereby repressing gene expression. (b) Activator-directed hyperacetylation of histone N-terminal tails. The DNA-binding domain of the activator Gcn4 interacts with specific upstream activating sequences (UAS) of the genes it regulates. The Gcn4 activation domain (AD) then interacts with a

multiprotein histone acetylase complex that includes the Gcn5 catalytic subunit. Subsequent hyperacetylation of histone N-terminal tails on nucleosomes in the vicinity of the Gcn4-binding site facilitates access by the general transcription factors required for initiation. Repression and activation of many genes in higher eukaryotes occur by similar mechanisms.

In yeast, the Sin3-Rpd3 complex (Rpd3L) functions as a **co-repressor**. Co-repressor complexes containing histone deacetylases have also been found associated with many repressors from mammalian cells (see Repressors Are the Functional Converse of Activators). Some of these complexes contain the mammalian homolog of Sin3 (mSin3), which interacts with the repression domain of repressors, as in yeast (see Figure 8-33a). Other histone deacetylase complexes identified in mammalian cells contain additional or different subunits that bind the repression domains of other eukaryotic repressors. These various repressor–co-repressor combinations mediate histone deacetylation at specific promoters by a mechanism similar to the yeast mechanism (see Figure 8-33a). In addition to repressing transcription through the formation of closed chromatin structures, some repression domains have also been found to inhibit the assembly of preinitiation complexes in experiments with purified general transcription factors in the absence of histones. This activity probably also functions in vivo, contributing to the repression of transcription by these repression domains.

Activators Can Direct Histone Acetylation at Specific Genes

Just as repressors function through co-repressors that bind to their repression domains, the activation domains of DNA-binding activators function by binding multisubunit **co-activator** complexes. One of the first co-activator complexes to be characterized was the yeast *SAGA complex*, which functions with the Gcn4 and Gal4 activator proteins described in Section 8.3. Early genetic studies indicated that full activity of the Gcn4 activator required a protein called *Gcn5*. The clue to Gcn5's function came from biochemical studies of a *histone acetylase* purified from the protozoan *Tetrahymena*, the first histone acetylase to be purified. Sequence analysis revealed homology between the *Tetrahymena* protein and yeast Gcn5, which was soon shown to have histone acetylase activity as well. Further genetic and biochemical studies revealed that Gcn5 is one subunit of the SAGA multiprotein co-activator complex. Another subunit of this histone acetylase complex binds to activation domains in multiple yeast activator proteins, including Gcn4. The model shown in Figure 8-33b is consistent with the observation that nucleosomes near the promoter region of a gene regulated by the Gcn4 activator are specifically hyperacetylated compared with most histones in the cell. This activator-directed hyperacetylation of nucleosomes near a promoter region facilitates binding of other proteins required for transcription initiation.

In addition to leading to the decondensation of chromatin, the acetylation of specific histone lysines generates binding sites for proteins containing **bromodomains**. A bromodomain is a sequence of about 110 amino acids that folds into a domain, which binds acetylated lysine. Bromodomains are found in several chromosome-associated proteins that contribute to transcriptional activation. For example, a subunit of the general transcription factor TFIID contains two bromodomains, which bind with high affinity to acetylated nucleosomes. Recall that TFIID binding to a promoter initiates assembly of an RNA polymerase II preinitiation complex (see Figure 8-13). Nucleosomes at promoter regions of virtually all active genes have acetylated lysines in their H3 and H4 histone tails.

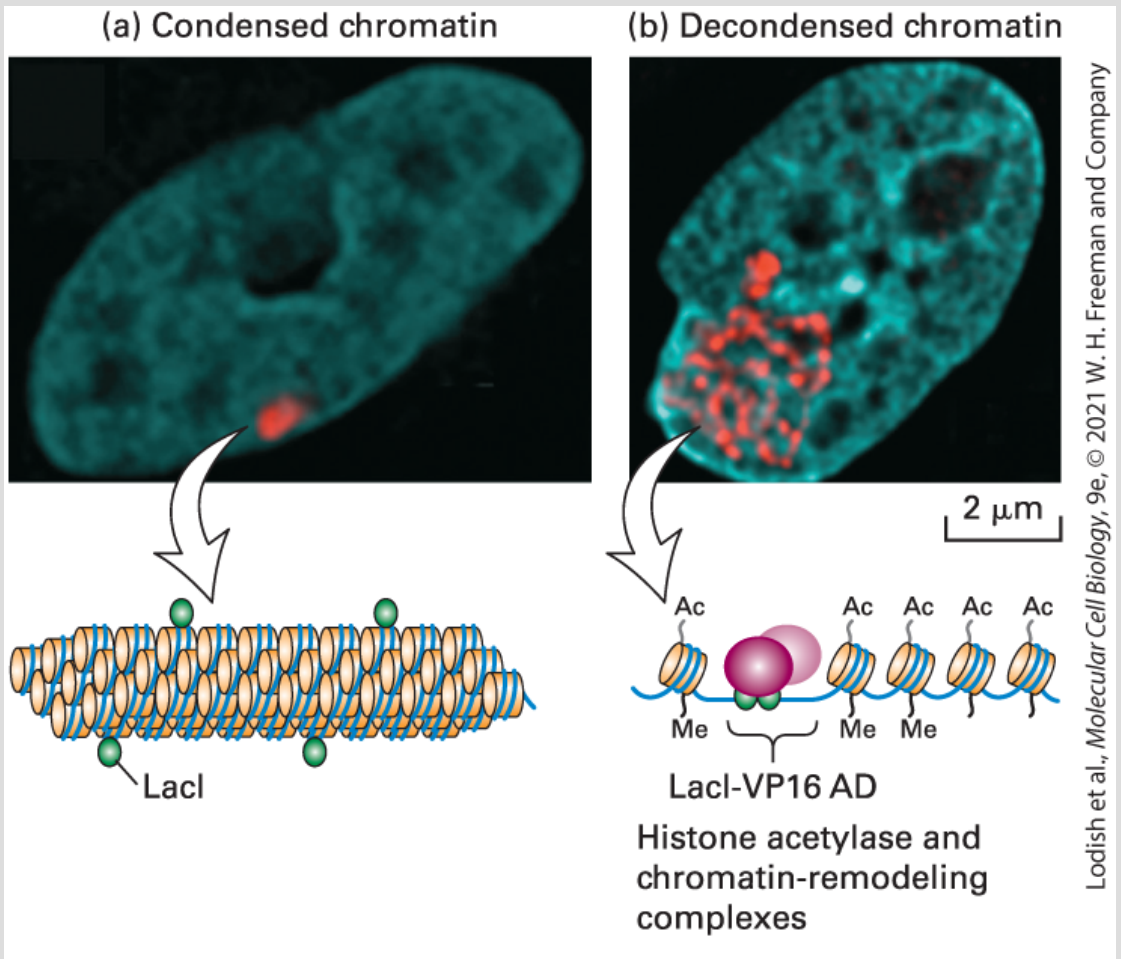
A similar activation mechanism operates in higher eukaryotes. Mammalian cells contain multisubunit histone acetylase co-activator complexes that are homologous to the yeast SAGA complex. Mammalian cells also express two related 300-kDa, multidomain proteins called *CBP* and *p300*, which acetylate histone tail lysines. As noted earlier, one domain of CBP binds the phosphorylated acidic activation domain in the CREB transcription factor (Figure 8-26a). Other domains of CBP interact with activation domains in other activators. CBP also has histone acetylase activity and the ability to associate with additional multisubunit histone acetylase complexes. CREB and many other mammalian activators function in part by directing CBP and p300 and their associated histone acetylase complexes to specific nucleosomes, where they

acetylate histone tails, thereby facilitating the interaction of general transcription factors with promoter DNA.

Chromatin-Remodeling Complexes Help Activate or Repress Transcription

In addition to histone acetylase complexes, multiprotein *chromatin-remodeling complexes* are required for activation at many promoters. The first of these complexes to be characterized was the yeast **SWI/SNF chromatin-remodeling complex**. One of the SWI/SNF subunits has homology to DNA helicases, enzymes that use energy from ATP hydrolysis to disrupt interactions between base-paired nucleic acids or between nucleic acids and proteins. The SWI/SNF complex is thought to pump or push DNA into the nucleosome so that DNA bound to the surface of the histone octamer transiently dissociates from the surface and translocates, causing the nucleosomes to slide along the DNA. This chromatin remodeling facilitates binding of transcription factors to specific DNA sequences in chromatin. Many activation domains bind to *chromatin-remodeling complexes*, which stimulate in vitro transcription from chromatin templates. Thus the SWI/SNF complex represents another type of co-activator complex. The experiment shown in Figure 8-34 demonstrates dramatically how an activation domain can cause decondensation of a region of chromatin. This decondensation

results from association of the activation domain with chromatin-remodeling and histone acetylase complexes.



EXPERIMENTAL FIGURE 8-34 Expression of fusion proteins demonstrates chromatin decondensation in response to an activation domain. A cultured hamster cell line was engineered to contain multiple copies of a tandem array of *E. coli lac* operator sequences integrated into a chromosome in a region of heterochromatin. (a) When an expression vector for the *lac* repressor (LacI) was transfected into these cells, *lac* repressor bound to the *lac* operator sites could be visualized in a region of condensed chromatin using an antibody against the *lac* repressor (red). DNA was visualized by staining with DAPI (blue), revealing the nucleus. A diagram of condensed chromatin is shown below. (b) When LacI fused to an activation domain was transfected into these cells, staining as in (a) revealed that the activation domain causes this region of chromatin to decondense into a thinner

chromatin fiber that fills a much larger volume of the nucleus. A diagram of a region of decondensed chromatin with bound LacI fusions to the VP16 activation domain (AD) and associated chromatin remodeling and histone acetylase complexes is shown below.

[From T. Tumbar, G. Sudlow, and A. S. Belmont, 1999, *J. Cell. Biol.* **145**(7):1341–1354: https://doi.org/10.1083/jcb.145.7.1341]

Chromatin-remodeling complexes are required for many processes involving DNA in eukaryotic cells, including transcriptional control, DNA replication, recombination, and DNA repair. Several types of chromatin-remodeling complexes are found in eukaryotic cells, all with homologous DNA helicase domains. SWI/SNF complexes and related chromatin-remodeling complexes in multicellular organisms contain subunits with bromodomains that bind to acetylated histone tails. Consequently, SWI/SNF complexes continue to associate with activated, acetylated regions of chromatin, presumably maintaining them in a decondensed conformation. Chromatin-remodeling complexes can also participate in transcriptional repression. These complexes bind to the repression domains of repressors and contribute to repression, presumably by folding chromatin into condensed structures. Much remains to be learned about how this important class of proteins alters chromatin structure to influence gene expression and other processes.

Pioneer Transcription Factors Initiate the Process of Gene Activation During

Cellular Differentiation

As cells differentiate during embryogenesis and during differentiation from stem cells in adult organisms (see Chapter 22), many of the genes induced during the process are initially in repressed regions of heterochromatin in undifferentiated progenitor cells. Activation of these genes requires that the chromatin environment of their transcription-control regions become decondensed so that transcription factors can bind to enhancers and promoter-proximal control elements and so that the general transcription factors and Pol II can bind to promoters. In many cases, this decondensation is initiated by special *pioneer transcription factors* that can bind to their cognate binding sites in DNA even when those sites are within repressed heterochromatic regions of chromatin. These pioneer transcription factors bind to the surface of both the DNA and the histones in a nucleosome, and in some cases, such as for Sox2, use the binding energy to begin to unwrap the DNA from the nucleosome surface (Figure 8-35). This allows these pioneer factors to bind to their specific binding sites while the DNA is wrapped around a histone octamer with the opposite side of the DNA against the surfaces of histones.

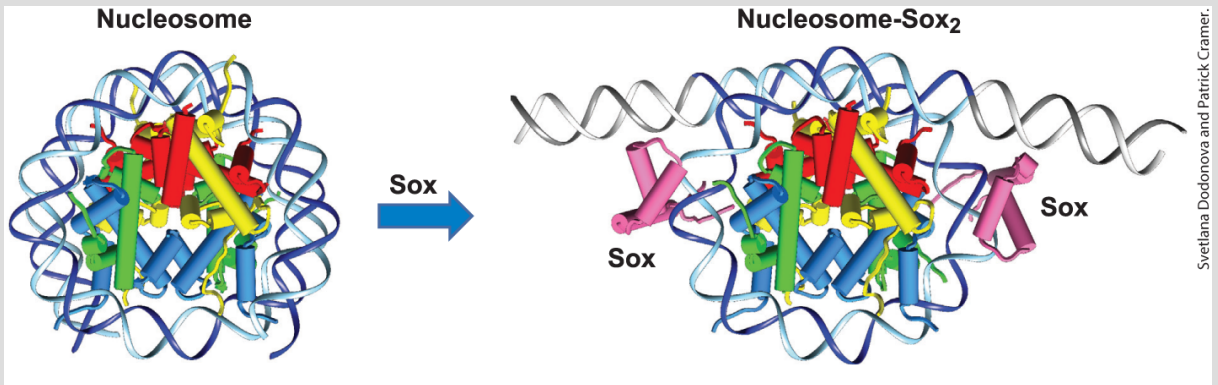


FIGURE 8-35 Structure of the Sox11-nucleosome complex. Like Sox2, the closely related transcription factor Sox11 can bind its cognate-binding sites even when they are wrapped around the surface of a nucleosome. *Left:* Model of a nucleosome with two Sox11-binding sites, one inside the beginning of the DNA 1 2/3 wraps around the surface of the nucleosome, and the other near the end of the DNA wrap around the nucleosome. Histone α helices are shown as cylinders with H2A subunits yellow; H2Bs red; H3s blue; and H4s green. *Right:* When Sox11 binds its two DNA-binding sites in this nucleosome, it unwinds the DNA from the surface of the nucleosome, making it available for interacting with other proteins.

An example of pioneer transcription factors that activate transcription from a promoter in condensed chromatin involves the liver-specific gene *Alb1*, encoding serum albumin, a major constituent of blood serum that is secreted into the blood by hepatocytes, the major cell type in the liver. In the developing mouse, the FoxA and GATA-4 or GATA-6 transcription factors are the first to bind to an *Alb1* enhancer in undifferentiated gut endodermal cells destined to develop into the liver. FoxA has a *winged helix* DNA-binding domain that binds to one side of the DNA helix containing the FoxA-binding site. GATA factors are also able to bind to their specific sites in DNA when those sites are included in nucleosomal DNA wrapped around a histone octamer. The FoxA and GATA-4/6

activation domains may then interact with chromatin remodeling complexes and histone acetylase complexes to decondense the chromatin of the 120-bp *Alb1* enhancer, allowing the observed subsequent binding of four additional transcription factors in the nascent liver bud that develops later.

The Mediator Complex Forms a Molecular Bridge Between Activation Domains and Pol II

Once the interaction of activation domains with histone acetylase complexes and chromatin-remodeling complexes decondense the chromatin of a promoter region to an "open" structure that allows the binding of general transcription factors, activation domains of the transcription factors bound to DNA control elements interact with another multisubunit co-activator complex, the Mediator complex (Figure 8-36). Activation domain–Mediator interactions stimulate assembly of a preinitiation complex on a promoter. Recent cryoelectron microscopy studies show that the head and middle domains of the Mediator complex interact directly with Pol II. Several Mediator subunits bind to activation domains in various activator proteins. Thus Mediator can form a molecular bridge between an activator bound to its cognate site in DNA and Pol II bound to a promoter.

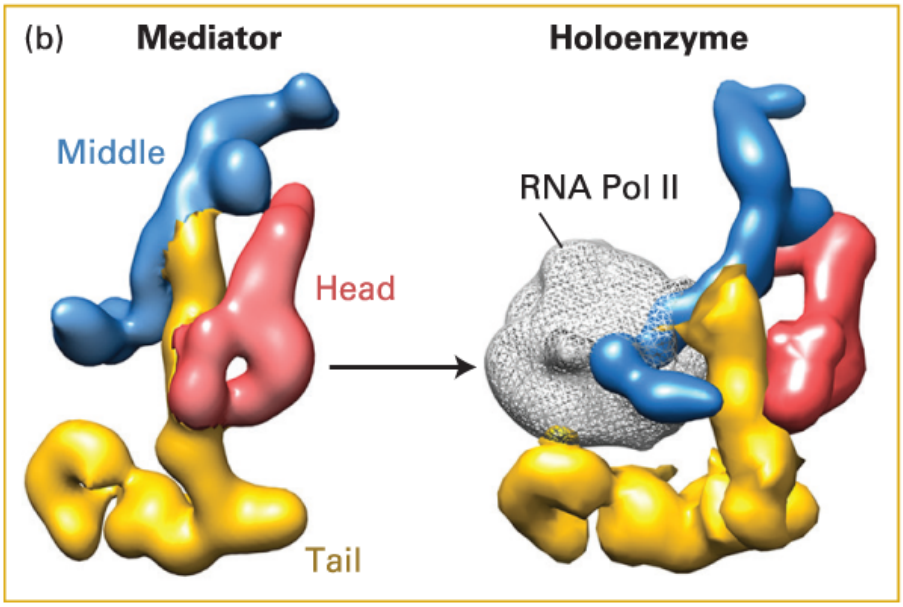
(a)

Yeast

Head	Middle	Tail	CKM
Med6	Med1	Med2	Med12
Med8	Med4	Med3	Med13
Med11	Med7	Med5	Cdk8
Med17	Med9	Med14	CycC
Med18	Med10	Med15	
Med20	Med19	Med16	
Med22	Med21		
	Med31		

Human

Head	Middle	Tail	CKM
MED6	MED1	MED14	MED12/12L
MED8	MED4	MED15	MED13/13L
MED11	MED7	MED16	Cdk8/CDK19
MED17	MED9	MED23	CycC
MED18	MED10	MED24	
MED20	MED19	MED25	
MED22	MED21		
MED27	MED31		
MED28	MED26		
MED29			
MED30			



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 8-36 Structure of yeast and human Mediator complexes. (a) Subunits of the *S. cerevisiae* and human Mediator complexes. The subunits constituting the head, middle, and tail modules of Mediator are indicated, as well as the subunits of the CDK8-kinase module (CKM) that associates with some Mediator complexes, blocking Pol II binding. (b) Cryoelectron microscopic structure of the yeast Mediator without the CKM. (*Left*) The head, middle, and tail modules composed of the subunits listed above are color-coded. (*Right*) The structure of a complex of Mediator with Pol II, called the *holoenzyme*, suggests that the Mediator modules rotate relative to one another as shown to create a surface that binds Pol II.

[Part (b) Republished with permission from Elsevier, from K. L. Tsai, 2014, “Subunit Architecture and Functional Modular Rearrangements of the Transcriptional

Mediator Complex." *Cell*, **157**(6): 1430–1444; permission conveyed through Copyright Clearance Center, Inc.]

Experiments with temperature-sensitive yeast mutants indicate that some Mediator subunits are required for transcription of virtually all yeast genes. These subunits help maintain the overall structure of the Mediator complex or bind to Pol II. In contrast, other Mediator subunits are required for activation or repression of specific subsets of genes. RNA-seq analysis (see Chapter 6) has been used to examine gene expression in yeast mutants with defects in nonessential Mediator subunits. Results indicate that each mutation influences transcription of 3–10 percent of all genes, meaning that the deletion either increases or decreases mRNA expression by a factor of twofold or more. In many cases, these Mediator subunits interact with specific activation domains. Thus when one Mediator subunit is defective, transcription of genes regulated by activators that bind to that subunit is severely depressed, but transcription of other genes is unaffected. Recent cryoelectron microscopy studies suggest that when activation domains interact with Mediator, the head, middle, and tail domains depicted in Figure 8-36b rotate relative to one another, creating a binding surface for RNA polymerase II. The polymerase-Mediator complex is referred to as the *holoenzyme*. The surfaces of the polymerase that interact with general transcription factors in the preinitiation complex (see Figure 8-13) remain exposed in the proposed model of the polymerase-Mediator complex.

The various experimental results indicating that individual Mediator subunits bind to specific activation domains suggest that multiple activators may influence transcription from a single promoter by interacting with a Mediator complex simultaneously or in rapid succession (Figure 8-37). Activators bound at enhancers or promoter-proximal elements can interact with Mediator associated with a promoter because chromatin, like DNA, is flexible and can form a loop, bringing the regulatory regions and the promoter close together (see Figure 8-30). The multiprotein complexes that form on eukaryotic promoters may comprise more than 100 polypeptides with a total mass of 3–5 megadaltons (MDa), which is as large as a ribosome.

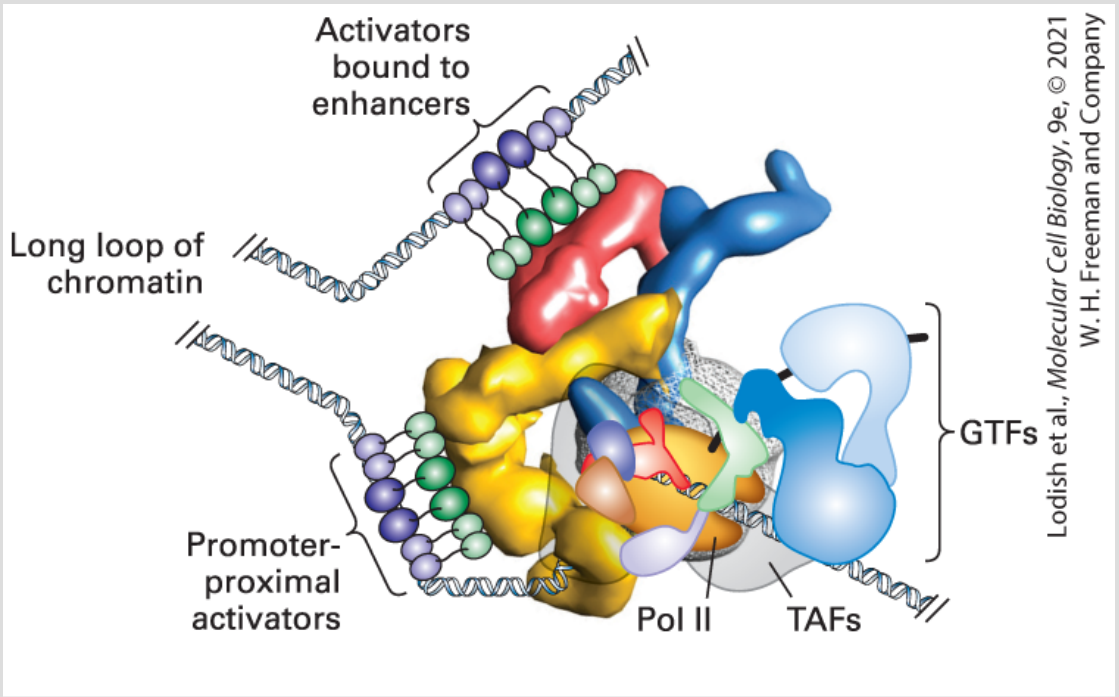


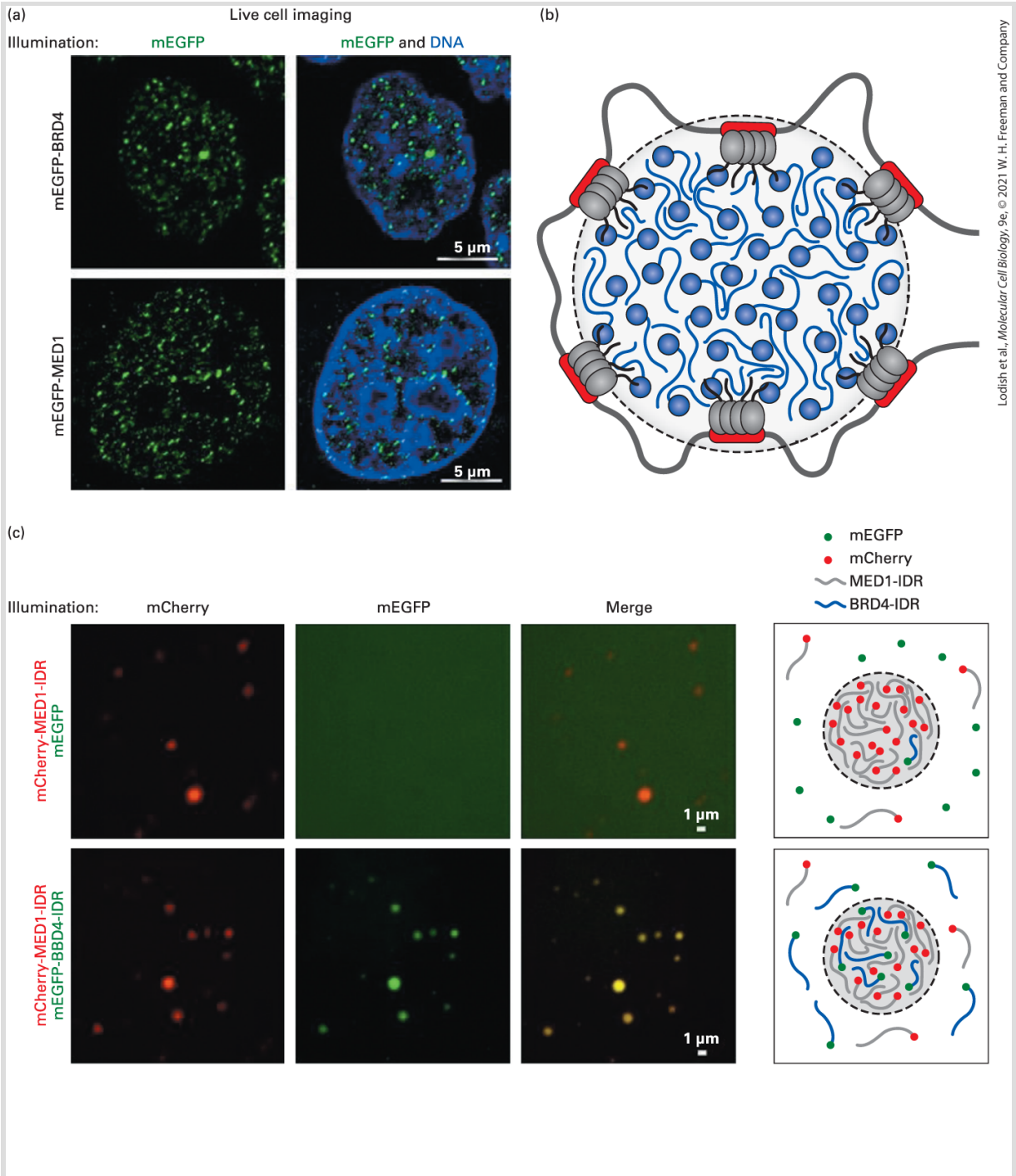
FIGURE 8-37 Model of several DNA-bound activators interacting with a single Mediator complex. The ability of different Mediator subunits to interact with specific

activation domains may contribute to the integration of signals from several activators at a single promoter. See the text for discussion.

Transcriptional Condensates Greatly Increase the Rate of Transcription Initiation

Recently developed high resolution microscopic methods have revealed that proteins involved in transcription, such as Mediator, often co-localize in nuclei in puncta that are on the order of 0.5–0.1 μm in diameter — far larger than the diameter of a chromatin fiber. BRD4, a protein with four bromodomains that stimulates transcription elongation in promoter proximal regions where histones are highly acetylated, also forms similar puncta in vivo (Figure 8-38a). These regions of high protein concentrations contain hundreds to thousands of Mediator and BRD4 molecules. The punctate structures are not delimited by a membrane but are continuous with the surrounding nucleoplasm. Their formation is dependent on intrinsically disordered regions (IDRs) of nuclear proteins, including transcriptional activators, repressors, and co-activators (see Figure 8-23). Multiple short regions of polypeptide in these IDRs form weak interactions with short peptides in the IDRs of other nuclear proteins, causing what is referred to as “condensation” of multiple proteins into a “transcriptional condensate” (diagrammed in Figure 8-38b). These interactions are transient, but

the large number of these possible multivalent interactions causes the proteins to associate in a liquid-like state. Consequently, the process that forms these protein concentrates has been called a "liquid-liquid phase separation."



EXPERIMENTAL FIGURE 8-38 Phase separation of co-activators compartmentalizes and concentrates the transcription apparatus. (a) Immunofluorescence (IF) imaging of BRD4 and MED1 in mouse embryonic stem cells (mESCs). Fluorescence signal is shown alone (*left*) and merged with Hoechst stain for DNA (*right*). (b) Super-enhancers are clusters of enhancers (red) bound by master transcription factors (grey ovals) that concentrate high densities of co-activators (blue circles) and the transcription apparatus to drive robust expression of genes that play prominent roles in cell identity. This is achieved by the phase separation of co-activators, which is driven in part by high-valency and low-affinity interactions of intrinsically disordered regions (blue lines). (c) The indicated mEGFP or mCherry fusion proteins were mixed at 10 mM each in a buffer containing 10 percent Ficoll-400 and 125 mM NaCl. Indicated fluorescence channels are presented for each mixture. Illustrations summarizing results are shown on the right.

[Republished with permission of American Association for the Advancement of Science, from B. R. Sabari et al., 2018, "Coactivator Condensation at Super-Enhancers Links Phase Separation and Gene Control," *Science* 361(6400): eaar3958; permission conveyed through Copyright Clearance Center, Inc.]

It has been possible to observe and study the formation of these protein condensates in vitro with purified proteins. Figure 8-38c shows an experiment with a purified fusion protein consisting of an ~200 amino acid–serine-rich IDR in the largest subunit of Mediator (MED1) fused to the red fluorescent protein called *mCherry* to allow the fusion protein to be detected by fluorescence microscopy. When incubated along with green-fluorescent protein that was not fused to an IDR in a buffer that promotes protein condensation, the mCherry-MED1-IDR became concentrated in droplets similar in size to the condensates observed in vivo (Figure 8-38c, *top*), while GFP was not concentrated in the droplets. On the other hand, when the same mCherry-MED1-IDR fusion protein was incubated with GFP

that was fused to an IDR from BRD4, the two fusion proteins were included in the same droplets (Figure 8-38c, *bottom*). This result shows that proteins with different IDRs can be concentrated in the same protein condensates. Further immunofluorescent staining has shown that condensates formed in a nuclear extract visualized with GFP-BRD4 and GFP-MED1 as in Figure 8-38a also contain hundreds to thousands of copies of many other proteins involved in transcription including RNA polymerase II with a hypo-phosphorylated CTD, the form of Pol II that can assemble into a preinitiation complex (see Figure 8-13). When IDRs required for formation of transcriptional condensates are deleted from an activator, the transcriptional activation function of the mutant activator is greatly diminished.

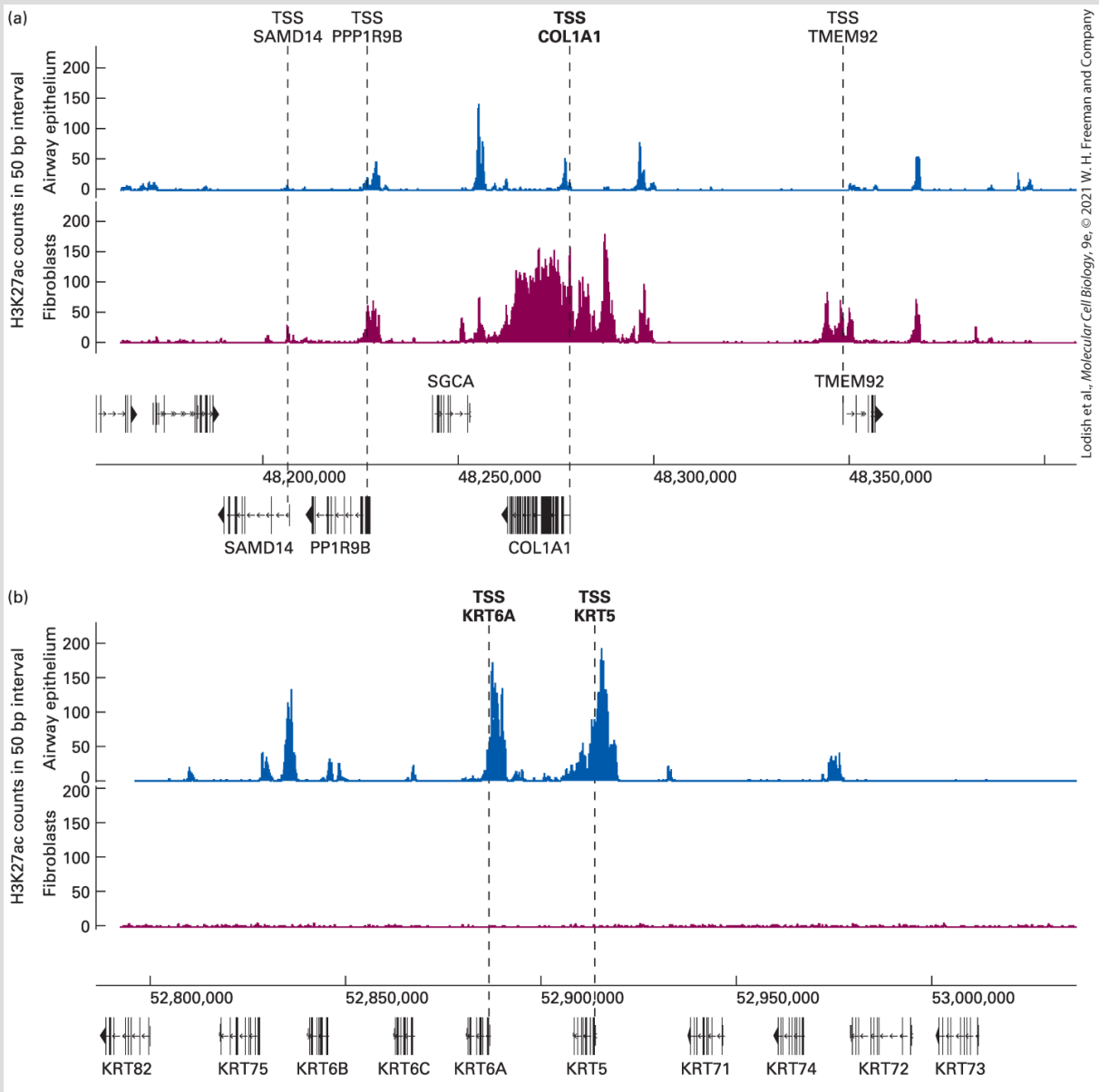
Super-Enhancers

Characterization of all the enhancers in cultured mammalian cells revealed that a small fraction of them are associated with regions containing multiple, closely spaced enhancers associated with highly acetylated histones over a region of 10 or more kilobases. These regions termed *super-enhancers* include ∼1.5–4 percent of all enhancers, depending on the cell type. Unlike typical enhancers, super-enhancers contain longer regions of DNA bound by transcription factors and their associated co-activators. This long region of bound transcription factors, many with IDRs, promotes the formation of transcriptional condensates on super-enhancers.

This leads to concentrations of proteins involved in transcription initiation within the condensate that are twenty- to a hundredfold higher than in other regions of the nucleoplasm, resulting in a huge increase in the rate of assembly of the multicomponent preinitiation complex, contributing greatly to the high rate of transcription induced by super-enhancers.

Super-enhancers often activate key cell identity genes expressed at high level that are important for normal mammalian development. For example, in fibroblasts, which secrete extremely high levels of collagens to provide tensile strength and flexibility to the extracellular matrix (see Chapter 20), a super-enhancer of ∼30–50 kb forms over the collagen 1A1 gene (*COL1A1*) and its upstream region (Figure 8-39a). The closely related histone acetylase co-activators p300 and CBP associate with chromatin throughout this region, presumably by interacting with multiple activators bound to their specific cognate-binding sequences. CBP/p300 acetylate histone H3 lysines (K in the one letter amino acid code) at positions 18 and 27. Acetylation at H3K27 was detected by chromatin immunoprecipitation with antibodies specific for H3 acetylated at K27 (Figure 8-39a, *top*). The length of the acetylated region associated with *COL1A1* was much longer than other regions of H3K27ac in this region. In contrast to fibroblasts, which express high levels of COL1A1, airway epithelial cells do not express high levels of COL1A1 and do not form a super-enhancer over the *COL1A1* gene (Figure 8-39a, *top*). Rather, in these epithelial cells, a super-

enhancer encompassing a 13-kb region of chromatin formed over the *KRT5* gene, resulting in high level expression of this protein that gives tensile strength and flexibility to a sheet of epithelial cells (see Chapter 20). A broad region of H3K27 acetylation was also associated with the *KRT6A* gene, expressed at lower but still substantial level in these cells, and small peaks of H3K27ac were associated with the *KRT6B* and *KRT6C* genes expressed at low level. None of the other keratin genes in this gene cluster were expressed in these cells, and consistently, they did not have peaks of H3K27ac associated with their TSSs, indicating that these promoters were not active in these cells (Figure 8-39b).



EXPERIMENTAL FIGURE 8-39 The COL1A1 and KRT 5 and 6A super-enhancers. (a) *Bottom:* Map of genes on human chromosome 17 between 48,200,000 and 48,400,000 base pairs from the left end, represented by the horizontal line demarcated in 50,000 base-pair intervals. Genes transcribed to the right are represented above the line; genes transcribed to the left, below the line. Exons are represented by vertical lines, with arrowheads indicating the $3'$ exon. Dashed vertical lines indicate the positions of transcription start sites (TSS) for genes expressed in airway epithelial cells. *Top:* Chromatin immunoprecipitation data (see Figure 8-12) for histone H3 lysine 27 acetylation (H3K27ac) from cultured human fibroblasts or airway epithelial cells, as indicated at the left. The data is reported as the number of times a sequence in each 50-bp interval along the chromosome was observed in cross-

linked chromatin fragments immunoprecipitated with antibody specific for an H3 peptide acetylated at lysine 27. Note that most peaks of H3K27ac were <5 kb in length. In contrast, COL1A1 is associated with a much longer, $\sim 30–50$-kb region of chromatin extensively acetylated on H3K27. This unusually long region of H3K27ac identifies this as a super-enhancer. (b) *Bottom:* Map of genes on human chromosome 12 between 52,800,000 and 53,000,000 base pairs from the left end. Dashed vertical lines indicate positions of TSSs for genes expressed in fibroblasts. *Top:* Chromatin immunoprecipitation data for histone H3K27ac from cultured human fibroblasts or airway epithelial cells.

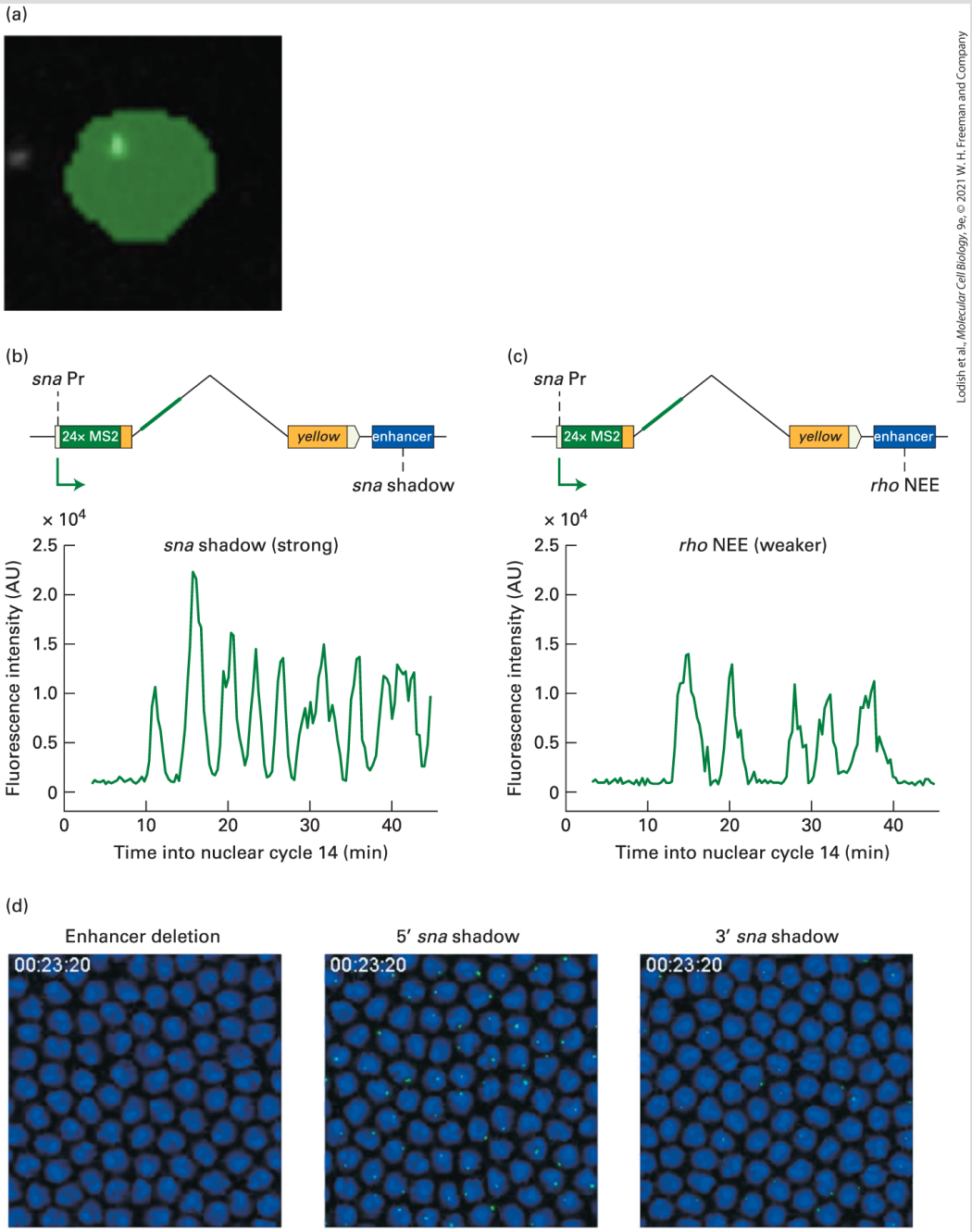
[Data from R. Ferrari et al., 2014, *Cell Host Microbe* **16**:663–676.]

In situ hybridization experiments have shown that transcriptional condensates (see Figure 8-38a) associate with the super-enhancers that are active in a particular cell type. Presumably, the high concentration of proteins involved in transcription initiation within these transcriptional condensates results in the very high rates of transcription initiation associated with super-enhancers.

Transcription Occurs in Bursts

High-resolution video microscopy recently revealed that many genes expressed at high level are not continuously transcribed at high rates. Rather, transcription initiation from highly transcribed genes occurs in bursts of multiple initiation events separated by periods of no transcription between transcriptional bursts. This was observed in *Drosophila* embryos containing constructs of reporter genes attached to multiple copies of engineered stem loops in their 5′-untranslated sequence. The stem loops bind to bacteriophage coat

proteins with very high affinity and specificity. In the experiment presented in Figure 8-40, 24 copies of the binding site for the bacteriophage MS2-coat protein were inserted immediately downstream of the promoter region of the *Drosophila snail (sna)* gene, which is expressed in the early *Drosophila* embryo. When transgenic flies were generated containing this reporter inserted into one site on one chromosome, no expression of the transgene was detected. But transcription was detected when the inserted DNA contained an enhancer downstream of the reporter gene, an enhancer that activates expression of *sna* in the early *Drosophila* embryo (Figure 8-40b). Transcription was detected by expressing MS2-coat protein fused to GFP (green fluorescent protein) in the same cells in which the reporter gene construct was inserted. When the 5′-end of the reporter gene was transcribed, MS2-coat protein-GFP bound to the 24 MS2-coat protein-binding sites, generating a signal that was visible by confocal fluorescence microscopy (Figure 8-40a). A video of images of the embryo during the 14th nuclear cycle (∼60 min long) revealed one region of GFP fluorescence in each cell at the site of the transcribed, inserted reporter gene. But the intensity of the GFP signal in each cell was not continuous. Rather, the intensity of each spot of fluorescence increased and decreased over several minutes (Figure 8-40d and associated movie, whose link is in the Figure 8-40 legend).



EXPERIMENTAL FIGURE 8-40 Highly expressed genes are transcribed in bursts. (a) Transcriptional focus detected by binding of MS2-coat protein-yellow fluorescent protein (MCP-YFP) to the 24 coat protein binding sites at the 5′-end of the RNA transcribed from the reporter gene. (b) *Top:* Schematic representation of reporter gene containing 100-bp of the *sna* promoter spanning the transcription start site and 24× MS2 RNA stem loops within the 5′ UTR. The strong 1.5-kb *sna* "shadow" enhancer was placed 7.5 kb downstream.

Bottom: Trace of fluorescence intensity over time of signal in an individual nucleus. (c) Same as part (b) but with the *rho* NE enhancer substituted for the *sna* shadow enhancer. (d) Frames from movies of transcriptional bursting. "$3'$ *sna* shadow" (*right*) shows yellow fluorescence bursts (which appear green) in living *Drosophila* embryos in which all cells have the integrated reporter gene indicated in part (b). The time shown (23 minutes and 20 seconds) is the time elapsed since the 14th nuclear division. "Enhancer deletion" (*left*) is the reporter construct in (b) but with no enhancer inserted. "$5'$ *sna* shadow" (*middle*) has the *sna* shadow enhancer inserted 1 kb upstream of the *sna* promoter. To watch the movie, go to https://www.cell.com/cms/10.1016/j.cell.2016.05.025/attachment/ 2184326d-3b6c-47d6-8121-c00c36c3ebeb/mmc2. The movie starts at nuclear cycle (nc) 14 of embryogenesis and continues until gastrulation. The cells also express a fusion of RFP to histone H2A shown in blue, as a marker for the nucleus. Anterior is to the left and ventral is facing up.

[Republished with permission by Elsevier, from T. Fukaya, B. Lim, and M. Levine (2016) Enhancer Control of Transcriptional Bursting. *Cell* **166** (2): 358–368; permission conveyed through Copyright Clearance Center, Inc.]

The intensity of a single fluorescent focus in one nucleus was plotted versus time for this construct with the strong *sna* "shadow" enhancer and for a second construct with the weaker *rho* NEE enhancer (Figures 8-40b and c, respectively). Bursts of fluorescence were observed extending for variable periods of time averaging $\sim$4 min (the width of the peaks). Based on the fluorescence intensity, each burst is thought to have resulted from 20–100 transcription initiation events over the $\sim$4 min period of a single transcriptional burst. For the construct with the strong *sna* shadow enhancer, bursts occurred frequently; for the construct with the weaker *rho* NEE enhancer, bursts of approximately the same height resulting from

20–100 transcription initiations also were observed, but the bursts occurred less frequently (Figure 8-40b, c).

Similar observations have been made for other *Drosophila* enhancers. In most cases, the strength of an enhancer is related primarily to how frequently transcriptional bursts are initiated. Additional research is needed to fully understand the mechanism(s) underlying transcriptional bursting and its regulation by enhancers. However, in vitro transcription experiments with yeast extracts indicate that after the first round of transcription from an added template DNA, TBP and possibly other general transcription factors (GTFs) remain associated with promoter DNA forming a partially assembled preinitiation complex (PIC) scaffold that allows very rapid assembly of the next preinitiation complex and subsequent initiation by Pol II. These observations and other in vitro transcription experiments with purified factors have led to the model that an interaction between enhancer-bound factors, Mediator, and transcription factors bound to promoter-proximal elements promotes the assembly of a Pol II preinitiation complex that initiates, leaving behind a scaffold of TBP bound to the promoter that can reassemble and initiate transcription up to 100 times before the scaffold dissociates from the promoter. This is postulated to result in the observed bursts of transcription. Strong enhancers are proposed to increase the frequency of these bursts by stimulating the rate of assembly of general transcription factors into the first PIC that initiates the transcriptional burst, while weak

enhancers are proposed to cause the assembly of PICs on the regulated promoter at a lower frequency.

KEY CONCEPTS OF SECTION 8.4

Molecular Mechanisms of Transcription Repression and Activation

- Eukaryotic transcription activators and repressors exert their effects largely by binding to multisubunit co-activators or co-repressors that influence the assembly of preinitiation complexes either by modulating chromatin structure or by interacting with Pol II and general transcription factors.
- The DNA in condensed regions of chromatin (heterochromatin) is relatively inaccessible to transcription factors and other proteins, so that gene expression in these regions is repressed.
- The interactions of several proteins with one another and with the hypoacetylated N-terminal tails of histones H3 and H4 are responsible for the chromatin-mediated repression of transcription that occurs in the telomeres and the silent mating-type loci in *S. cerevisiae* (see Figure 8-32).
- Some repression domains function by interacting with co-repressors that are histone deacetylase complexes. The subsequent deacetylation of histone N-terminal tails in nucleosomes near the repressor-binding site inhibits interaction between the promoter DNA and general transcription factors, thereby repressing transcription initiation (see Figure 8-33a).
- Activation domains function by binding multiprotein co-activator complexes such as histone acetylase complexes. The subsequent hyperacetylation of histone N-terminal tails in nucleosomes near the activator-binding site facilitates interactions between the promoter DNA and general transcription factors, thereby stimulating transcription initiation (see Figure 8-33b).
- SWI/SNF chromatin-remodeling factors constitute another type of co-activator. These multisubunit complexes can transiently dissociate DNA from histone cores in an ATP-dependent reaction and may also decondense regions of chromatin, thereby promoting the binding of DNA-binding proteins needed for transcription initiation.
- The Mediator complex, another type of co-activator, is a roughly 30-subunit complex that forms a molecular bridge between activation domains and RNA polymerase II by binding directly to the polymerase and activation domains. By binding to several

different activators either simultaneously or in rapid succession, Mediator probably helps integrate the effects of multiple activators on a single promoter (see Figure 8-37).

- Activators bound to a distant enhancer can interact with transcription factors bound to a promoter because chromatin is flexible and the intervening chromatin can form a large loop.

8.5 Regulation of Transcription-Factor Activity

We have seen in the preceding discussion how combinations of transcription factors binding to specific DNA regulatory sequences control transcription of eukaryotic genes. Consequently, the nuclear concentrations and activities of the transcription factors that interact with the transcription-control regions of a gene are crucial to determining whether a specific gene in a multicellular organism is expressed in a particular cell at a particular time during development or in differentiated cells. As we have seen, local chromatin structure also contributes to transcription regulation in eukaryotes, as in the case of Sir protein repression at telomeres and the silent mating-type loci in yeast (see Figure 8-32). In the next section, we discuss other examples of chromatin-mediated control that can result in *transcriptional memory*. Here we discuss regulation of transcription-factor expression and activity.

Which transcription factors are expressed in a particular cell type is determined by previous interactions between other transcription factors and control regions in transcription-factor genes that occur during development and differentiation of that cell type. We now have a high-resolution view of how transcription-factor binding changes during development and differentiation of multiple human cell types. This understanding is due to the identification of DNase I hypersensitive sites on a genomic scale.

DNase I Hypersensitive Sites Reflect the Developmental History of Cellular Differentiation

In Chapter 7, we learned that an expressed gene is far more sensitive to digestion by the nuclease DNase I than when the gene is not expressed (see Figure 7-27). In addition to this general increase in DNase I sensitivity over the length of a gene, researchers have found that specific short regions of the genome are extremely sensitive to DNase I digestion compared to the vast majority of the DNA in chromatin. These stretches of DNA are on the order of a hundred base pairs in length, and are the first regions cut when isolated nuclei are treated with low levels of DNase I. These cut regions are known as *DNase I hypersensitive sites (DHSs)*. High-throughput sequencing methods have allowed mapping of DHSs across the genome in multiple differentiated and embryonic cell types. Briefly, after digestion of isolated nuclei with low levels of DNase I, DNA is isolated from the treated chromatin. Oligonucleotide linkers of a known sequence are ligated to the DNA ends generated by DNase I digestion. The DNA is sheared into small fragments by sonication, amplified by PCR, and sequenced. Using this approach, human DNA sequences adjacent to the oligonucleotide linker sequence were thus identified as DHSs.

Figure 8-41a shows plots of the number of times a DHS was sequenced in samples from the human cell types indicated at the

left. The frequency at which genomic DNA within an interval of 1 kb along the length of a chromosome is observed by this method is a measure of the DNase sensitivity of this 1 kb region of the genome. The height of each vertical bar represents the degree of sensitivity of the DNA sequence at that position to digestion in nuclei isolated from each of the cell types. A roughly 600-kb region of the genome on chromosome 9 is shown.

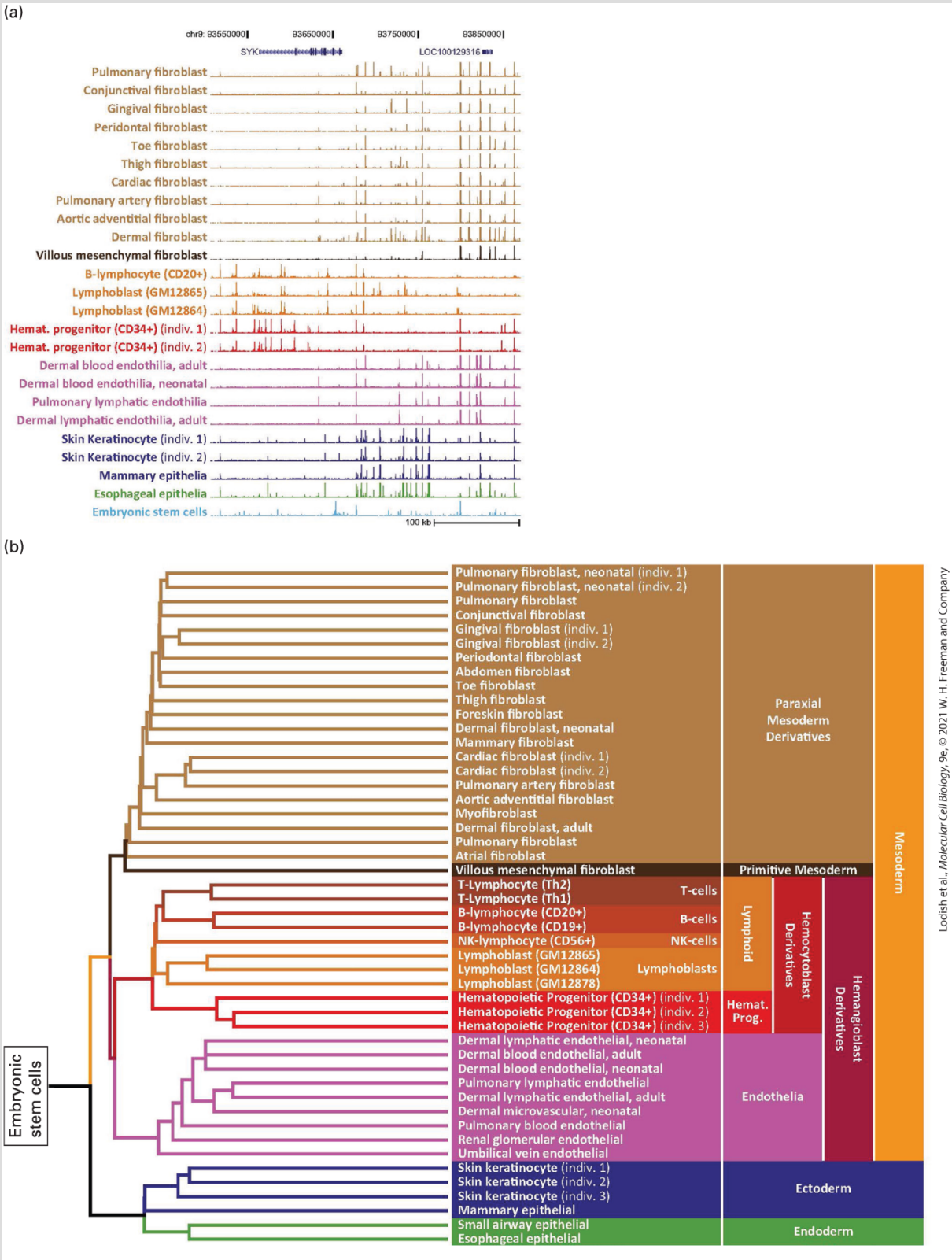


FIGURE 8-41 Maps of DNase I hypersensitive sites in embryonic and adult cells reflect their developmental history. (a) DHSs from each of the human cell types shown at the left

are mapped in an interval on chromosome 9. The height of each vertical bar in the figure represents the number of times a sequence in a 50-bp interval at that position was sequenced and represents DNase sensitivity. The plots are color-coded according to the embryonic tissue from which they developed. (b) Dendrogram showing the relationships among the DHS maps for each cell type across the entire genome. The embryonic tissue from which each of these cell types develops is shown at the right. Embryonic stem cells form the root of the dendrogram. The DHS maps for all other cell types are derived from those for the embryonic stem cell by loss of some DHSs and the acquisition of other DHSs. The dendrogram, based on how closely DHS maps from two cell types are related, parallels the developmental relationships among the cell types.

[Republished with permission from Elsevier, from A. B. Stergachis et al., 2013, "Developmental Fate and Cellular Maturity Encoded in Human Regulatory DNA Landscapes," *Cell*, **154**:888–903; permission conveyed through Copyright Clearance Center, Inc.]

DHSs appear to be sites where sequence-specific binding proteins, such as transcription factors, bind to DNA. Consequently, the DHS pattern in a region of chromatin indicates the positions of bound transcription factors, although the transcription factors bound are not directly identified.

In Figure 8-41a, the type of tissue from which the DHS data were determined is shown on the left, and the embryonic tissues from which these tissue types developed are color-coded as indicated in Figure 8-41b. It is apparent that more closely related cell types, such as fibroblasts from different regions of the body, or endothelial cells that line the inner surfaces of blood vessels from different organs, have more similar DHSs than more distantly related cell types. With computer methods, it is possible to compare the similarity of the DHS maps for each of these cell types across the entire genome.

With these computational methods, a dendrogram can be generated showing how closely the DHS map from one cell type resembles those of other cell types (see Figure 8-41b). This dendrogram is similar to the dendrograms used to show the relatedness, and hence the evolution, of gene sequences (see Figure 6-24b).

Importantly, the DHS pattern of embryonic stem cells is at the root of the DHS dendrogram for all cell types (see Figure 8-41b). These cells from the inner cell mass of the early mammalian embryo (discussed in Chapter 22; see Figure 22-4) are the progenitors of all cells in the adult organism. Embryonic stem cells have the largest number of DHSs, suggesting that they have the most complex transcriptional control of all cells: about 257,000 DHSs in one study, compared with 90,000–150,000 in differentiated cells. This difference probably reflects the developmental potential of embryonic stem cells. Approximately 50,000–100,000 DHSs that are not found in embryonic stem cells arise during development. Note that for each adult cell type, a different set of new DHSs arises. These DHS patterns reveal the complexity of the combinations of transcription factors that regulate each gene.

Approximately a million distinct DHSs were characterized in the cell types shown in Figure 8-41. Since DHSs include clustered transcription-factor binding sites typical of enhancers, these results suggest that on average, combinations of four or five enhancers regulate the transcription of each of the roughly 19,000 genes in the human genome. The maps of DHSs reveal where binding of early

embryonic transcription factors is lost and where new cell-type-specific combinations of transcription factors bind as a cell differentiates. Even this estimate fails to capture the complexity of transcriptional control, since many transcription-factor-binding sites detected as one DHS are, in fact, bound by related transcription factors with similar DNA-binding domains in different cell types. The related transcription factors regulate the appropriate level of transcription for that cell type.

Nuclear Receptors Are Regulated by Lipid-Soluble Hormones

In addition to controlling the expression of transcription factors, cells also regulate the activity of many of their transcription factors once they have been expressed. For example, the activity of many transcription factors is regulated by signals received from other cells. Interactions between extracellular domains of transmembrane-receptor proteins on the surface of the cell and specific protein ligands for these receptors secreted by other cells, or expressed on the surface of neighboring cells, activate intracellular domains of the transmembrane proteins. The result is that the signal received on the outside of the cell is transduced to a signal on the inside of the cell. The intracellular signal regulates the activities of enzymes that modify transcription factors by post-translational protein modifications, such as phosphorylation,

acetylation, and proteosomal degradation (see Figure 3-32). These post-translational modifications activate or inhibit transcription factors in the nucleus. In Chapters 15 and 16, we describe the major types of cell-surface receptors, the protein ligands they bind, and the intracellular signaling pathways that regulate transcription-factor activity.

Here we discuss a group of extracellular signals that use a (somewhat) different pathway to regulate the activities of transcription factors: small, lipid-soluble hormones, including steroid hormones, retinoids, and thyroid hormones. These lipid-soluble hormones diffuse through the plasma and nuclear membranes and interact directly with the transcription factors they control (Figure 8-42). One group of transcription factors that we mentioned already, the *nuclear-receptor superfamily*, needs to bind to lipid-soluble hormones to convert the transcription factors into transcriptional activators. We will now consider these interactions in more detail.

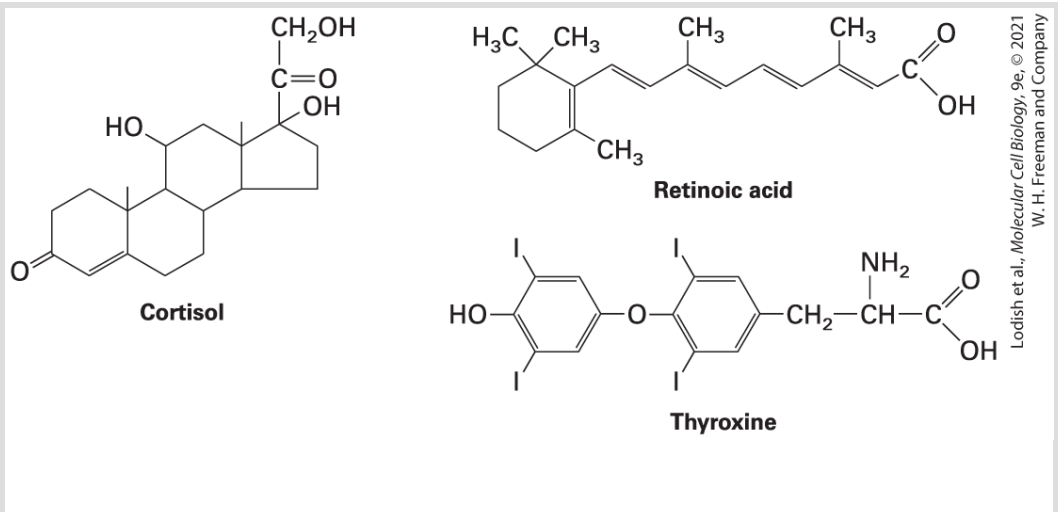


FIGURE 8-42 Examples of hormones that bind to nuclear receptors. These and related lipid-soluble hormones diffuse through the plasma and nuclear membranes and bind to receptors located in the cytosol or nucleus. The ligand-receptor complex functions as a transcription activator.

All Nuclear Receptors Share a Common Domain Structure

cDNA sequencing of mRNAs encoding various nuclear receptors has revealed a remarkable conservation in their amino acid sequences. It has also revealed that each of these receptors has three functional regions (Figure 8-43). The first is a unique N-terminal region of variable length (100–500 amino acids). Portions of this variable region function as activation domains in most nuclear receptors. The second is a DNA-binding domain that maps near the center of the primary sequence and contains a repeat of the C_4 zinc-finger motif (see Figure 8-25b). The third region, the hormone-binding domain, located near the C-terminal end, contains a hormone-dependent activation domain (see Figure 8-26b, c). In some nuclear receptors, the hormone-binding domain functions as a repression domain in the absence of ligand.

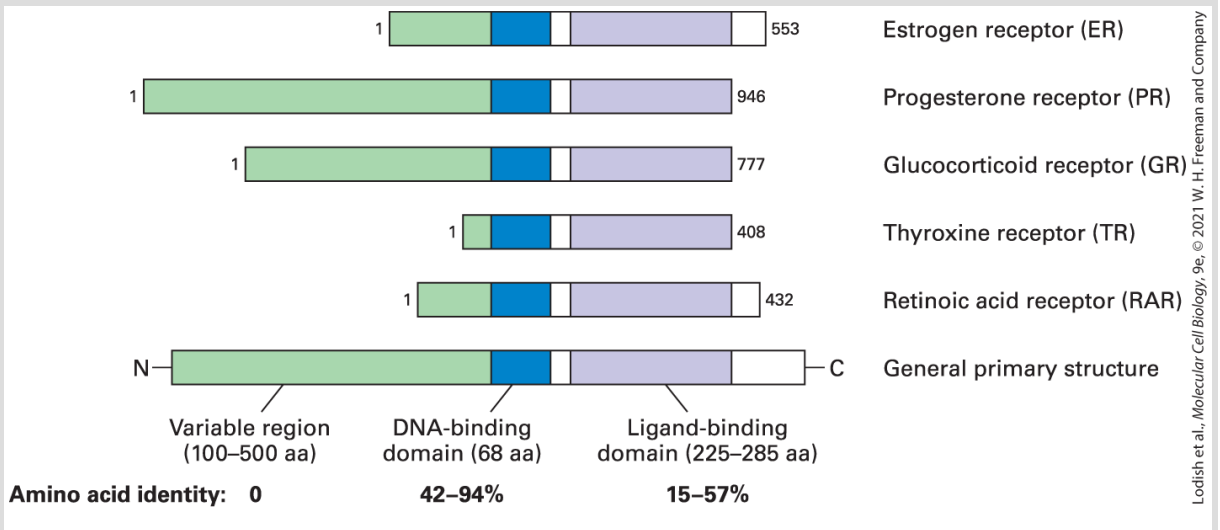
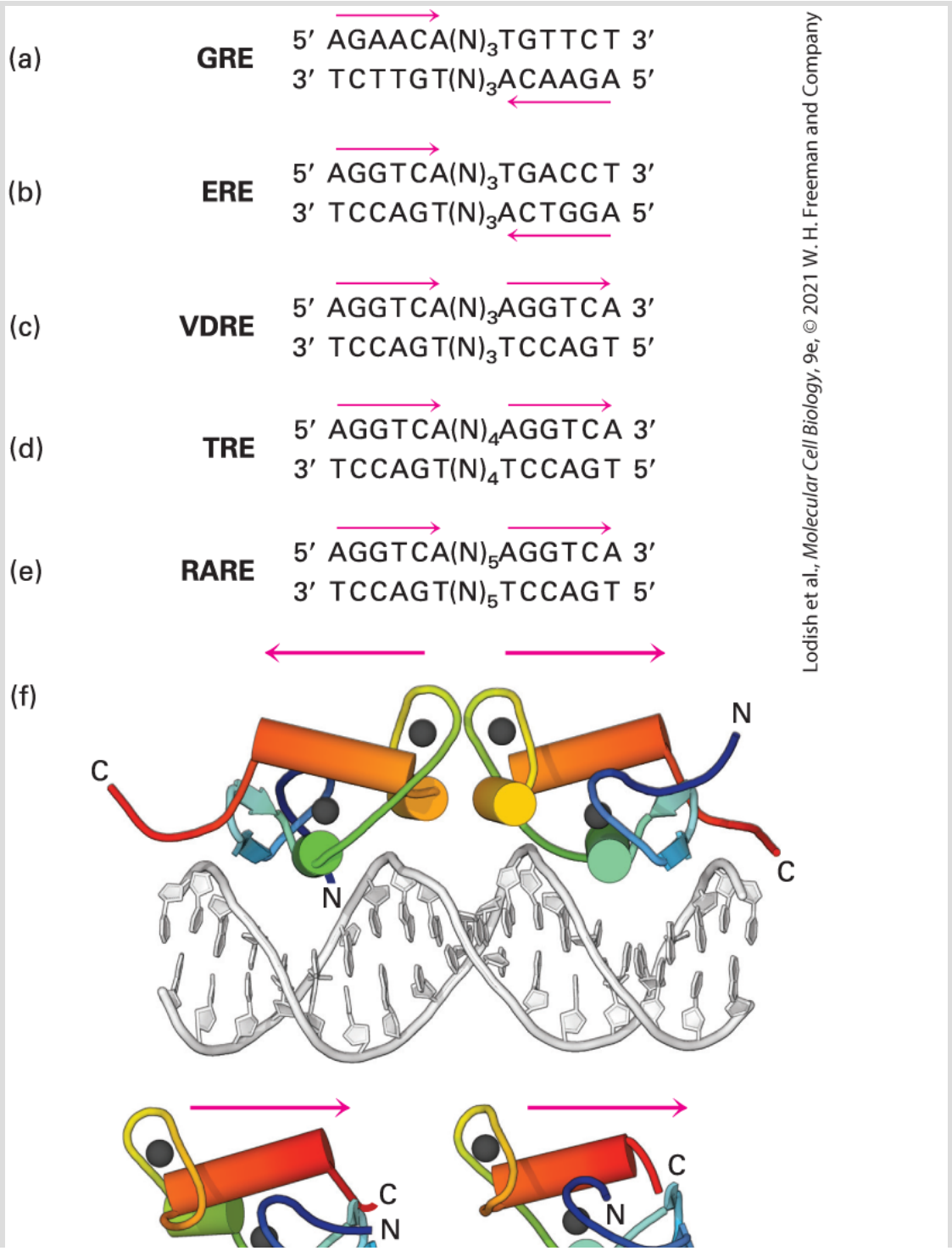


FIGURE 8-43 General design of transcription factors in the nuclear-receptor superfamily. The centrally located DNA-binding domain exhibits considerable sequence homology among different receptors and contains two copies of the C_4 zinc-finger motif (see Figure 8-25b). The C-terminal hormone-binding domain exhibits somewhat less homology. The N-terminal regions of various receptors vary in length, have unique sequences, and may contain one or more activation domains. See R. M. Evans, 1988, *Science* **240**:889.

Nuclear-Receptor Response Elements Contain Inverted or Direct Repeats

The DNA sites to which nuclear receptors bind are called *response elements*. The nucleotide sequences of several response elements have been determined. The consensus sequences of response elements for two steroid hormone receptors, the glucocorticoid receptor response element (GRE) and the estrogen receptor response element (ERE) are 6-bp inverted repeats separated by any three base pairs (Figure 8-44a, b). This finding suggested that the cognate steroid hormone receptors would bind to DNA as symmetric

dimers. This was confirmed by x-ray crystallographic analysis of the homodimeric glucocorticoid receptor's C_4 zinc-finger DNA-binding domain (see Figure 8-25b).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

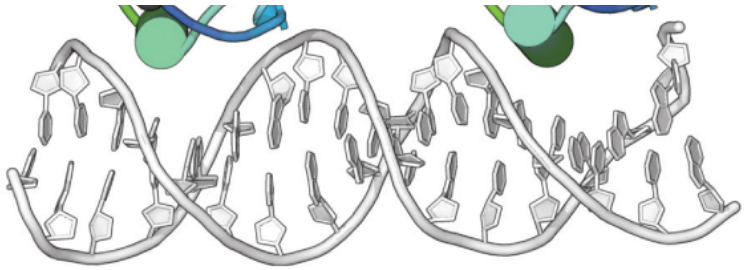


FIGURE 8-44 Consensus sequences of DNA response elements that bind five nuclear receptors. (a, b) The glucocorticoid and estrogen receptors are twofold symmetric dimers that bind, respectively, to the glucocorticoid receptor response element (GRE) and the estrogen receptor response element (ERE). Each of these response elements contains inverted repeats separated by three base pairs. (c–e) The heterodimeric nuclear receptors each contain one RXR subunit associated with another nuclear-receptor subunit that defines the hormone response. RXR-VDR mediates responses to vitamin D_3 by binding to a direct repeat separated by three base pairs (a VDRE). RXR-TR mediates responses to thyroid hormone by binding to the same DNA bases in a direct repeat separated by four base pairs (a TRE). Similarly, RXR-RAR mediates a response to retinoic acid by binding to the same direct repeat separated by five base pairs, comprising a RARE. The repeat sequences bound by the reading helices of these receptors are indicated by red arrows. (f) Crystal structures of the glucocorticoid receptor bound to DNA containing a GRE (*top*) and of the RXR-TR heterodimer bound to DNA containing a TRE (*bottom*). Red arrows indicate the orientation from N to C of the reading helices. Note that in the twofold symmetric glucocorticoid receptor, the reading helices are inverted relative to each other so that they "read" an AGAACA on the top strand of the left half-site and on the bottom strand of the right half-site, separated by 3 base pairs. Consequently, the binding site for the glucocorticoid receptor and other twofold symmetric homodimers such as the estrogen receptor is an inverted repeat [see (a) and (b)]. In contrast, the reading helices in the RXR-TR heterodimer are in the same orientation. Consequently, they read an AGGTCA sequence in the same orientation in the two-half sites separated by four base pairs, a direct-repeat binding site. The interface between the RXR subunit and the vitamin D_3 receptor (VDR) subunit bound to a VDRE brings the two reading helices closer together so that they bind to the same half-sites separated by three rather than four base pairs. Similarly, the interface between the RXR and RAR subunits bound to a RARE positions the two reading helices in the heterodimer farther apart than in the RXR-TR, so that they bind the same AGGTCA sequences

separated by five base pairs. See K. Umesono et al., 1991, *Cell* **65**:1255, and A. M. Naar et al., 1991, *Cell* **65**:1267.

[Part (f) Data from F. Rastinejad et al., 1995, *Nature* **375**:203.]

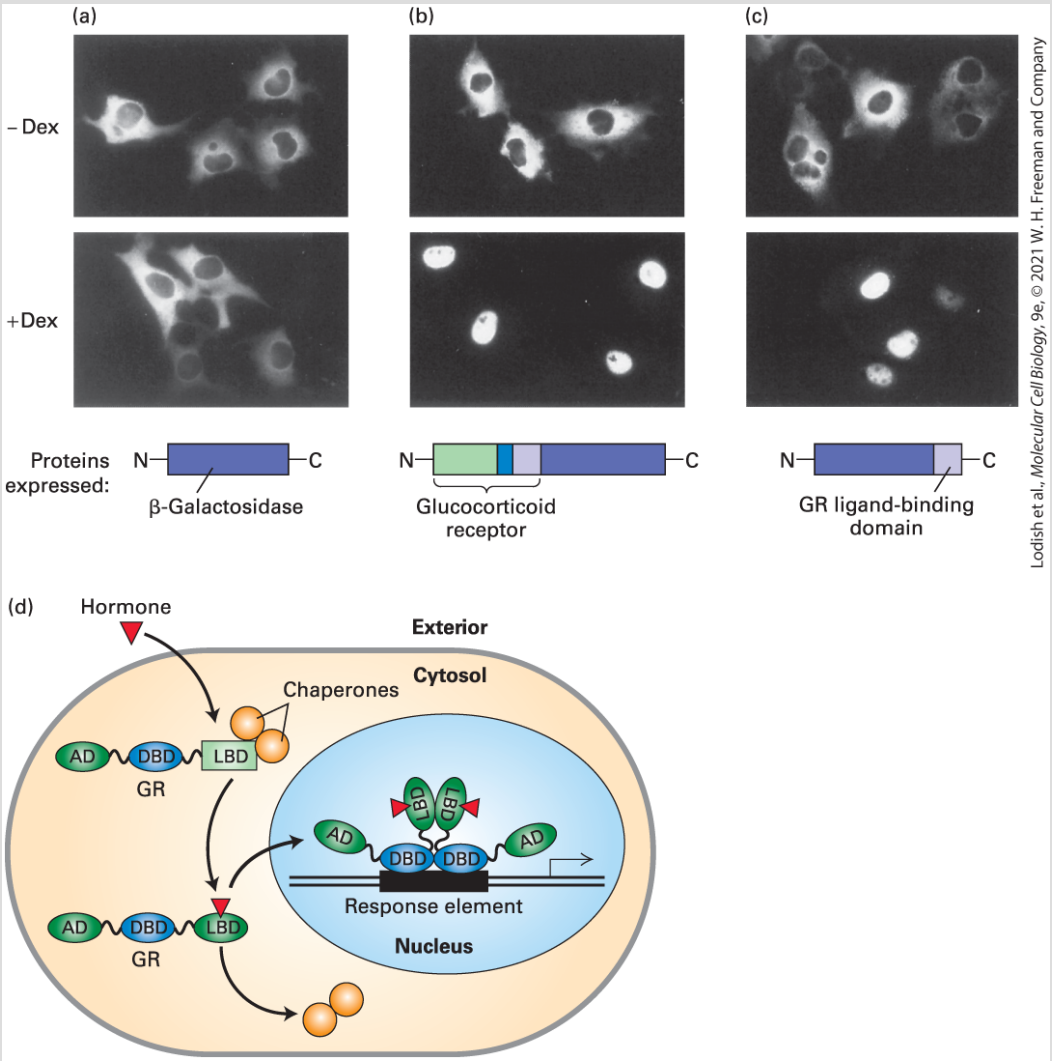
Some nuclear-receptor response elements, such as those for the receptors that bind nonsteroids such as vitamin D_3, thyroid hormone, and retinoic acid, are direct repeats of the same sequence that is recognized by the estrogen receptor, separated by three, four, or five base pairs (Figure 8-44c–e). The specificity of these response elements is determined by the spacing between the repeats. The nuclear receptors that bind to these direct-repeat response elements do so as heterodimers, all of which share a monomer called *RXR*. The vitamin D_3 response element (VDRE), for example, is bound by the RXR-VDR heterodimer, and the retinoic acid response element (RARE) is bound by RXR-RAR. The monomers composing these heterodimers interact with each other in such a way that the two DNA-binding domains lie in the same orientation, allowing the RXR heterodimers to bind to direct repeats of the binding site for each monomer (Figure 8-44f). In contrast, the monomers in homodimeric nuclear receptors (e.g., GRE and ERE) have an inverted orientation.

Hormone Binding to a Nuclear Receptor Regulates Its Activity as a Transcription Factor

The mechanism by which hormone binding controls the activity of nuclear receptors differs between heterodimeric and homodimeric receptors. Heterodimeric nuclear receptors (e.g., RXR-VDR, RXR-TR, and RXR-RAR) are located in the nucleus. In the absence of their hormone ligand, they repress transcription when bound to their cognate sites in DNA. They do so by directing histone deacetylation at nearby nucleosomes by associating with histone deacetylase complexes, as described earlier for other repressors (see Figure 8-33a). When a heterodimeric nuclear receptor binds its ligand, the receptor undergoes a conformational change, and as a consequence, it dissociates from histone deacetylase complexes and binds histone acetylase complexes, thereby reversing its own repressing effects. In the presence of ligand, the ligand-bound conformation of the receptor also binds Mediator, stimulating preinitiation complex assembly.

In contrast to heterodimeric nuclear receptors, homodimeric receptors are found in the cytoplasm in the absence of their hormone ligands. When a hormone binds to its receptor, the hormone-receptor complex translocates into the nucleus. The hormone-dependent translocation of the homodimeric glucocorticoid receptor (GR) was demonstrated in the experiments shown in Figure 8-45a–c. The GR hormone-binding domain alone mediates this transport. Subsequent studies showed that in the absence of hormone, GR cannot be transported into the nucleus because its ligand-binding domain is partially unfolded and associates reversibly with Hsp70, a major cellular protein chaperone

that assists folding of nascent polypeptides in the cytoplasm (see Figure 3-19a). As long as the receptor is confined to the cytoplasm, it cannot interact with target genes and hence cannot activate transcription. Hormone binding promotes an exchange of GR from Hsp70 to Hsp90 (see Figure 3-19b), which, with coupled hydrolysis of ATP, refolds the GR ligand-binding domain, increasing the affinity for hormone and releasing GR from Hsp70 so that it can enter the nucleus. Once in the nucleus in the conformation induced by ligand binding, the GR can bind to response elements associated with target genes (Figure 8-45d). Once the receptor with bound hormone binds to a response element, it activates transcription by interacting with chromatin-remodeling and histone acetylase complexes and Mediator.



EXPERIMENTAL FIGURE 8-45 Fusion proteins from expression vectors demonstrate that the hormone-binding domain of the glucocorticoid receptor mediates translocation to the nucleus in the presence of hormone. Cultured animal cells were transfected with expression vectors encoding the proteins diagrammed in part (d). Immunofluorescence with a labeled antibody specific for β-galactosidase was used to detect the expressed proteins in transfected cells. (a) In cells that expressed β-galactosidase alone, the enzyme was localized to the cytoplasm in the presence and absence of the glucocorticoid hormone dexamethasone (Dex). (b) In cells that expressed a fusion protein consisting of β-galactosidase and the entire glucocorticoid receptor (GR), the fusion protein was present in the cytoplasm in the absence of hormone but was transported to the nucleus in the presence of hormone. (c) Cells that expressed a fusion protein composed of β-galactosidase and only the

GR ligand-binding domain (light purple) also exhibited hormone-dependent transport of the fusion protein to the nucleus. (d) Model of hormone-dependent gene activation by a homodimeric nuclear receptor. In the absence of hormone, the receptor is kept in the cytoplasm by interaction between its ligand-binding domain (LBD) and chaperone proteins. When hormone is present, it diffuses through the plasma membrane and binds to the ligand-binding domain, causing a conformational change that releases the receptor from the chaperone proteins. The receptor with bound ligand is then translocated into the nucleus, where its DNA-binding domain (DBD) binds to response elements, allowing the ligand-binding domain and an additional activation domain (AD) at the N-terminus to stimulate transcription of target genes.

[Parts (a)–(c) Republished with permission of John Wiley & Sons — Books, from D. Picard and K. R. Yamamoto, 1987, "Two Signals Mediate Hormone-Dependent Nuclear Localization of the Glucocorticoid Receptor," *EMBO J.*, **6**(11): 3333–3340, Fig. 3; permission conveyed through Copyright Clearance Center, Inc.]

Metazoans Regulate the RNA Polymerase II Transition from Initiation to Elongation

A recent unexpected discovery that resulted from application of the chromatin immunoprecipitation technique is that a large fraction of genes in metazoans have a paused elongating RNA polymerase II within about 50–100 bp of the transcription start site. Thus expression of the encoded protein is controlled not only by transcription initiation but also by transcription elongation early in the transcription unit. The first genes discovered to be regulated by

control of transcription elongation were **heat-shock genes** (e.g., *hsp70*), which encode molecular chaperones that refold denatured proteins and other proteins that help the cell deal with the effects of heat shock. When heat shock occurs, the heat-shock transcription factor (HSF1) is activated (see Chapter 21). Binding of activated HSF1 to specific sites in the promoter-proximal region of heat-shock genes stimulates the paused polymerase to continue chain elongation and promotes rapid reinitiation by additional Pol II molecules, leading to many transcription initiations per minute. This mechanism of transcriptional control permits a rapid response: these genes are always paused in a state of suspended transcription and therefore, when an emergency arises, no time is required to remodel and acetylate chromatin at the promoter and assemble a transcription preinitiation complex.

Another transcription factor known to regulate transcription elongation is MYC, which functions in the regulation of cell growth and division. MYC is often expressed at high levels in cancer cells and is a key transcription factor in the reprogramming of somatic cells into pluripotent stem cells capable of differentiation into any cell type. The ability to induce differentiated cells to convert to pluripotent stem cells has elicited enormous research interest because of its potential for the development of therapeutic treatments for traumatic injuries to the nervous system and degenerative diseases (see Chapter 22).

Termination of Transcription Is Also Regulated

Once Pol II has transcribed about 50 nucleotides from the transcription start site, elongation through most genes is highly processive. Chromatin immunoprecipitation with antibody to Pol II, however, indicates that the amount of Pol II at various positions in a transcription unit in a population of cells varies greatly (see Figure 8-12b, *right*). This finding indicates that the enzyme can elongate through some regions much more rapidly than others. In most cases, Pol II does not terminate transcription until after a sequence is transcribed that directs cleavage and polyadenylation of the RNA at the 3′ end of the encoded mRNA. Pol II can then terminate transcription at multiple sites located 0.5–2 kb beyond this poly(A) addition site. Experiments with mutant genes show that termination is coupled to the process that cleaves and polyadenylates the 3′ end of a transcript, which is discussed in the next chapter.

KEY CONCEPTS OF SECTION 8.5

Regulation of Transcription-Factor Activity

- The activities of many transcription factors are indirectly regulated by binding of extracellular proteins and peptides to cell-surface receptors. These receptors activate intracellular signal transduction pathways that regulate specific transcription factors through a variety of mechanisms discussed in Chapter 16.
- Nuclear receptors constitute a superfamily of dimeric C_4 zinc-finger transcription factors that bind lipid-soluble hormones and interact with specific response elements in DNA (see Figures 8-41 and 8-43).

- Hormone binding to nuclear receptors induces conformational changes that modify the interactions of these receptors with other proteins (see Figure 8-26b, c).
- Heterodimeric nuclear receptors (e.g., those for retinoids, vitamin D, and thyroid hormone) are found only in the nucleus. In the absence of hormone, they repress transcription of target genes with the corresponding response element. When bound to their ligands, they activate transcription.
- Steroid hormone receptors are homodimeric nuclear receptors. In the absence of hormone, they are trapped in the cytoplasm by molecular chaperones. When bound to their ligands, they can translocate to the nucleus and activate transcription of target genes (see Figure 8-44).
- In metazoans, RNA polymerase II often pauses during elongation within approximately 50–100 base pairs from the transcription start site. Release from this pause contributes to the regulation of gene transcription.
- Termination by RNA polymerase II is coupled to cleavage of transcripts at the end of the final exon followed by polyadenylation. Cleavage and polyadenylation are considered further in Chapter 9.

8.6 Epigenetic Regulation of Transcription

The term **epigenetics** refers to the study of inherited changes in the phenotype of a cell that do not result from changes in DNA sequence. For example, during the differentiation of bone marrow stem cells into the several different types of blood cells, a hematopoietic stem cell divides into two daughter cells, one of which continues to have the properties of a hematopoietic stem cell, including the potential to differentiate into all the different types of blood cells. But the other daughter cell becomes a type of progenitor cell that normally can differentiate into lymphocytes or myeloid cells (see **Figure 22-18**). Lymphoid progenitor cells generate daughter cells that differentiate into lymphocytes, which perform many of the functions involved in immune responses to pathogens (see Chapter 24). Myeloid progenitor cells divide into daughter cells that are committed to differentiating into red blood cells, different kinds of phagocytic white blood cells, or the cells that generate platelets involved in blood clotting. Lymphoid and myeloid progenitor cells both have the same DNA sequence as the zygote (generated by fertilization of an egg cell by a sperm cell) from which they developed, but they have restricted developmental potential because of epigenetic differences between them.

Such epigenetic changes are initially the consequence of the expression of specific master-transcription factors that are

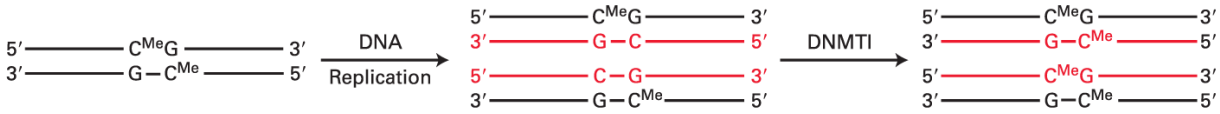
regulators of cellular differentiation, controlling the expression of other genes that encode transcription factors and proteins involved in cell-cell communication in complex networks of gene control. Understanding these networks is currently the subject of intense investigation. Changes in gene expression initiated by transcription factors are often reinforced and maintained over multiple cell divisions by post-translational modifications of histones and methylation of DNA at position 5 of the cytosine pyrimidine ring (see Figure 2-17). These changes to chromatin are maintained and propagated to daughter cells when cells divide. Consequently, the term *epigenetic marks* is used to refer to such post-translational modifications of histones and 5-methyl C modification of DNA.

DNA Methylation Regulates Transcription

Recall that most promoters in mammals fall into the CpG island class. Active CpG island promoters have unmethylated cytosines in their CG sequences. Unmethylated CpG island promoters have reduced affinity for histone octamers. But nucleosomes immediately neighboring the unmethylated promoters are methylated at histone H3 lysine 4. These methylated nucleosomes are associated with Pol II molecules that have paused during transcription of both the sense and antisense template-DNA strands (see Figures 8-12 and 8-13). Recent research indicates that methylation of histone H3 lysine 4

occurs in mouse cells because a protein named Cfp1 (CXXC finger protein 1) binds unmethylated CpG-rich DNA through a zinc-finger domain (CXXC) and associates with a histone methylase (Setd1) specific for histone H3 lysine 4. Chromatin-remodeling complexes and the general transcription factor TFIID, which initiates Pol II preinitiation complex assembly (see Figure 8-13), associate with nucleosomes bearing the H3 lysine 4 trimethyl mark, thus promoting Pol II transcription initiation.

In differentiated cells, a small percentage of specific CpG island promoters, depending on the cell type, have CpGs marked by 5-methyl C. A family of proteins binds to DNA that is rich in 5-methyl C–modified CpGs. Once bound, these methyl CpG-binding proteins, or MBDs, associate with histone deacetylases and repressive chromatin-remodeling complexes that condense chromatin, resulting in transcriptional repression. The 5-methyl C is added to the CpGs by DNA methyl transferases named DNMT3a and DNMT3b. They are referred to as *de novo DNA methyl transferases* because they methylate Cs in a region where there is no 5-methyl C in either DNA strand. Much remains to be learned about how DNMT3a and b are directed to specific CpG islands. But once they have methylated a DNA sequence, methylation at that C is passed on through DNA replication through the action of the ubiquitous *maintenance* methyl transferase DNMT1 (red indicates daughter strands):



As a consequence, once a CpG island promoter is methylated by DNMT3a or b, it continues to be methylated by DNMT1 in subsequent daughter cells. Consequently, the promoter remains repressed in all subsequent daughter cells through interactions with MBDs, even after the stimulus for the initial C-methylation by DNMT3a or b has ceased. Therefore, repression of C-methylated promoters is inherited through cell division. This mechanism of epigenetic repression is being intensely investigated because tumor-suppressor genes encoding proteins that function to suppress the development of cancer are often inactivated in cancer cells by abnormal CpG methylation of their promoter regions. This is discussed further in Chapter 25.

Methylation of Specific Histone Lysines Is Linked to Epigenetic Mechanisms of Gene Repression

Figure 7-26b summarized the major different types of post-translational modifications that are found on histones, including acetylation and methylation of lysines. The acetylation state at a specific histone lysine on a particular nucleosome results from a dynamic equilibrium between acetylation and deacetylation by

histone acetylases and histone deacetylases. Acetylation of histones in a localized region of chromatin predominates when local DNA-bound activators transiently bind histone acetylase complexes. Deacetylation predominates when repressors transiently bind histone deacetylase complexes. Pulse-chase radiolabeling experiments have shown that acetyl groups on histone lysines turn over rapidly through the sequential actions of these competing enzymes.

In contrast, methyl groups on histones are much more stable. Methylation of lysine occurs on the nitrogen atom of the terminal ε-amino group of the lysine side chain (see Figure 2-14). Lysines can be modified by the addition of one, two, or three methyl groups to this terminal nitrogen atom, generating mono-, di-, and trimethylated lysine, all of which carry a single positive charge. Histone lysine methyl groups can be removed by *histone lysine demethylases.* But the resulting turnover of histone lysine methyl groups is much slower than the turnover of histone lysine acetyl groups, which makes methylation a more appropriate post-translational modification for propagating epigenetic information.

Three other common post-translational modifications of histones have been characterized (see Figure 7-26b). Through study of these histone modifications and their effects on transcription and other processes like the formation of mitotic chromosomes, a picture of chromatin has emerged. We know that the histone tails extending as

random coils from the chromatin fiber are modified to generate one of many possible combinations of modifications that regulate chromatin-based processes by affecting the binding of a large number of different protein complexes. This control of protein interactions with specific regions of chromatin that results from the combined influence of post-translational modifications of histones has been called a *histone code.* Some of these modifications, such as histone lysine acetylation, are rapidly reversible, whereas others, such as histone lysine methylation, can be epigenetically inherited. Table 8-2 summarizes the post-translational modifications that make up the histone code.

TABLE 8-2 • Histone Post-Translational Modifications Associated with Active and Repressed Genes

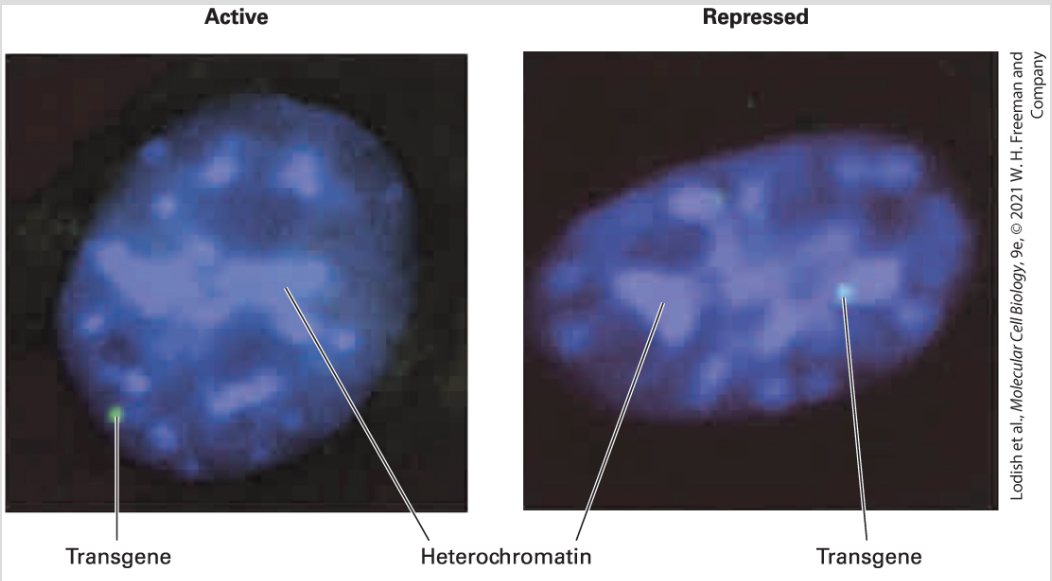
Modification	Sites of Modification	Effect on Transcription
Acetylated lysine	H3 (K9, K14, K18, K23, K27)	Activation
	H4 (K5, K8, K12, K16)	Activation
	H2A (K5, K8, K13)	Activation
	H2B (K5, K12, K15, K20)	Activation
Hypoacetylated lysine		Repression
Phosphorylated serine/threonine	H3 (T3, S10, T11, S28)	Activation
	H2A (S1, T120)	Activation

	H2B (S14)	Activation
Methylated arginine	H3 (R2, R17, R26)	Activation
	H4 (R3)	Activation
Methylated lysine	H3 (K4) Me3 in promoter region	Activation
	H3 (K4) Me1 in enhancers	Activation
	H3 (K36, K79) in transcribed region	Elongation
	H3 (K9, K27) Me2/3	Repression
	H4 (K20) Me2/3	Repression
Ubiquitinylated lysine	H2B (K120 in mammals, K123 in *S. cerevisiae*)	Activation
	H2A (K119 in mammals)	Repression

Histone H3 Lysine 9 Methylation in Heterochromatin

In most eukaryotes, some co-repressor complexes contain histone methyl transferase subunits that methylate histone H3 at lysine 9, generating di- and trimethyl lysines. These methylated lysines are binding sites for isoforms of HP1 protein, which functions in the condensation of heterochromatin, as discussed in Chapter 7 (see Figure 7-28). For example, the KAP1 co-repressor complex functions with a class of more than 200 zinc-finger–transcriptional repressors

encoded in the human genome. The co-repressor complex includes a subunit with H3 lysine 9 methyl transferase activity that methylates nucleosomes over the promoter regions of genes targeted for repression by these zinc-finger repressors. This results in HP1 binding to these promoter regions, chromatin condensation, and transcriptional repression. An integrated transgene in cultured mouse fibroblasts that was repressed through the action of the KAP1 co-repressor was associated with heterochromatin in most cells, whereas the active form of the same transgene was associated with euchromatin (Figure 8-46). Chromatin immunoprecipitation assays showed that the repressed gene was associated with histone H3 methylated at lysine 9 and with HP1, whereas the active gene was not.



EXPERIMENTAL FIGURE 8-46 Association of a repressed transgene with heterochromatin. Mouse fibroblasts were stably transformed with a transgene that contained binding sites for an engineered repressor. The repressor was a fusion between a DNA-binding domain, a repression domain that interacts with

the KAP1 co-repressor complex, and the ligand-binding domain of a nuclear receptor that allows the nuclear import of the fusion protein to be controlled experimentally (see Figure 8-45). DNA was stained blue with the dye DAPI. Brighter-staining regions are regions of heterochromatin, where the DNA concentration is higher than in euchromatin. The transgene was detected by hybridization of a fluorescently labeled complementary probe (green). When the recombinant repressor was retained in the cytoplasm, the transgene was transcribed (*left*) and was associated with euchromatin in most cells. When hormone was added so that the recombinant repressor entered the nucleus, the transgene was repressed (*right*) and associated with heterochromatin. Chromatin immunoprecipitation assays (see Figure 8-12) showed that the repressed gene was associated with histone H3 methylated at lysine 9 and HP1, whereas the active gene was not.

[Republished with permission from Cold Spring Harbor Laboratory Press, from K. Ayyanathan et al., 2003, "Regulated Recruitment of HP1 to a Euchromatic Gene Induces Mitotically Heritable, Epigenetic Gene Silencing: A Mammalian Cell Culture Model of Gene Variegation," *Genes Dev.*, **17**:1855–1869; Fig. 6.]

Importantly, H3 lysine 9 methylation is maintained following chromosome replication by the mechanism diagrammed in Figure 8-47. When a methylated region of DNA is replicated in S phase, the nucleosomes associated with the parent DNA are randomly distributed to the daughter DNA molecules. New histone octamers that are not methylated on H3 lysine 9 also associate randomly with the new daughter chromosomes to reestablish the density of nucleosomes before DNA replication, but since the parent nucleosomes are associated with both daughter chromosomes, approximately half of the daughter chromosomes' nucleosomes are methylated on lysine 9. The repressor complex also contains a subunit with a *chromo* domain that binds histone H3 trimethylated at

lysine 9 (H3K9me3). This causes the co-repressor complex to associate with the newly replicated chromatin in which about half the nucleosomes contain H3K9me3. As a result, the methylase subunit of the co-repressor complex is brought into sufficiently close proximity to the H3 histone N-terminal tails of the newly assembled histone octamers that they methylate all the H3 N-termini associated with that region of chromatin, leading to HP1 association and chromatin condensation (see Figure 7-28). Repetition of this process with each cell division results in maintenance of H3 lysine 9 methylation of this region of the chromosome.

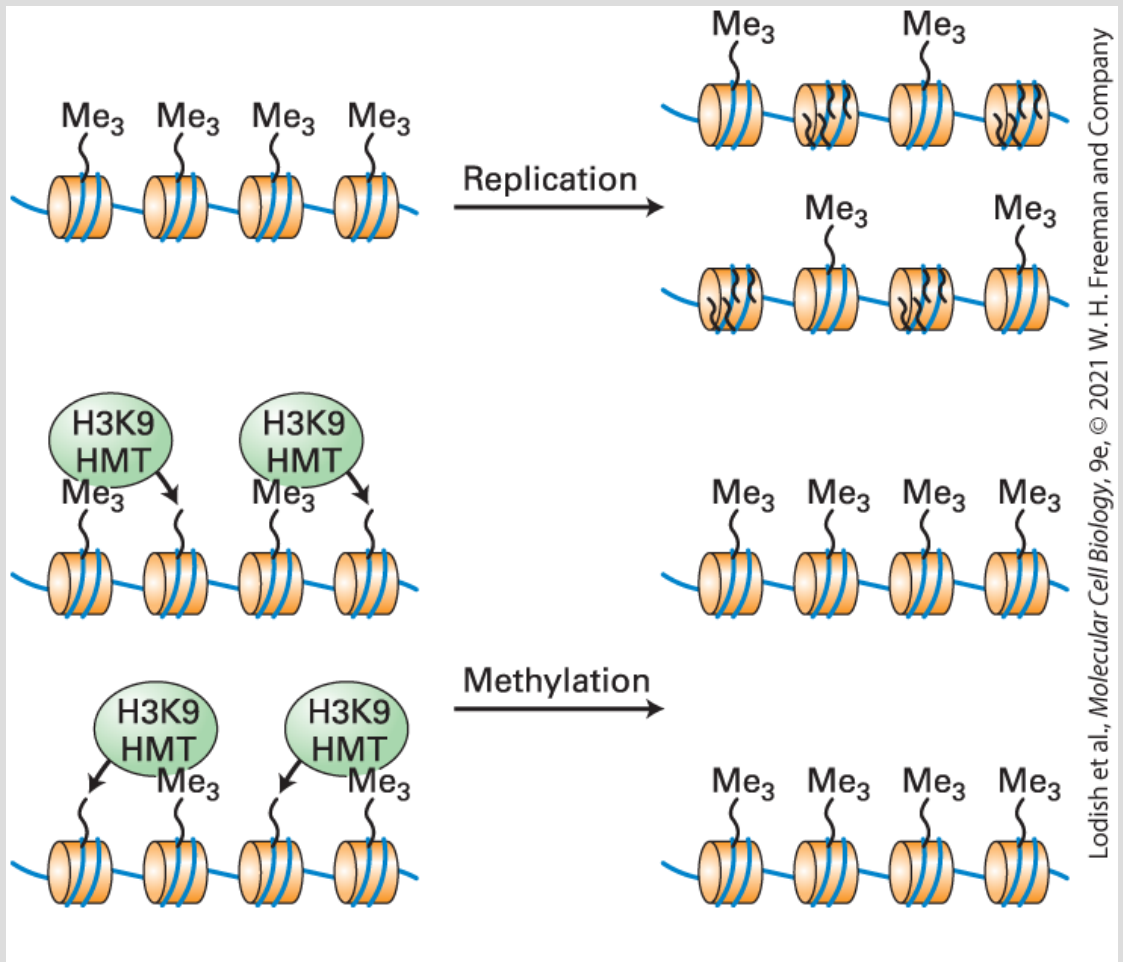


FIGURE 8-47 Maintenance of histone H3 lysine 9 methylation during chromosome replication. When chromosomal DNA is replicated, the parent histones randomly associate with the two daughter chromosomes, while unmethylated histones synthesized during S phase are assembled into other nucleosomes in those same daughter chromosomes. Association of histone H3 lysine 9 methyl transferases (H3K9 HMT) with parent nucleosomes bearing the histone 3 lysine 9 di- or trimethylation mark methylates the newly added unmodified nucleosomes. Consequently, histone H3 lysine 9 methylation marks are maintained during repeated cell divisions unless they are specifically removed by a histone demethylase.

Epigenetic Control by Polycomb and Trithorax Complexes

A critical type of epigenetic mark found in multicellular animals and plants involves a set of proteins known collectively as *Polycomb* and a counteracting set of proteins known as *Trithorax*. These names were derived from the phenotypes of mutations in the genes encoding these proteins in *Drosophila*, in which they were first discovered. Polycomb is essential for maintaining the repression of genes in specific types of cells, and in all of the cells that descend from them, throughout the life of an organism. Important genes that are repressed by Polycomb proteins include the Hox genes, which encode master regulatory transcription factors. Different combinations of Hox transcription factors help to direct the development of specific tissues and organs in a developing embryo. Early in embryogenesis, expression of Hox genes is controlled by

typical activator and repressor proteins. However, the expression of these activators and repressors stops at an early point in embryogenesis. Correct expression of Hox genes in the descendants of early embryonic cells is maintained for the remainder of embryogenesis and into adult life by Polycomb and Trithorax proteins. Polycomb proteins maintain the *repression* of specific Hox genes in all cells that descend from early embryonic cells where the Hox gene was repressed in the early embryo. Trithorax proteins perform the opposite function; they maintain the *expression* of Hox genes in the descendants of early embryonic cells that expressed the Hox gene. Polycomb and Trithorax proteins control thousands of genes, including genes that regulate cell growth and division. Polycomb and Trithorax genes are often mutated in cancer cells, contributing to the abnormal properties of these cells (see Chapter 25).

Remarkably, virtually all cells in the developing embryo and adult express a similar set of Polycomb and Trithorax proteins, and all cells contain the same set of Hox genes. Yet only the Hox genes in descendants of cells where they were initially repressed in early embryogenesis remain repressed, even though the same Hox genes in other cells in a neighboring region of the embryo remain active in the presence of the same Polycomb proteins. Consequently, as in the case of the yeast silent mating-type loci (see Section 8.4), the expression of Hox genes is regulated by a process that involves more

than specific DNA sequences interacting with proteins that diffuse through the nucleoplasm.

A current model for repression by Polycomb proteins is depicted in Figure 8-48. Most Polycomb proteins are subunits of one of two classes of multiprotein *P*olycomb *r*epressive *c*omplexes: PRC1 and PRC2. The PRC2 complexes are thought to act initially by associating with the repression domains of specific repressors bound to their cognate DNA sequences early in embryogenesis, or with ribonucleoprotein complexes containing long noncoding RNAs, which will be discussed in the next section. The PRC2 complexes contain histone deacetylases that inhibit transcription. They also contain a subunit [E(z) in *Drosophila*, EZH2 in mammals] with a *SET domain*, which is the catalytic domain of several histone methyl transferases. The SET domain in PRC2 methylates histone H3 on lysine 27, generating di- and trimethyl lysines (H3K27me2 and H3K27me3). A PRC1 complex then binds the nucleosomes with H3K27me2 and 3 through dimeric Pc subunits (CBXs in mammals), each containing a methyl lysine–binding chromodomain specific for methylated H3 lysine 27. It has been proposed that binding of the dimeric Pc to neighboring nucleosomes condenses the chromatin into a structure that inhibits transcription. This proposal is supported by electron microscopy studies showing that PRC1 complexes cause nucleosomes to associate in vitro (see Figure 8-48d, e).

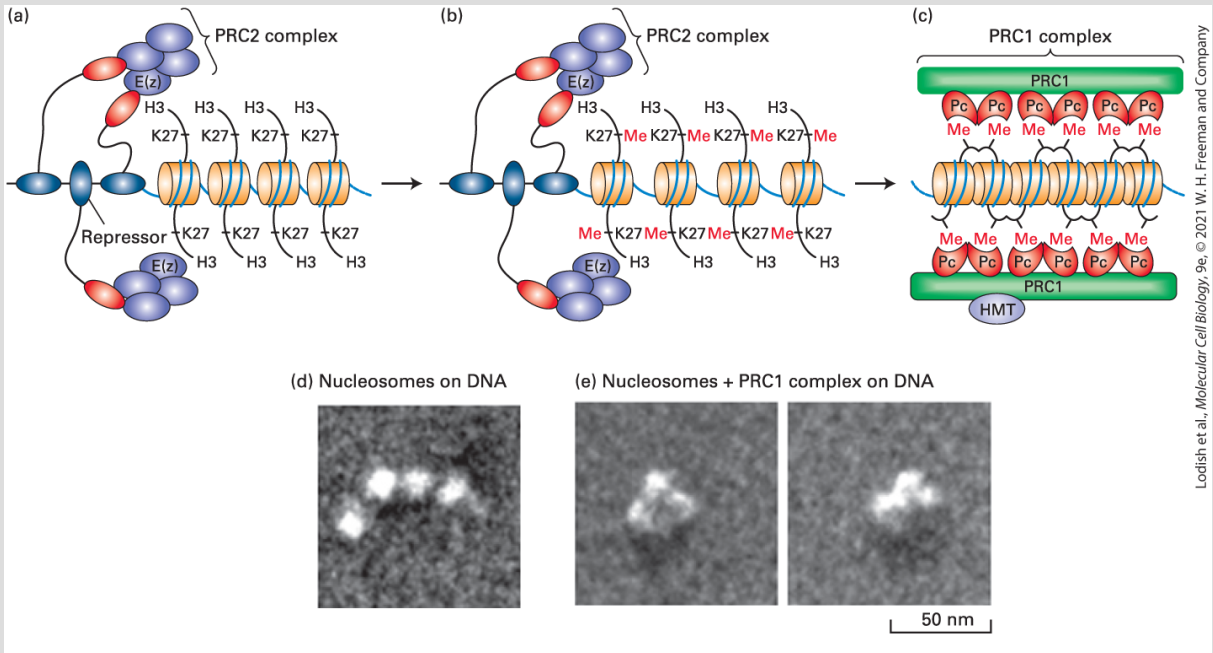


FIGURE 8-48 Model for repression by Polycomb complexes. (a) During early embryogenesis, repressors associate with the PRC2 complex. (b) This association results in methylation (Me) of neighboring nucleosomes on histone H3 lysine 27 (K27) by the SET domain–containing subunit E(z). (c) The PRC1 complex binds nucleosomes methylated at H3 lysine 27 through a dimeric, chromodomain-containing subunit Pc. The PRC1 complex condenses the chromatin into a repressed chromatin structure. PRC2 complexes associate with PRC1 complexes to maintain H3 lysine 27 methylation of neighboring histones. As a consequence, PRC1 and PRC2 association with the region is maintained when expression of the repressor proteins in (a) ceases. (d, e) Electron micrograph of a 1-kb fragment of DNA bound by four nucleosomes in the absence (d) and presence (e) of one PRC1 complex per five nucleosomes. See A. H. Lund and M. van Lohuizen, 2004, *Curr. Opin. Cell Biol.* **16**:239; and N. J. Francis, R. E. Kingston, and C. L. Woodcock, 2004, *Science* **306**:1574.

[Part (d)–(e) Republished with permission from AAAS, from N. J. Francis et al., 2004, "Chromatin Compaction by a Polycomb Group Protein Complex, " *Science* **306**(5701):1574–1577; permission conveyed through Copyright Clearance Center, Inc.]

PRC1 complexes also repress transcription through additional mechanisms. The PRC1 complex contains a ubiquitin ligase that *mono*ubiquitinylates histone H2A at lysine 119 in the H2A C-terminal

tail (see Figure 7-26b). This modification of H2A blocks elongation by inhibiting a histone chaperone that removes histone octamers from DNA as Pol II transcribes through a nucleosome and then replaces them as the polymerase passes. PRC1 also associates with a histone demethylase that specifically removes methyl groups from lysine 4 of histone H3, an activating mark discussed below with Trithorax proteins.

PRC2 complexes associate with nucleosomes bearing the histone H3K27me3 mark, maintaining methylation of H3 lysine 27 in nucleosomes in the region. This methylation results in association of chromatin with PRC1 and PRC2 complexes even after expression of the initial repressor proteins shown in Figure 8-48a, b has ceased. This association maintains H3 lysine 27 methylation by a mechanism analogous to that diagrammed in Figure 8-47. This mechanism is a key feature of Polycomb repression, which is maintained through successive cell divisions for the life of an organism (∼100 years for some vertebrates and 5000 years for a bristle cone pine!).

Trithorax proteins counteract the repression mechanism of Polycomb proteins, as shown in studies of expression of the Hox transcription factor Abd-B in the *Drosophila* embryo (Figure 8-49). Abd-B is normally expressed only in posterior segments of the developing embryo. When the Polycomb system is defective, Abd-B is expressed in all cells of the embryo. When the Trithorax system is

defective and cannot counteract repression by the Polycomb system, Abd-B is repressed in most cells, except those in the very posterior of the embryo. Trithorax complexes include a histone methyl transferase that trimethylates histone H3 lysine 4, a histone methylation that is associated with the promoters of actively transcribed genes. This histone modification creates a binding site for histone acetylase and for chromatin-remodeling complexes that promote transcription, as well as for TFIID, the general transcription factor that initiates preinitiation-complex assembly (see Figure 8-13). Nucleosomes with H3 lysine 4 methylation are also binding sites for specific histone demethylases that remove H3 histone lysine 9 and lysine 27 methylation, preventing the binding of HP1 and the Polycomb repressive complexes. Nucleosomes marked with H3 lysine 4 methylation are also thought to be distributed to both daughter DNA molecules during DNA replication, resulting in maintenance of this epigenetic mark by a strategy similar to that diagrammed in Figure 8-47.

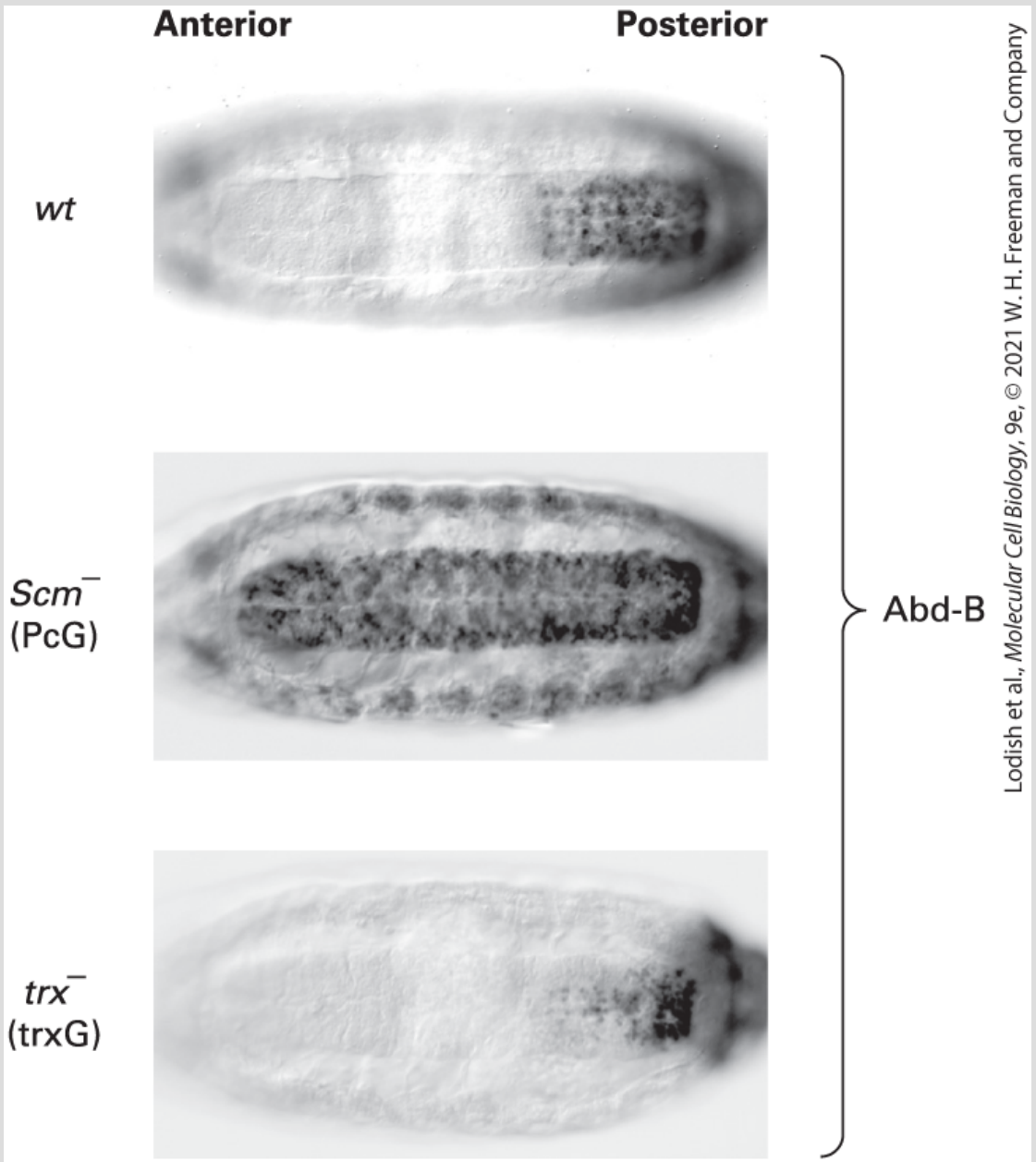


FIGURE 8-49 Opposing influence of Polycomb and Trithorax complexes on expression of the Hox transcription factor Abd-B in *Drosophila* embryos. At the stage of *Drosophila* embryogenesis shown, Abd-B is normally expressed only in posterior segments of the developing embryo, as shown at the top (*wt*) by immunostaining with a specific anti–Abd-B antibody. In embryos with homozygous mutations of *Scm*, a Polycomb gene (PcG) encoding a protein associated with the PRC1 complex, Abd-B expression is derepressed in all embryo segments. In contrast, in homozygous mutants of *trx*, a Trithorax gene (trxG), Abd-B repression is increased

so that the protein is expressed at high concentrations only in the most posterior segment.

[Republished with permission from John Wiley & Sons — Books, from T. Klymenko and J. Muller, 2004, "The Histone Methyltransferases Trithorax and Ash1 Prevent Transcriptional Silencing by Polycomb Group Proteins," *EMBO Rep.* **5**(4):373–377; permission conveyed through Copyright Clearance Center, Inc.]

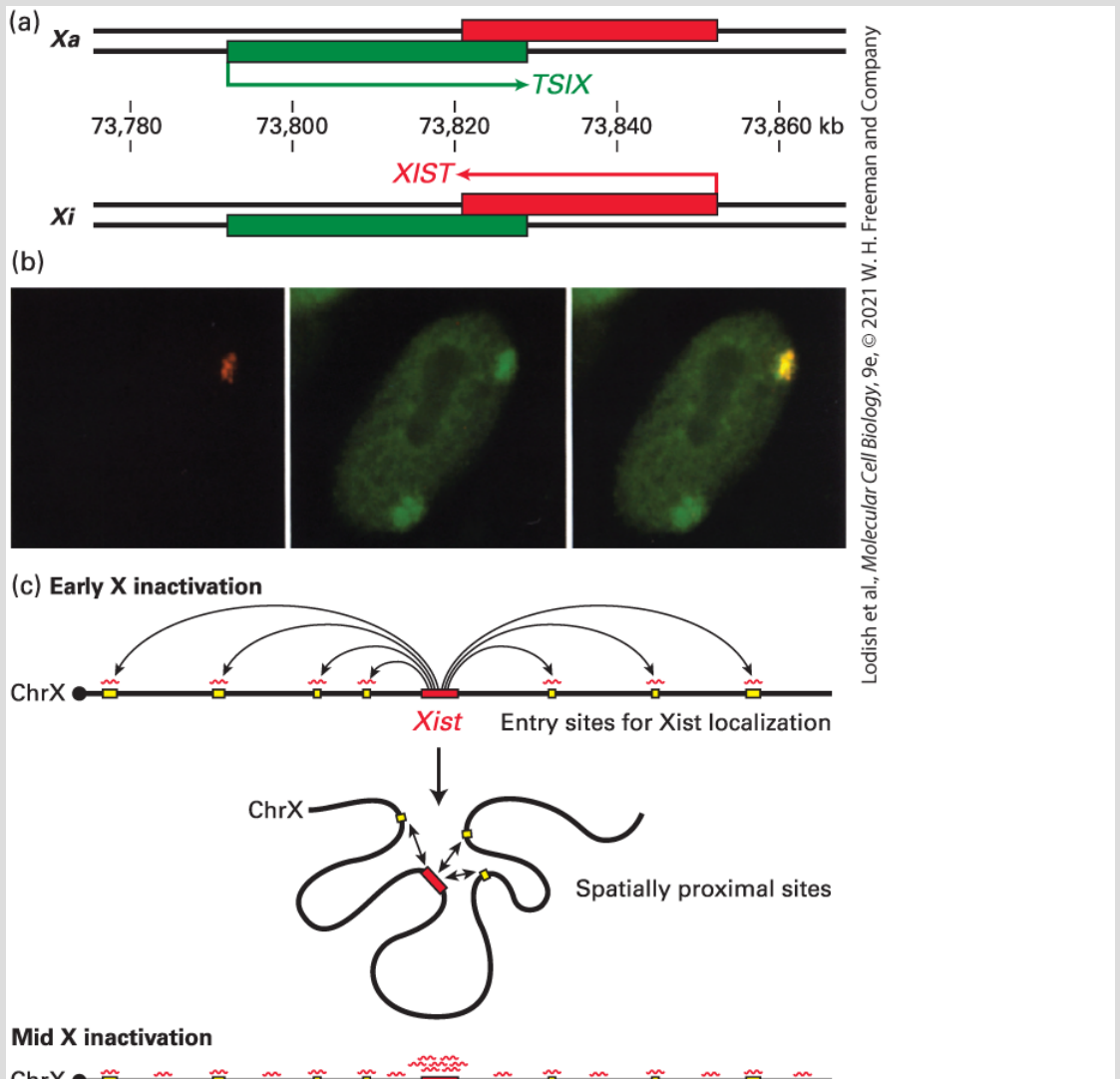
Long Noncoding RNAs Direct Epigenetic Repression in Metazoans

Complexes have been discovered that are composed of multiple proteins that function to repress transcription bound to long RNAs. These RNAs, which can be many kilobases in length, lack long open reading frames and are consequently called **long noncoding RNAs** or **lncRNAs**. In some cases, these lncRNA-protein complexes repress genes on the same chromosome from which the RNA is transcribed, as in the case of X-chromosome inactivation in female mammals. In other cases, these RNA-protein complexes act in trans, repressing genes on chromosomes other than those from which the lncRNA is transcribed.

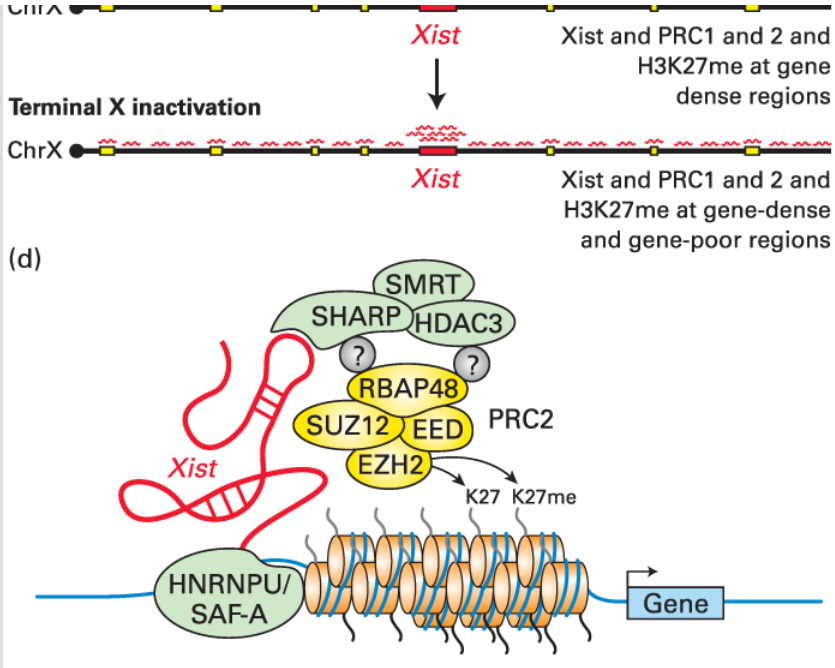
X-Chromosome Inactivation in Mammals

X-chromosome inactivation in female mammals (see Chapter 7) is one of the most intensely studied examples of epigenetic repression mediated by a lncRNA. X inactivation is controlled by an ∼100-kb

domain on the X chromosome called the *X-inactivation center*. Remarkably, this region encodes several lncRNAs required for the random inactivation of one entire X chromosome early in the development of female mammals. The functions of these lncRNAs are only partially understood. Some are transcribed from complementary DNA strands near the middle of the X-inactivation center: the 40-kb Tsix lncRNA and the *Xist* RNA. The latter is spliced and polyadenylated into an RNA of about 17 kb that remains in the nucleus (Figure 8-50a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company



EXPERIMENTAL FIGURE 8-50 The *Xist* long noncoding RNA encoded in the X-inactivation center coats the inactive X chromosome in cells of mammalian females, repressing transcription of most genes on the inactive X. (a) The region of the human X-inactivation center encoding the noncoding RNAs *Xist* (transcribed from the inactive X), and Tsix (transcribed from the active X). Numbers are bp from the left end of the X-chromosome. (b) A cultured fibroblast from a human female was analyzed by in situ hybridization with a probe complementary to *Xist* RNA labeled with a red fluorescent dye (*left*), a chromosome paint set of probes for the X chromosome labeled with a green fluorescent dye (*center*), and an overlay of the two fluorescent micrographs. The condensed inactive X chromosome is associated with *Xist* RNA. (c) Model for the spreading of the *Xist* lncRNA-protein complex on the inactive X chromosome during early differentiation of female embryonic stem cells. See E. Heard and A.-V. Gendrel, 2014, *Annu. Rev. Cell Dev. Biol.* **30**:561. (d) Proteins associated with *Xist* lncRNA. Question marks indicate that it is not yet known how PRC2 complexes (yellow) associate with HDAC3 and the RNA-binding protein SHARP although this might be due to formation of transcriptional condensates. See C. A. McHugh et al., 2015, *Nature* **521**:232.

[Part (b) C. M. Clemson et al., 1996, *J. Cell Biol.* **132**(3):259–275. https://doi.org/10.1083/jcb.132.3.259.]

In differentiated female cells, the inactive X chromosome is associated with 50–100 multimolecular *Xist* RNA-protein complexes dispersed along most of the length of the X-chromosome (Figure 8-50b, c). Proteins associated with these complexes include co-repressors and their associated histone deacetylase complexes, repressive chromatin remodeling complexes and polycomb complexes discussed earlier (Figure 8-48). These proteins are thought to associate through multiple weak interactions between their intrinsically disordered regions forming a repressive transcriptional condensate by the same principles discussed earlier for the formation of activating transcriptional condensates (see Figure 8-48). These repressive transcriptional condensates scattered along the X-chromosome are thought to condense the inactive X-chromosome into a chromatin structure that inhibits transcription. Targeted deletion of the *Xist* gene in cultured embryonic stem cells showed that it is required for X inactivation. Unlike most protein-coding genes on the inactive X chromosome, the *Xist* gene is actively transcribed. The *Xist* RNA-protein complexes do not diffuse to interact with the active X or other chromosomes, but remain associated with the inactive X chromosome. In contrast to *Xist*, Tsix is transcribed from the active X chromosome, not from the inactive X chromosome.

In the early female embryo, made up of embryonic stem cells capable of differentiating into all cell types (see Chapter 22), genes on both X chromosomes are transcribed, and the 40-kb Tsix lncRNA

(Figure 8-50a) is transcribed from both copies of the X chromosome. Experiments employing engineered deletions in the X-inactivation center showed that Tsix transcription prevents significant transcription of the *Xist* RNA from the complementary DNA strand. Later in development, as cells begin to differentiate, Tsix transcription is repressed on one of the X chromosomes. This repression occurs randomly in different cells on either of the X chromosomes. This inhibition of Tsix transcription determines which of the X chromosomes will be inactivated as the cells differentiate further, because inhibition of Tsix transcription allows transcription of *Xist* lncRNA to occur on that chromosome.

Recent studies indicate that *Xist* lncRNA-protein complexes first associate with regions of the X chromosome localized near the X-inactivation center in the three-dimensional, folded structure of the future inactive X (Figure 8-50c), as shown by chromosome conformation capture assays (see Figure 7-30). These initial sites of *Xist* association are in gene-rich regions of the X chromosome and are postulated to serve as entry sites where additional copies of the *Xist* lncRNA-protein complexes first bind and then spread to neighboring regions. The mechanism of spreading is not currently understood. The inactive X chromosome also becomes associated with PRC2 complexes, which catalyze the trimethylation of histone H3 lysine 27. This methylation results in association of the PRC1 complex and transcriptional repression, as discussed in the previous section. These mechanisms of transcriptional repression must be

redundant, however, because repression still occurs in the absence of the Polycomb proteins essential for the assembly of PRC1 and PRC2. At the same time, transcription of Tsix continues from the active X chromosome, which is associated with repression of *Xist* transcription from that X chromosome, and consequently prevents *Xist*-mediated repression of the active X. *Xist* and PRC1 and 2 complexes are then observed to associate with gene-poor regions of the inactive X chromosome as well as with gene-rich regions, silencing ∼1000 genes on the inactive X chromosome.

Xist RNA-protein complexes were isolated from cultured mouse embryonic stem cells undergoing the initiation phase of X inactivation. This was done by hybridizing several biotinylated oligonucleotides complementary to regions of *Xist* to nuclear extracts prepared from these cells, and then binding to a column with covalently bound streptavidin, a bacterial protein with very high affinity for biotin. Proteins eluted from this affinity column were analyzed by mass spectrometry (see Chapter 3) to identify proteins associated with *Xist* lncRNA as diagrammed in Figure 8-50d. Thus the *Xist* RNA is postulated to function as a large molecular scaffold for the assembly of a multiprotein complex. One of the proteins identified in the complex was *SMRT*, a co-repressor of the thyroid hormone nuclear receptor. During X inactivation, SMRT also interacts with the histone deacetylase HDAC3. Subsequent knockdown experiments with siRNAs directed against SMRT and HDAC3 showed that they are required for X inactivation, as are other

RNA and chromatin-binding proteins, such as SHARP, that link SMRT to *Xist* RNA (Figure 8-50d). These proteins are also required for the association of *Xist* RNA with the inactive X chromosome. These studies of *Xist* have led to the suggestion that other stable lncRNAs may function similarly as scaffolds for the assembly of other multiprotein-RNA complexes.

Shortly after X inactivation by histone deacetylation and polycomb complexes, the DNA of the inactive X chromosome becomes methylated at most of its CpG island promoters. In addition, specialized histone octamers in which histone H2A is replaced by a paralog of H2A called *macroH2A* also associate with the inactive X chromosome. DNA methylation and macroH2A contribute to the stable repression of the inactive X chromosome throughout embryogenesis and adult life.

Cis Activation by Long Noncoding RNAs

Examples of lncRNAs involved in gene activation have been characterized recently. For example, HOTTIP lncRNA, which is transcribed from the 5′ end of the *HOXA* locus, is proposed to coordinate the activation of *HOXA* genes by binding to a histone H3 lysine 4 methylase. In addition, nascent transcripts of lncRNA genes have been reported to activate transcription from promoters several kilobases away by interacting with the Mediator complex and

delivering it to the promoter by looping of the intervening chromatin.

In humans, but not in mice, the lncRNA called *XACT* associates with multiple sites along the entire length of the active X chromosome. XACT may contribute to maintenance of gene activity on that chromosome. XACT is also remarkable for being one of the longest characterized RNAs: 252 kb! It is mostly unspliced.

In *Drosophila*, equal expression of genes encoded on the X chromosome in males and females (dosage compensation) does not result from inactivating one X chromosome in females. Rather, a generalized twofold increase in transcriptional activation of genes on the single X chromosome in males is controlled by two lncRNAs, roX1 and roX2, transcribed from the X chromosome only in males. The roX1 and roX2 RNAs associate with several proteins encoded by *MSL* (male-specific-lethal) genes and spread over the X chromosome specifically, much as *Xist* lncRNA-protein complexes spread over the inactive X in mammals.

Recently, sequencing of total cellular RNA in multiple types of human cells identified roughly 15,000 human lncRNAs. Many of these lncRNAs have sequences that are evolutionarily conserved in most mammals, while about 5000 are found only in primates. This conservation of sequence strongly suggests that these lncRNAs, like *Xist*, have important functions. Many lncRNAs are expressed only in

specific cell types at specific times during development. For example, multiple lncRNAs are expressed primarily in differentiating red blood cells. Knockdown (see Figure 6-41 and Chapter 9) of several of these lncRNAs inhibits normal red blood cell development, but precisely how these lncRNAs perform their essential functions is not yet clear. The study of these conserved long noncoding RNAs and how they influence gene expression is another area of intense current investigation.

ENCODE (Encyclopedia of DNA Elements) encompasses a consortium of international research groups organized and funded by the US National Human Genome Research Institute with the goal of building a comprehensive, publicly available database of functional elements in the human genome

These elements include DNA control elements and the transcription factors that bind to them, histone post-translational modifications mapped by chromatin immunoprecipitation/next-gen sequencing (ChIP-seq) and related methods, DNase I hypersensitive sites, and regulatory lncRNAs and their sites of association in the genome, as well as newly discovered regulatory elements that control cells and circumstances in which a gene is active. Data sets from human cells and cells of model organisms that are too large to be published are also made publicly available at a site called *GEO (Gene Expression*

Omnibus) maintained by the US National Center for Bioinformatics (NCBI). Most journals that publish research based on genomic methods such as RNA-seq and ChIP-seq require that authors upload their original data to GEO. Worldwide public access to these data sets is greatly accelerating the pace of discovery in the area of gene regulation.

KEY CONCEPTS OF SECTION 8.6

Epigenetic Regulation of Transcription

- Epigenetic control of transcription refers to repression or activation that is maintained after cells replicate as the result of DNA methylation or post-translational modification of histones, especially histone methylation.
- Methylation of CpG sequences in CpG island promoters in mammals generates binding sites for a family of methyl-binding proteins (MBDs) that associate with histone deacetylases, inducing hypoacetylation of the promoter regions and transcriptional repression.
- Histone H3 lysine 9 di- and trimethylation creates binding sites for the heterochromatin-associated protein HP1, which results in condensation of chromatin and transcriptional repression. These post-translational modifications are perpetuated following chromosome replication because the methylated histones randomly associate with the daughter DNA molecules and associate with histone H3 lysine 9 methyl transferases that methylate histone 3 lysine 9 on newly synthesized histone octamers assembled on the daughter DNA.
- Polycomb complexes maintain repression of genes initially repressed by sequence-specific repressors expressed early during embryogenesis. One class of Polycomb complexes, PRC2 complexes, associates with these repressors in early embryonic cells, resulting in methylation of histone H3 lysine 27. This methylation creates binding sites for subunits in the PRC2 complex as well as for PRC1 complexes, which condense chromatin, inhibit the assembly of preinitiation complexes, and inhibit elongation. Since parent histone octamers with H3 methylated at lysine 27 are distributed to both daughter DNA molecules following DNA replication, PRC2 complexes that associate with these nucleosomes maintain histone H3 lysine 27 methylation through cell division.

- Trithorax complexes oppose repression by Polycomb complexes by methylating H3 at lysine 4 and maintaining this activating mark through chromosome replication.
- X-chromosome inactivation in female mammals requires a long noncoding RNA (lncRNA) called *Xist* that is transcribed from the X-inactivation center of one X chromosome and then spreads along the length of the same chromosome. *Xist* interacts with a co-repressor that binds a histone deacetylase and PRC2 complexes at an early stage of embryogenesis, initiating X inactivation. X inactivation is maintained throughout the remainder of embryogenesis and adult life by continued association with Polycomb complexes and DNA methylation of CpG island promoters on the inactive X.
- Some lncRNAs cause repression of genes in trans (i.e., on other chromosomes), as opposed to the cis inactivation imposed by *Xist*. Repression is initiated by their interaction with PRC2 complexes.
- Some lncRNAs are associated with gene activation. Much remains to be learned about how lncRNAs are targeted to specific chromosomal regions, but the discovery of about 15,000 nuclear lncRNAs expressed in various human cells during specific stages of their differentiation suggests that lncRNAs are central to widely used mechanisms of transcription regulation.

8.7 Other Eukaryotic Transcription Systems

We conclude this chapter with a brief discussion of transcription initiation by the other two eukaryotic nuclear RNA polymerases, Pol I and Pol III. Although these systems, and particularly their regulation, are less thoroughly understood than transcription by RNA polymerase II, they are equally fundamental to the life of eukaryotic cells. The polymerases that transcribe mitochondrial and chloroplast DNA will be discussed in Chapter 12.

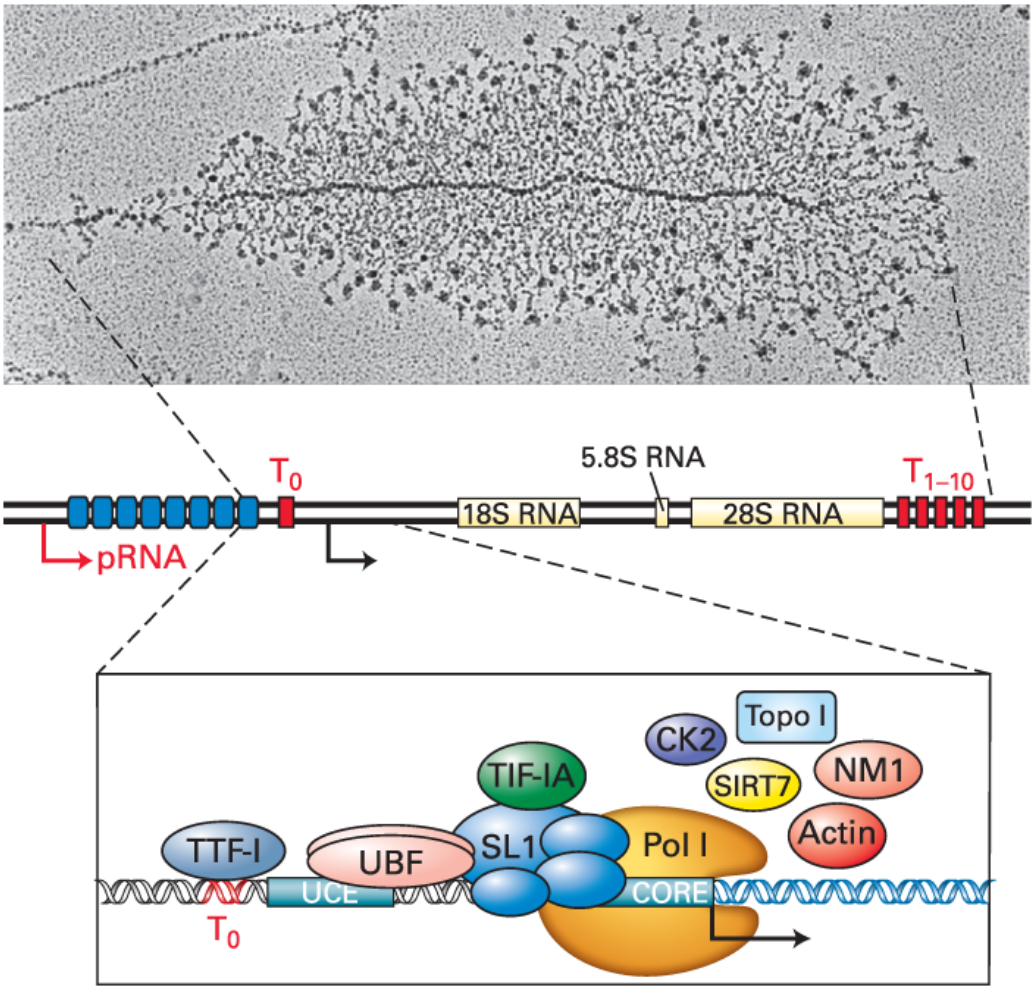
Transcription Initiation by Pol I and Pol III Is Analogous to That by Pol II

The formation of transcription initiation complexes involving Pol I and Pol III is similar in some respects to assembly of Pol II initiation complexes (see Figure 8-13). However, each of the three eukaryotic nuclear RNA polymerases requires its own polymerase-specific general transcription factors and each recognizes different DNA control elements. Moreover, neither Pol I nor Pol III requires ATP hydrolysis by a DNA helicase to help melt the DNA template strands to initiate transcription, whereas Pol II does. As we discuss in Chapter 21, transcription initiation by Pol I, which synthesizes pre-rRNA, and by Pol III, which synthesizes tRNAs, 5S rRNA, and other

small stable RNAs (see Table 8-1), is tightly coupled to the rate of cell growth and proliferation.

Initiation by Pol I

The regulatory elements directing Pol I initiation are similarly located relative to the transcription start site in yeast and in mammals. A *core element* spanning the transcription start site from −40 to +5 is essential for Pol I transcription. An additional *upstream control element* extending from roughly −155 to −60 increases in vitro Pol I transcription tenfold. In humans, assembly of the Pol I preinitiation complex (Figure 8-51) is initiated by the cooperative binding of UBF (upstream binding factor) and SL1 (selectivity factor), a multisubunit factor containing TBP and four Pol I–specific TBP-associated factors (TAF_Is), to the Pol I promoter region. The TAF_I subunits interact directly with Pol I–specific subunits, directing this specific nuclear RNA polymerase to the transcription start site. TIF-1A, the mammalian homolog of *S. cerevisiae* RRN3, is another required factor, as are the abundant nuclear protein kinase CK2 (casein kinase 2), nuclear actin, nuclear myosin, the protein deacetylase SIRT7, and topoisomerase I, which prevents DNA supercoils (see Figure 5-8) from forming during rapid Pol I transcription of the 14-kb transcription unit.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Electron micrograph courtesy of Ann L. Beyer.

FIGURE 8-51 Transcription of the rRNA precursor RNA by RNA polymerase I. *Top*: Electron micrograph of RNA-protein complexes transcribed from repeated rRNA genes. *Middle:* A single Pol I transcription unit. Enhancers that stimulate Pol I transcription from a single transcription start site are represented by blue boxes. Pol I transcription termination sites (T_0, $T_1 - T_{10}$) bound by the Pol I–specific termination factor TTF-1 are shown as red rectangles. pRNA indicates transcription of the noncoding pRNA required for transcriptional silencing. The black arrow indicates the start site of the primary transcript. Regions of DNA shown as parallel black lines are contained in the primary transcript, but are removed and degraded during rRNA processing. *Bottom:* The core promoter element and upstream control element are shown with the location of Pol I and its general transcription factors UBF, SL1, and TIF-1A represented, as well as other proteins required for Pol I elongation and control.

Transcription of the 14-kb precursor of 18S, 5.8S, and 28S rRNAs (see Chapter 9) is regulated to coordinate ribosome synthesis with cell growth and division. This coordination is achieved in several ways: regulation of the activities of Pol I initiation factors by post-translational modifications, including phosphorylation and acetylation; control of the rate of Pol I elongation; and control of the number of rRNA genes that are transcriptionally active by epigenetic mechanisms that assemble inactive copies into heterochromatin. At active rRNA genes, transcription terminates at terminator sequencs T_0 and T_1, T_2, T_3, … T_{10}, positioned as shown in Figure 8-51. This nonsymmetrical sequence functions to terminate Pol I transcription only when it is in the correct orientation. The terminator sequences are bound by a termination factor that causes termination and Pol I release from the template.

Switching between the active and heterochromatic silent states of rRNA genes is accomplished by a multisubunit chromatin-remodeling complex called *NoRC* (*No* for nucleolus, the site of rRNA transcription within nuclei). NoRC localizes a nucleosome over the Pol I transcription start site, blocking preinitiation complex assembly. It also interacts with a DNA methyl transferase that methylates a critical CpG in the upstream control element, thus inhibiting binding by UBF. NoRC also interacts with histone deacetylases and with histone methyl transferases that di- and trimethylate histone H3 lysine 9, creating binding sites for heterochromatic HP1. Moreover, a roughly 250-nt noncoding RNA

called *pRNA* (promoter-associated RNA) transcribed by Pol I from about 2 kb upstream of the rRNA transcription unit (red arrow in Figure 8-51) is bound by a subunit of NoRC and is required for transcriptional silencing by assembly of the gene into heterochromatin. The pRNA is believed to target NoRC to Pol I promoter regions by forming an RNA:DNA triplex with the T_0 terminator sequence. This creates a binding site for the DNA methyl transferase DNMT3b, which methylates the critical CpG in the upstream promoter element.

Initiation by Pol III

Unlike promoters for protein-coding genes and pre-rRNA genes, the promoter regions of tRNA and 5S-rRNA genes lie entirely within the transcribed sequence (Figure 8-52a, b). Two such *internal* promoter elements, termed the *A box* and the *B box*, are present in all tRNA genes. These highly conserved sequences not only function as promoters but also encode two invariant portions of eukaryotic tRNAs that are required for protein synthesis. In 5S-rRNA genes, a single internal control region, the *C box*, acts as a promoter.

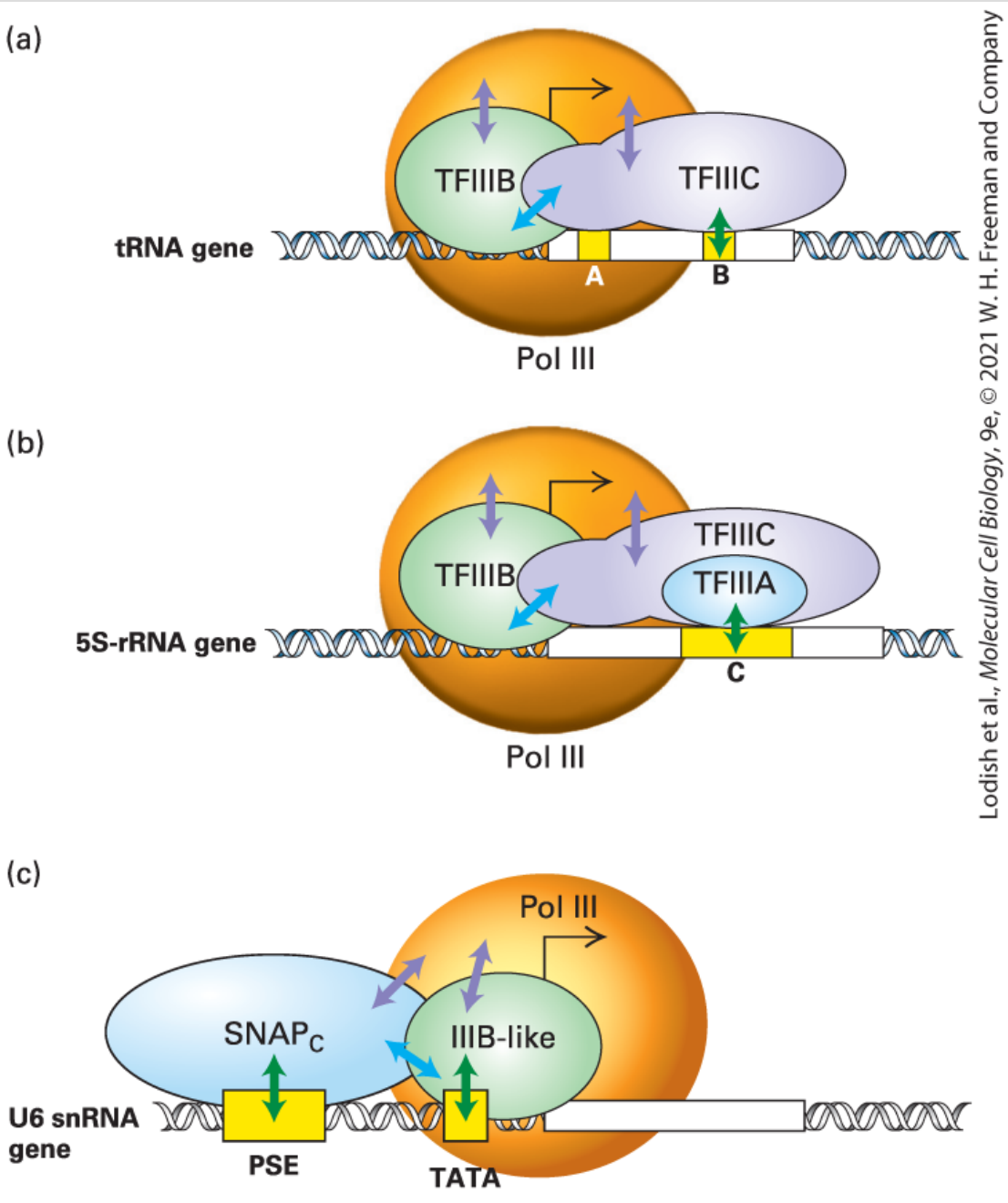


FIGURE 8-52 Transcription-control elements in genes transcribed by RNA polymerase III. Both tRNA (a) and 5S-rRNA (b) genes contain internal promoter elements (yellow) located downstream from the start site and named A, B, and C boxes, as indicated. Assembly of transcription initiation complexes on these genes begins with the binding of Pol III–specific general transcription factors TFIIIA, TFIIIB, and TFIIIC to these control elements. Green arrows indicate strong, sequence-specific protein-DNA interactions. Blue arrows indicate interactions between general

transcription factors. Purple arrows indicate interactions between general transcription factors and Pol III. (c) Transcription of the U6 snRNA gene in mammals is controlled by an upstream promoter with a TATA box bound by the TBP subunit of a specialized form of TFIIIB with an alternative BRF subunit and an upstream regulatory element called the *PSE* bound by a multisubunit factor called $SNAP_C$. See L. Schramm and N. Hernandez, 2002, *Genes Dev.* **16**:2593.

Three general transcription factors are required for Pol III to initiate transcription of tRNA and 5S-rRNA genes in vitro. Two multimeric factors, TFIIIC and TFIIIB, participate in initiation at both tRNA and 5S-rRNA promoters, while a third factor, TFIIIA, is required for initiation at 5S-rRNA promoters. As with assembly of Pol I and Pol II initiation complexes, the Pol III general transcription factors bind to promoter DNA in a defined sequence.

The N-terminal half of one TFIIIB subunit, called *BRF* (for TFII*B*-*r*elated *f*actor), is similar in sequence to TFIIB (a Pol II factor). This similarity suggests that BRF and TFIIB perform a similar function in initiation, namely, to assist in separating the template DNA strands at the transcription start site. Once TFIIIB has bound to either a tRNA or a 5S-rRNA gene, Pol III can bind and initiate transcription in the presence of ribonucleoside triphosphates. The BRF subunit of TFIIIB interacts specifically with one of the polymerase subunits unique to Pol III, accounting for initiation by this specific nuclear RNA polymerase.

Another of the three subunits composing TFIIIB is TBP, which we can now see is a component of a general transcription factor for all three eukaryotic nuclear RNA polymerases. The finding that TBP participates in transcription initiation by Pol I and Pol III was surprising, since the promoters recognized by these enzymes often do not contain TATA boxes. Nonetheless, in the case of Pol III transcription, the TBP subunit of TFIIIB interacts with DNA about 30 bp upstream of the transcription start site. The protein-DNA interaction is similar to the way TBP interacts with TATA boxes.

Pol III also transcribes genes for small stable RNAs with upstream promoters containing a TATA box. One example is the gene for U6 snRNA, which is involved in pre-mRNA splicing (see Chapter 9). In mammals, this gene contains an upstream promoter element called the *PSE* in addition to the TATA box (Figure 8-52c). The PSE is bound by a multisubunit complex called $SNAP_C$, while the TATA box is bound by the TBP subunit of a specialized form of TFIIIB containing an alternative BRF subunit.

MAF1 is a specific inhibitor of Pol III transcription that functions by interacting with the BRF subunit of TFIIIB and with Pol III. Its function is regulated by controlling its import from the cytoplasm into the nucleus. MAF1 transport into the nucleus is activated by phosphorylations of MAF1 in response to the mTORC1 protein kinase that responds to cell stress and nutrient deprivation (see Chapters 16 and 25). In mammals, Pol III transcription is also

repressed by the tumor suppressors p53 and the retinoblastoma (Rb) family of proteins. In humans, there are two genes encoding the Pol III subunit RPC32. One of these is expressed specifically in replicating cells, and its forced expression can contribute to oncogenic transformation of cultured human fibroblasts, suggesting that altered regulation of transcription by Pol III with this subunit is oncogenic.

KEY CONCEPTS OF SECTION 8.7

Other Eukaryotic Transcription Systems

- The process of transcription initiation by Pol I and Pol III is similar to that by Pol II but requires different general transcription factors, is directed by different promoter elements, and does not require hydrolysis of ATP β-γ phosphodiester bonds to separate the DNA strands at the start site as Pol II transcription does.
- Pol I transcribes only a single RNA, the 45S precursor of 18S, 5.8S, and 28S rRNA, from multiple copies of the pre-rRNA gene.
- Pol III transcribes tRNAs from promoters within the genes that encode the tRNA regions common to all tRNAs. This internal promoter is bound by transcription factor TFIIIC, which in turn binds TFIIIB, a multisubunit factor that includes the TATA box–binding protein, TBP, which associates with the tRNA gene about 30 bp upstream of the transcription start site.
- Pol III transcribes 5s rRNA directed by a promoter within the 5S-rRNA coding region that is bound by transcription factor TFIIIA. TFIIIA then associates with TFIIIC and TFIIIB, which interact with Pol III in a manner similar to their interactions in tRNA transcription.
- Additional small stable RNAs, several with as yet unknown functions, are transcribed by Pol III as directed by TBP-containing transcription factors that bind immediately upstream of the genes (see Figure 8-52).
- Pol III transcription is regulated by a specific inhibitor, MAF1, whose transport from the cytoplasm into the nucleus is controlled in response to nutrient availability.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Analyze the Data 8-1: Coordination of Eukaryotic Polymerases
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

activation domain
activator
basic helix-loop-helix (bHLH)
bromodomain
carboxy-terminal domain (CTD)
chromatin
chromatin immunoprecipitation
chromatin-mediated repression
co-activator
constitutive
co-repressor
CpG island

DNA-binding domain
DNase I footprinting
downstream
electrophoretic mobility shift assay (EMSA)
elongation
enhanceosome
enhancer
epigenetics
euchromatin
gene control
heat-shock gene
helicase
heterochromatin
histone deacetylation
homeodomain
initiation
initiator
leucine zipper
long noncoding RNA (lncRNA)
Mediator
nuclear receptor
nucleosome
promoter
promoter-proximal element
reporter gene
repression domain
repressor
RNA polymerase II

silencer sequence
SWI/SNF chromatin- remodeling complex
Tat
TATA box
TATA box–binding protein (TBP)
termination
transcription-control region
transcription factor
upstream
upstream activating sequence (UAS)
zinc finger

Review the Concepts

1. What types of genes are transcribed by RNA polymerases I, II, and III? Design an experiment to determine whether a specific gene is transcribed by RNA polymerase II.
2. The CTD of the largest subunit of RNA polymerase II can be phosphorylated at multiple serine residues. What are the conditions that lead to the phosphorylated versus nonphosphorylated RNA polymerase II CTD?
3. What do TATA boxes, initiators, and CpG islands have in common? Which was the first of these to be identified? Why?
4. Describe the methods used to identify the location of transcription-control elements in promoter-proximal regions of genes.

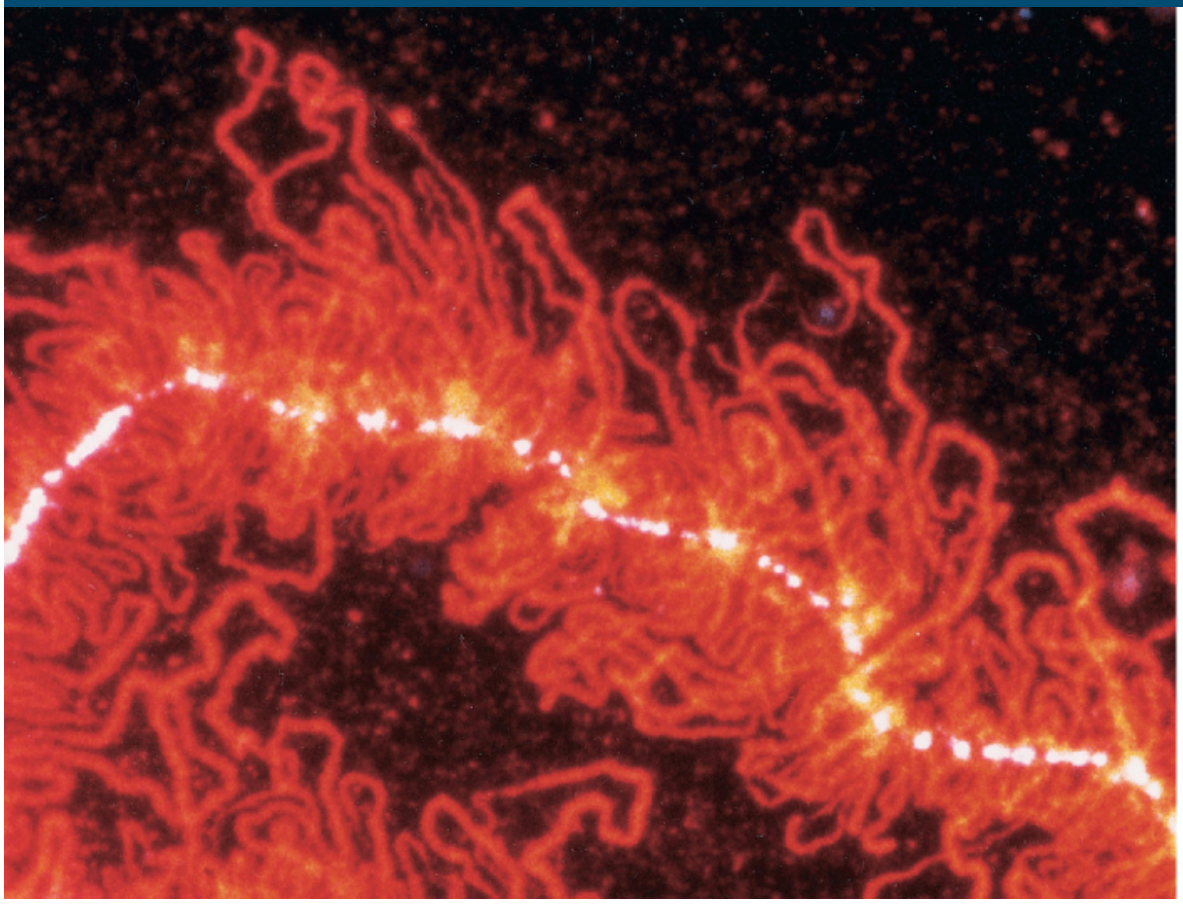
5. What is the difference between a promoter-proximal element and a distal enhancer? What are the similarities?
6. Describe the methods used to identify the location of DNA-binding proteins in the regulatory regions of genes.
7. Describe the structural features of transcription activator and repressor proteins.
8. Give two examples of how gene expression may be repressed without altering the coding sequence.
9. Using CREB and nuclear receptors as examples, compare and contrast the structural changes that take place when these transcription factors bind to their co-activators.
10. What general transcription factors associate with an RNA polymerase II promoter in addition to the polymerase? In what order do they bind in vitro? What structural change occurs in the DNA when an "open" transcription initiation complex is formed?
11. Expression of recombinant proteins in yeast is an important tool for biotechnology companies that produce new drugs for human use. In an attempt to get a new gene *X* expressed in yeast, a researcher has integrated gene *X* into the yeast genome near a telomere. Will this strategy result in good expression of gene *X*? Why or why not? Would the outcome of this experiment differ if the experiment had been performed in a yeast line containing mutations in the H3 or H4 histone tails?
12. You have isolated a new protein called *STICKY*. You can predict from comparisons with other known proteins that STICKY contains a bHLH domain and a Sin3-interacting

domain. Predict the function of STICKY and explain the importance of these domains in STICKY function.

13. Lower eukaryotes such as yeast have transcription-control elements called *upstream activating sequences*. What are the comparable sequences found in higher eukaryotic species?
14. You are curious to identify the region of the gene *X* sequence that serves as an enhancer for gene expression. Design an experiment to investigate this issue.
15. Some organisms have mechanisms in place that will override transcription termination. One such mechanism using the Tat protein is employed by the HIV retrovirus. Explain why Tat is therefore a good target for HIV vaccination.
16. Upon identification of the DNA-regulatory sequence responsible for translating a given gene, you note that it is enriched with CG sequences. Is the corresponding gene likely to be a highly expressed transcript?
17. Name four major classes of DNA-binding proteins that are responsible for controlling transcription, and describe their structural features.

CHAPTER 9

Post-Transcriptional Gene Control



Courtesy of M. Roth and J. Gall.

Portion of a "lampbrush chromosome" from an oocyte of the newt *Nophthalmus viridescens*. The hnRNP protein associated with nascent RNA transcripts fluoresces red after staining with a monoclonal antibody.

OUTLINE

9.1 Processing of Eukaryotic Pre-mRNA
9.2 Regulation of Pre-mRNA Processing
9.3 Transport of mRNA Across the Nuclear Envelope
9.4 Cytoplasmic Mechanisms of Post-Transcriptional Control
9.5 Processing of rRNA and tRNA
9.6 Nuclear Bodies Are Functionally Specialized Nuclear Domains

---

In the previous chapter, we saw that most genes are regulated at the first step in gene expression — transcription — by regulating the assembly of the transcription preinitiation complex on a promoter-DNA sequence and regulating transcription elongation in the promoter-proximal region. Once transcription has been initiated, synthesis of the encoded RNA requires that RNA polymerase transcribe the entire gene and not terminate prematurely. Moreover, the initial primary transcripts produced from eukaryotic genes must undergo various processing reactions to yield the corresponding functional RNAs. For mRNAs, the 5′ cap structure necessary for translation must be added (see Figure 5-26), introns must be spliced out of **pre-mRNAs** (Table 9-1), and the 3′ end must be polyadenylated (see Figure 5-27). Once formed in the nucleus, mature, functional RNAs are exported to the cytoplasm as components of ribonucleoproteins. Processing of mRNAs in the nucleus, their export from the nucleus, and their transport to specific loci in the cytoplasm offer opportunities for further regulating gene expression after the initiation of transcription.

TABLE 9-1 • RNAs Discussed in Chapter 9

mRNA	Fully processed messenger RNA with $5'$ cap, introns removed by RNA splicing, and a poly(A) tail.
pre-mRNA	An mRNA precursor containing introns and not cleaved at the poly(A) site.
hnRNA	Heterogeneous nuclear RNAs. These RNAs include pre-mRNAs and RNA-processing intermediates containing one or more introns.
snRNA	Five small nuclear RNAs that function in the removal of introns from pre-mRNAs by RNA splicing, plus two small nuclear RNAs that substitute for the first two at rare introns.
pre-tRNA	A tRNA precursor containing additional transcribed bases at the $5'$ and $3'$ ends compared with the mature tRNA. Some pre-tRNAs also contain an intron in the anticodon loop.
pre-rRNA	The precursor to mature 18S, 5.8S, and 28S ribosomal RNAs. The mature rRNAs are processed from this long precursor RNA molecule by cleavage, removal of bases from the ends of the cleaved products, and modification of specific bases.
snoRNA	Small nucleolar RNAs. These RNAs base-pair with complementary regions of the pre-rRNA molecule, directing cleavage of the RNA chain and modification of bases during maturation of the rRNAs.
siRNA	Short interfering RNAs, ~ 22 bases long, that are each perfectly complementary to a sequence in an mRNA. Together with associated proteins, siRNAs cause cleavage of the target RNA, leading to its rapid degradation.
miRNA	MicroRNAs, ~ 22 bases long, that base-pair extensively, but not completely, with mRNAs, especially over bases 2 to 7 at the $5'$ end of the miRNA (the “seed” sequence). This pairing inhibits translation of the target mRNA and targets it for degradation.

Recently, the vast amount of sequence data of human mRNAs expressed in different tissues and at various times during embryogenesis and cellular differentiation has revealed that ~95 percent of human genes give rise to alternatively spliced mRNAs. These alternatively spliced mRNAs encode related proteins with differences in sequences limited to specific functional domains. In many cases, alternative RNA splicing is regulated to meet the need for a specific protein isoform in a specific cell type. Given the complexity of pre-mRNA splicing, it is not surprising that mistakes are occasionally made, giving rise to mRNA precursors with improperly spliced exons. However, eukaryotic cells have evolved RNA-surveillance mechanisms that prevent the transport of incorrectly processed RNAs to the cytoplasm or lead to their degradation if they are transported.

Additional control of gene expression can occur in the cytoplasm. In the case of protein-coding genes, for instance, the amount of protein produced depends on the stability of the corresponding mRNAs in the cytoplasm and the rate of their translation. For example, during an immune response, lymphocytes communicate by secreting polypeptide hormones called *cytokines* that signal neighboring lymphocytes through cytokine receptors that span their plasma membranes (Chapter 24). It is important for lymphocytes to synthesize and secrete cytokines in short bursts. This is possible because cytokine mRNAs are extremely unstable. Consequently, the concentration of the mRNA in the cytoplasm falls rapidly once its

synthesis is stopped. In contrast, mRNAs encoding proteins required in large amounts that function over long periods, such as ribosomal proteins, are extremely stable so that multiple polypeptides can be transcribed from each mRNA.

In addition to regulation of pre-mRNA processing, nuclear export, and translation, the cellular locations of many, if not most, mRNAs are regulated so that newly synthesized protein is concentrated where it is needed. Particularly striking examples of this occur in the nervous systems of multicellular animals. Some neurons in the human brain generate more than 1000 separate synapses with other neurons. During the process of learning, synapses that fire more frequently than others increase in size many times, while other less active synapses made by the same neuron do not. This can occur because mRNAs encoding critical synapse proteins are stored at all synapses, but translation of these localized, stored mRNAs is regulated at each synapse independently by the frequency at which that synapse fires. In this way, synthesis of synapse-associated proteins can be regulated independently at each of the many synapses made by the same neuron (see Chapter 23).

Another type of gene regulation involves microRNAs (miRNAs), which regulate the stability and translation of specific target mRNAs to which they base pair in multicellular animals and plants. Analyses of these short miRNAs in various human tissues indicate that ~1900 miRNAs are expressed in the multiple types of human

cells. Although some have recently been discovered to function through inhibition of target gene expression in the appropriate tissue and at the appropriate time in development, the functions of the vast majority of human miRNAs are unknown and are the subject of a growing area of research. A closely related process called *RNA interference (RNAi)* is used by some organisms to degrade viral RNAs following infection and in the degradation of transposon-encoded RNAs. RNAi is extensively utilized in biomedical research: designed short interfering RNAs (siRNA) can be used to inhibit the translation of specific mRNAs experimentally by a process called *RNA knockdown*. This makes it possible to inhibit the function of any desired gene, even in organisms that are not amenable to classic genetic methods for isolating mutants.

We refer to all the mechanisms that regulate gene expression following transcription as *post-transcriptional gene control* (Figure 9-1). Because the stability and translation rate of an mRNA contribute to the amount of protein expressed from a gene, these post-transcriptional processes are important components of gene control. Indeed, the protein output of a gene is regulated at every step in the life of an mRNA from the initiation of its synthesis to its degradation. Thus genetic regulatory processes act on RNA as well as DNA. In this chapter, we consider the events that occur in the processing of mRNA following transcription initiation and promoter-proximal elongation as well as mechanisms discovered that regulate these steps in mRNA processing. In the last section, we

briefly discuss the processing of primary transcripts produced from genes encoding rRNAs and tRNAs.

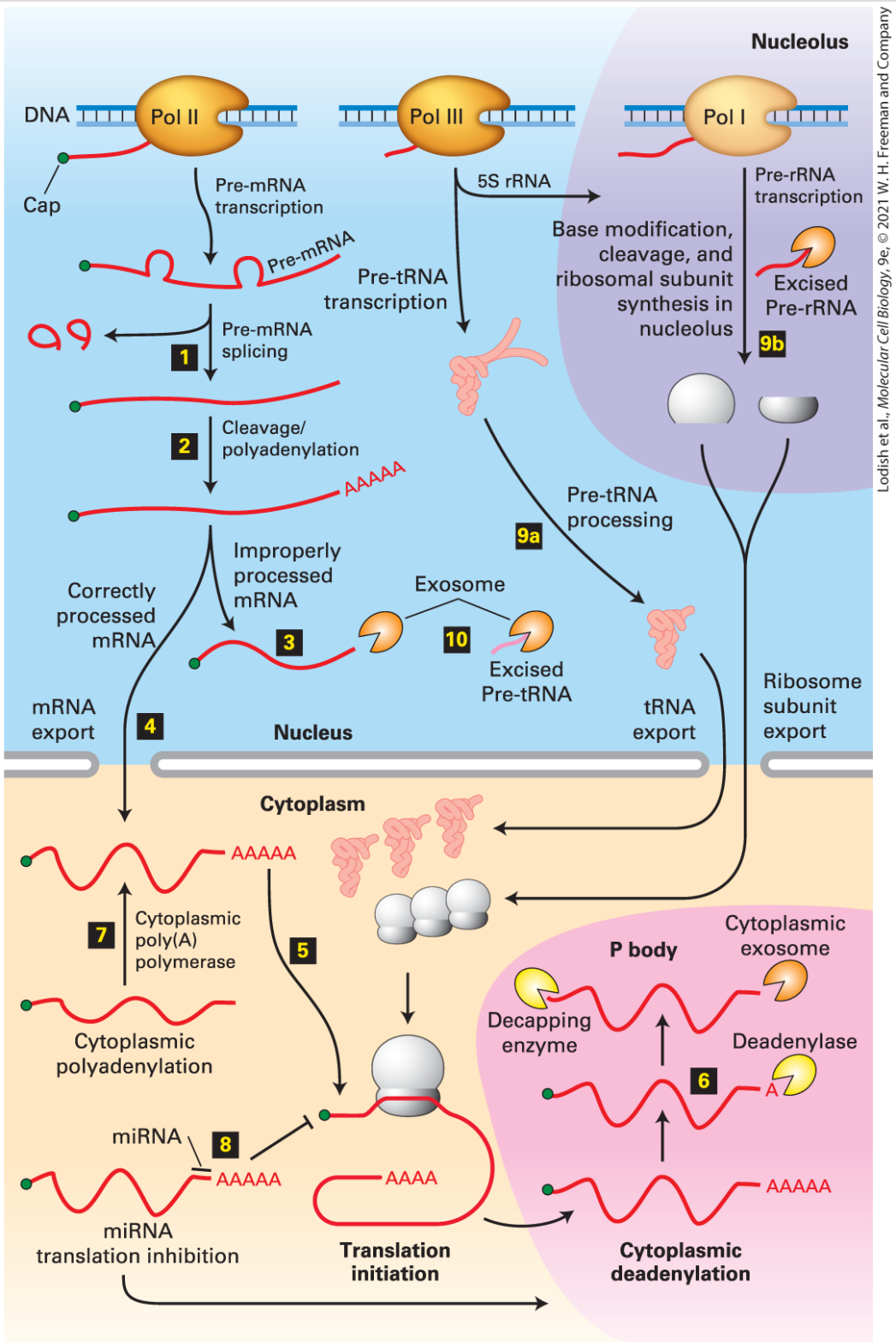


FIGURE 9-1 Overview of RNA processing and post-transcriptional gene control. Nearly all cytoplasmic RNAs are processed from primary transcripts in

the nucleus before they are exported to the cytoplasm. Major processing steps are diagrammed for RNAs transcribed by RNA polymerases I, II, and III. For protein-coding genes transcribed by RNA polymerase II, post-transcriptional gene control can be exerted through (step **1**) the choice of alternative exons during pre-mRNA splicing and (step **2**) the choice of alternative poly(A) sites. Improperly processed mRNAs are blocked from export to the cytoplasm and degraded (step **3**) by a large complex called the *exosome* that contains internal ribonucleases. Once exported to the cytoplasm (step **4**), translation initiation factors bind to the mRNA $5'$-cap cooperatively with poly(A)-binding protein C bound to the poly(A) tail and initiate translation (see Figure 5-36) (step **5**). (Step **6**) mRNA is degraded in the cytoplasm by deadenylation and decapping followed by degradation by cytoplasmic exosomes. Much of this mRNA degradation takes place in regions of the cytoplasm with high RNA concentration called "P bodies." The degradation rate of each mRNA is controlled, thereby regulating the mRNA concentration and, consequently, the amount of protein translated. Some mRNAs are synthesized without long poly(A) tails. Their translation is regulated by (step **7**) controlling the synthesis of a long poly(A) tail by a cytoplasmic poly(A) polymerase. Translation is also regulated by other mechanisms including miRNAs (step **8**). When expressed, these ~ 22-nucleotide RNAs inhibit translation of mRNAs to which they hybridize, usually in the $3'$ untranslated region. tRNAs and rRNAs are also synthesized as precursor RNAs that must be (step **9a**, **b**) processed before they are functional. Regions of precursors cleaved from the mature RNAs are degraded by nuclear exosomes (step **10**). See J. Houseley et al., 2006, *Nat. Rev. Mol. Cell Biol.* **7**:529.

9.1 Processing of Eukaryotic Pre-mRNA

In this section, we take a closer look at how eukaryotic cells convert the initial primary transcript synthesized by RNA polymerase II into a functional mRNA. Three major events occur during the process: **5′ capping**, **3′ cleavage/polyadenylation**, and **RNA splicing** (Figure 9-2). Adding these specific modifications to the 5′ and 3′ ends of the pre-mRNA is important to protect it from enzymes that quickly digest uncapped RNAs generated by RNA processing, such as spliced-out introns and RNA transcribed downstream from a polyadenylation site. The 5′ cap and 3′ poly(A) tail distinguish pre-mRNA molecules from the many other kinds of RNAs in the nucleus. Pre-mRNA molecules (see Table 9-1) are bound by nuclear proteins that function in mRNA export to the cytoplasm. After mRNAs are exported to the cytoplasm, they are bound by a set of cytoplasmic proteins that stimulate translation and are critical for mRNA stability in the cytoplasm. Prior to nuclear export, introns must be removed to generate the correct coding region of the mRNA. In vertebrates, including humans, **alternative splicing** is intricately regulated in order to include alternative functional domains in closely related proteins, producing a considerable expansion of the proteome of these organisms.

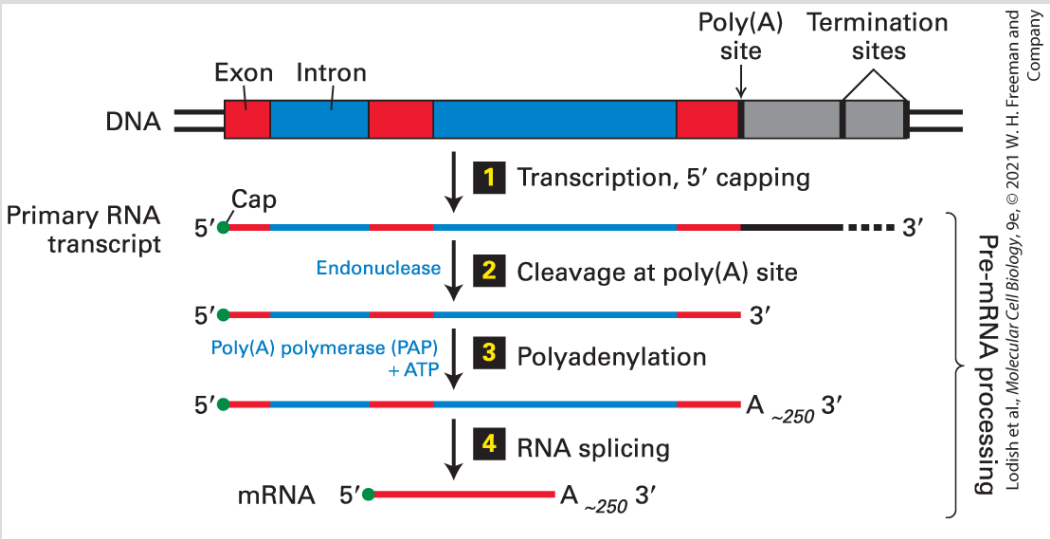


FIGURE 9-2 Overview of mRNA processing in eukaryotes. Shortly after RNA polymerase II initiates transcription at the first nucleotide of the first exon of a gene, the 5′ end of the nascent RNA is capped with 7-methylguanylate (step 1). Transcription by RNA polymerase II terminates at any one of multiple termination sites downstream from the poly(A) site, which is located at the 3′ end of the final exon. After the primary transcript is cleaved at the poly(A) site (step 2), a string of adenosine (A) residues is added (step 3). The poly(A) tail contains ~250 A residues in mammals, ~150 in insects, and ~100 in yeasts. For short primary transcripts with few introns, splicing (step 4) usually follows cleavage and polyadenylation, as shown. For large genes with multiple introns, introns often are spliced out of the nascent RNA during its transcription, i.e., before transcription of the gene is complete. Note that the 5′ cap and sequence adjacent to the poly(A) tail are retained in mature mRNAs. The diagram shown represents processing of human β-globin RNA.

The pre-mRNA processing events of capping, splicing, and polyadenylation occur in the nucleus as the nascent mRNA precursor is being transcribed. Thus pre-mRNA processing is *co-transcriptional*: it occurs while the RNA is being transcribed. As the RNA emerges from the surface of RNA polymerase II, its 5′ end is

immediately modified by the addition of the 5′ cap structure found on all mRNAs (see Figure 5-26). As the nascent pre-mRNA continues to emerge from the surface of the polymerase, it is immediately bound by members of a complex group of abundant RNA-binding proteins called *hnRNP proteins* that assist in RNA splicing and export of the fully processed mRNA through nuclear pore complexes into the cytoplasm. Some of these proteins remain associated with the mRNA in the cytoplasm, but most either remain in the nucleus or shuttle back into the nucleus shortly after the mRNA is exported to the cytoplasm. Cytoplasmic RNA-binding proteins are exchanged for the nuclear ones. Consequently, eukaryotic mRNAs never occur as free RNA molecules in the cell but are always associated with protein as ***ribonucleoprotein (RNP) complexes***.

Nuclear nascent *pre-mRNAs* associated with hnRNP proteins are referred to as *pre-mRNPs.* They are capped and spliced as they are transcribed. Then, following cleavage and polyadenylation at the 3′ end, they are referred to as *nuclear mRNPs.* Following the exchange of proteins that accompany export to the cytoplasm, they are called *cytoplasmic mRNPs.* Although we frequently refer to pre-mRNAs and mRNAs, it is important to remember that in eukaryotes they are always associated with proteins as RNP complexes.

The 5′ Cap Is Added to Nascent RNAs Shortly After Transcription Initiation

As the nascent RNA transcript emerges from the RNA exit channel of RNA polymerase II and reaches a length of ~25 nucleotides, a protective cap composed of 7-methylguanosine and methylated riboses is added to the 5′ end of eukaryotic mRNAs (see Figure 5-26). The 5′ *cap* marks RNA molecules as mRNA precursors and protects them from RNA-digesting enzymes (5′-exoribonucleases) in the nucleus and cytoplasm. This initial step in RNA processing is catalyzed by a dimeric capping enzyme, which associates with the phosphorylated carboxyl-terminal domain (CTD) of RNA polymerase II. Recall that the CTD becomes phosphorylated by the TFIIH general transcription factor at multiple serines at the 5 position in the CTD heptapeptide (seven residue) repeat during transcription initiation (see Figure 8-7(b) red arrow, Figure 8-8). Binding to the phosphorylated CTD stimulates the activity of the capping enzymes so that they target RNAs containing a 5′-triphosphate that emerge from RNA polymerase II and not RNAs transcribed by RNA polymerases I or III, which do not have a CTD. This is important because pre-mRNA synthesis accounts for only ~80 percent of the total RNA synthesized in replicating cells. The other 20 percent is pre-ribosomal RNA (**pre-rRNA**), which is transcribed by RNA polymerase I, and 5S rRNA, tRNAs, and other stable small RNAs transcribed by RNA polymerase III.

One subunit of the capping enzyme removes the γ-phosphate from the 5′ end of the nascent RNA (Figure 9-3). Another domain of this subunit transfers the GMP moiety from GTP to the 5′-diphosphate of

the nascent transcript, creating the unusual guanosine 5′, 5′-triphosphate ester structure. In the final steps, separate enzymes transfer methyl groups from *S*-adenosylmethionine to the N_7 position of the guanine and one or two of the 2′ oxygens of the ribose at the 5′ end of the nascent RNA.

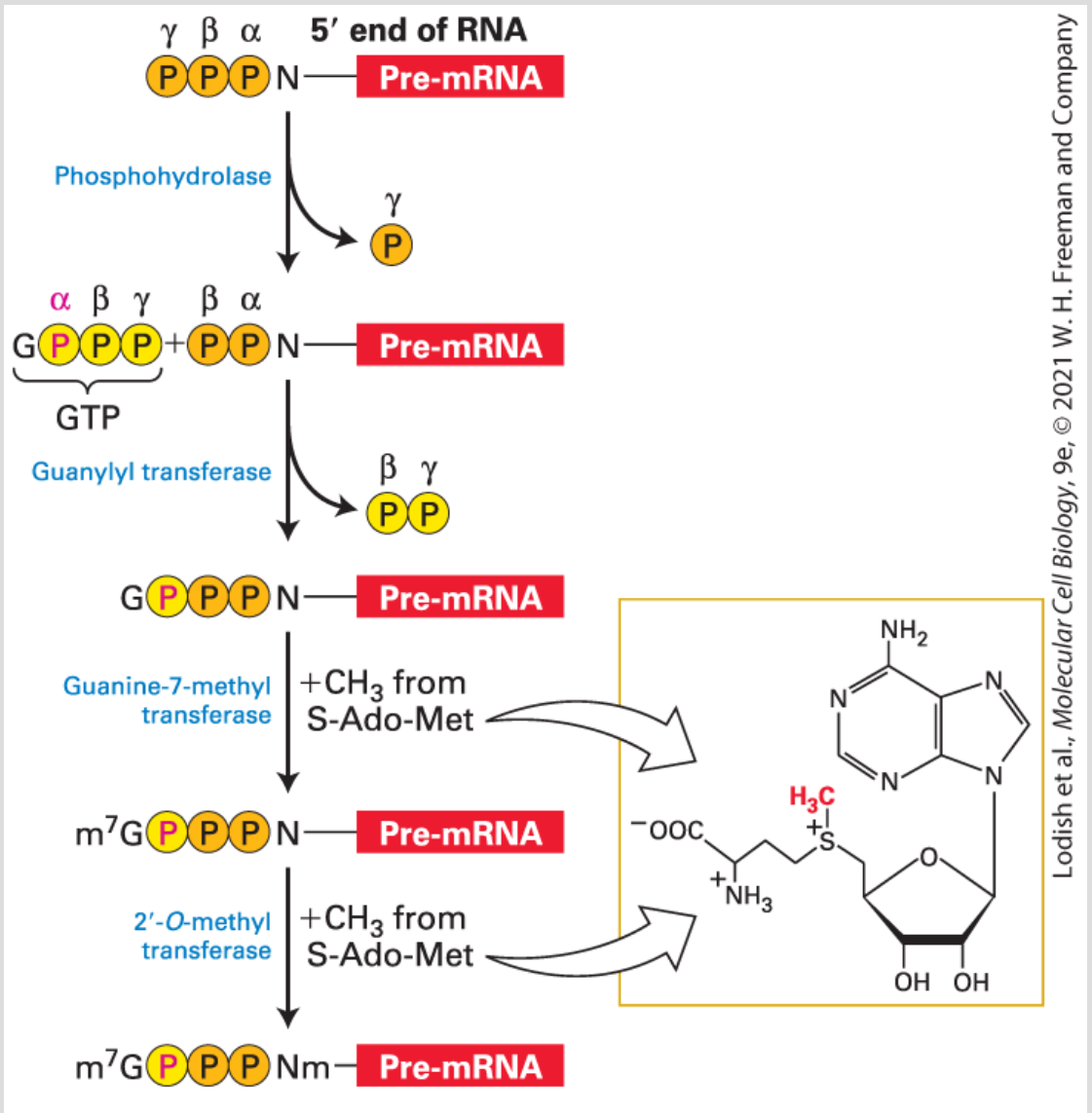


FIGURE 9-3 Synthesis of 5′-cap on eukaryotic mRNAs. The 5′ end of a nascent RNA contains a 5′-triphosphate from the initiating NTP. The γ-phosphate is removed in the first step of capping, while the remaining α- and β-phosphates (orange) remain

associated with the cap. The third phosphate of the 5′,5′-triphosphate ester is derived from the α-phosphate of the GTP that donates the guanine. The methyl donor for methylation of the cap guanine and the first one or two riboses of the mRNA is S-adenosylmethionine (S-Ado-Met). See S. Venkatesan and B. Moss, 1982, *Proc. Nat'l. Acad. Sci. USA* **79**:304.

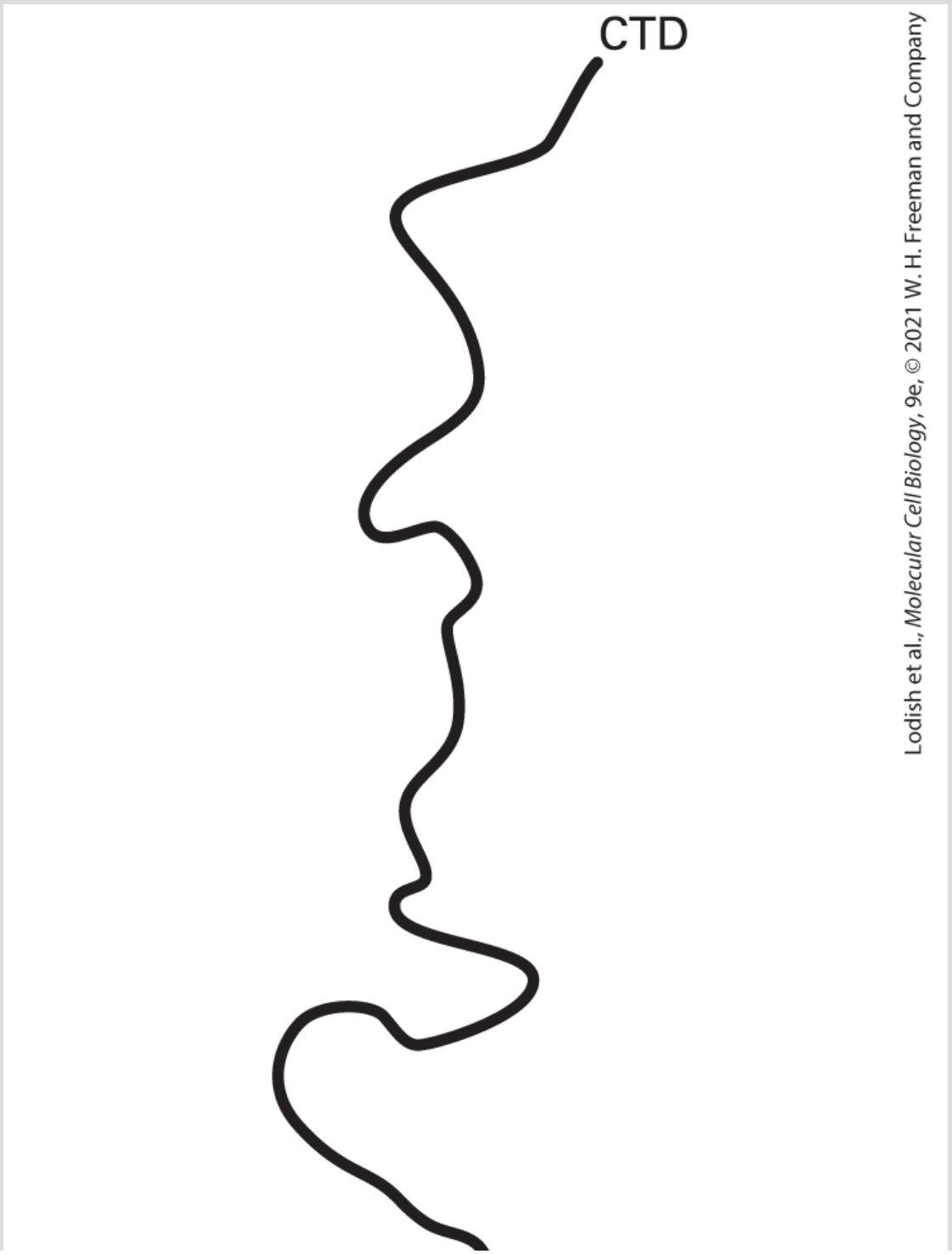
Considerable evidence indicates that capping of the nascent transcript is coupled to elongation by RNA polymerase II so that all transcripts generated by this polymerase are capped during the earliest phase of elongation. As discussed in Chapter 8, during the initial phase of transcription in metazoans, the polymerase elongates the nascent transcript very slowly due to association of NELF (*n*egative *el*ongation *f*actor) with a complex of RNA polymerase II and DSIF in the promoter-proximal region (see Figure 8-3 *bottom*, Figure 8-15a). Once the 5′ end of the nascent RNA is capped, phosphorylation of the RNA polymerase CTD at position 2 in the heptapeptide repeat and of NELF and DSIF by the cyclin T-CDK9 protein kinase causes the release of NELF (Figure 8-16). This allows the PAF elongation complex to associate with RNA polymerase II (Figure 8-15b), which then enters into a faster mode of elongation that rapidly transcribes away from the promoter. The net effect of this mechanism is that the polymerase waits for the nascent RNA to be capped before elongating at a rapid rate.

Chain Elongation by RNA Polymerase II Is Coupled to the Presence of RNA

Processing Factors

How is RNA processing coupled with the transcription of a pre-mRNA so that only Pol II transcripts are subjected to 5′ capping, RNA splicing, and polyadenylation? The key lies in the long carboxyl-terminal domain (CTD) of RNA polymerase II, which, as discussed in Chapter 8, is composed of multiple repeats of a heptapeptide sequence. When fully extended, the CTD domain in the yeast enzyme is about 65 nm long (Figure 9-4); the CTD in human RNA polymerase II is about twice as long. The remarkable length of the CTD apparently allows multiple proteins to associate simultaneously with a single RNA polymerase II molecule. As mentioned earlier, the enzymes that add the 5′ cap to nascent transcripts associate with the CTD phosphorylated on multiple serines at the fifth position in the heptapeptide repeat (Ser-5) during or shortly after transcription initiation by a subunit of TFIIH. In addition, RNA splicing and polyadenylation factors associate with the phosphorylated CTD. As a consequence, these processing factors are present at *high local concentrations* when splice sites and poly(A) signals are transcribed by the polymerase, enhancing the rate and specificity of RNA processing. In a reciprocal fashion, the association of hnRNP proteins, discussed below, with the nascent RNA enhances the interaction of RNA polymerase II with elongation factors such as DSIF and cyclin T-CDK9 (also called “P-TEFb”) (Figure 8-16), increasing the rate of transcription. As a consequence, the rate of transcription is coordinated with the rate of nascent RNA

association with hnRNP proteins and RNA-processing factors. This mechanism ensures that a pre-mRNA is not synthesized unless the machinery for processing it is properly positioned.



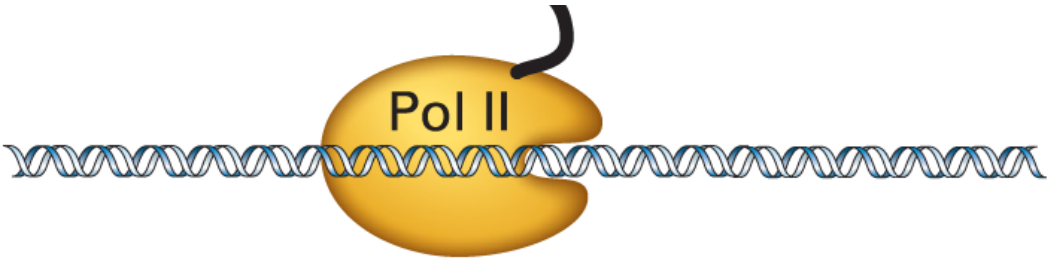


FIGURE 9-4 Schematic diagram of the length of the RNA polymerase II CTD relative to the globular domain of the polymerase. The length of the yeast RNA polymerase II carboxyl-terminal domain (CTD) and the linker region that connects it to the polymerase is shown relative to the globular domain of the polymerase. The CTD of mammalian RNA polymerase II is twice as long. In its extended form, the CTD can associate with multiple RNA-processing factors simultaneously. See P. Cramer, D. A. Bushnell, and R. D. Kornberg, 2001, *Science* **292**:1863.

A Diverse Set of Proteins with Conserved RNA-Binding Domains Associate with Pre-mRNAs

As noted earlier, neither nascent RNA transcripts from protein-coding genes nor the intermediates of mRNA processing, collectively referred to as *pre-mRNA*, exist as free RNA molecules in the nuclei of eukaryotic cells. From the time nascent transcripts first emerge from RNA polymerase II until mature mRNAs are transported into the cytoplasm, the RNA molecules are associated with nuclear proteins. The most abundant of these are the major protein components of *heterogeneous ribonucleoprotein particles (hnRNPs),* which contain *heterogeneous nuclear RNA (hnRNA),* a

collective term referring to pre-mRNA- and mRNA-processing intermediates of various sizes. These hnRNP proteins contribute to further steps in RNA processing, including splicing, polyadenylation, and export through nuclear pore complexes to the cytoplasm.

Researchers identified hnRNP proteins by first exposing cultured cells to high-dose UV irradiation, which causes covalent cross-links to form between RNA bases and closely associated proteins. Chromatography of nuclear extracts from treated cells on an oligo-dT cellulose column, which binds RNAs with a poly(A) tail, was used to recover the proteins that had become cross-linked to nuclear polyadenylated RNA. These included a complex set of abundant hnRNP proteins ranging in size from ~30 to ~120 kDa, collectively referred to as *hnRNP proteins*.

Like transcription factors, most hnRNP proteins have a modular structure. They contain one or more *RNA-binding domains* and at least one other domain, often an intrinsically disordered protein domain (see Figure 3-13 and associated text), that interacts with other proteins. Several different classes of RNA-binding domains are associated with hnRNP proteins, identified by testing deletion mutants for their ability to bind RNA.

Functions of hnRNP Proteins

These RNA-associated proteins have three key functions. First, the association of pre-mRNAs with the RNA-binding domains of hnRNP proteins prevents the pre-mRNAs from forming short secondary structures dependent on base pairing of closely spaced, short complementary regions. Pre-mRNAs associated with hnRNP proteins present a more uniform substrate for subsequent processing steps than would free, unbound pre-mRNAs, in which each mRNA forms a unique secondary structure due to its specific sequence. The RNA-binding domains of hnRNPs usually interact preferentially with short RNA sequences of three to four bases.

The second key function is in regulating pre-mRNA splicing. In mammals, alternatively spliced mRNAs are expressed from almost all genes (~95 percent). Regulation of the use of one or the other of the alternative splice sites present in a pre-mRNA transcribed from these genes depends in part on binding of specific hnRNP proteins. The binding of hnRNP proteins to a pre-mRNA near an RNA splice site can inhibit binding of factors required for pre-mRNA splicing and hence direct splicing to an alternative splice site (see below).

The third function is in mRNA transport. Cell-fusion experiments have shown that some hnRNP proteins remain localized in the nucleus, whereas others cycle in and out of the cytoplasm. This was the initial evidence suggesting that the hnRNPs that cycle in and out of the nucleus function in the transport of mRNA into the cytoplasm (Figure 9-5).

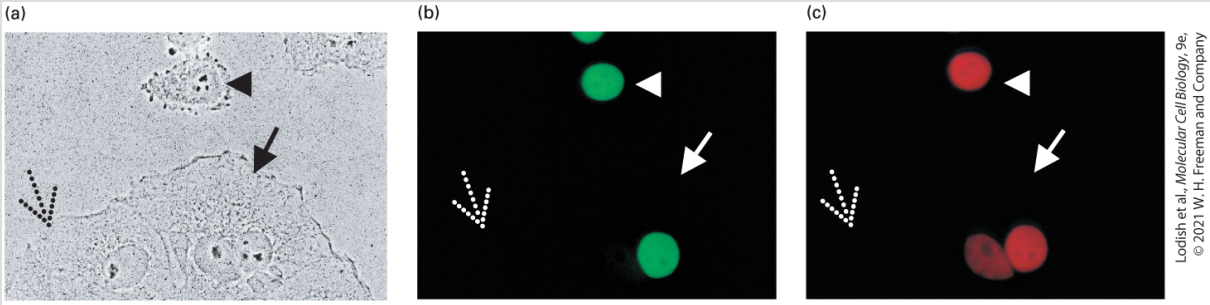


FIGURE 9-5 Human hnRNP A1 protein can cycle in and out of the nucleus, but human hnRNP C protein cannot. Cultured HeLa cells and *Xenopus* cells were fused by treatment with polyethylene glycol, producing heterokaryons containing nuclei from each cell type. Cyclohexamide was added immediately after fusion to prevent protein synthesis. After 2 hours, the cells were fixed and stained with fluorescent-labeled antibodies specific for human hnRNP C (green) and A1 proteins (red). These antibodies do not bind to the homologous Xenopus proteins. (a) A fixed preparation viewed by phase-contrast microscopy includes one unfused HeLa cell (arrowhead) and one unfused Xenopus cell (dotted arrow). Note that the nucleus in the Xenopus cell contains one prominent nucleolus and is more oval in shape than the round HeLa cell nucleus containing three nuclei. A fused heterokaryon (solid arrow) in this micrograph contains a round HeLa cell nucleus to the right of the oval-shaped Xenopus nucleus with one nucleolus. (b, c) When the same preparation was viewed by fluorescence microscopy, human hnRNP C protein appeared green and human hnRNP A1 protein appeared red. Note that the unfused Xenopus cell on the left is unstained, confirming that the antibodies are specific for the human proteins. In the heterokaryon, human hnRNP C protein (green) appears only in the HeLa-cell nucleus (b), whereas the human A1 protein (red) appears in both the HeLa-cell nucleus and the Xenopus nucleus in the heterokaryon (c). Since protein synthesis was blocked after cell fusion, some of the human hnRNP A1 protein must have left the HeLa-cell nucleus, moved through the cytoplasm, and entered the Xenopus nucleus in the heterokaryon.

[Reprinted by permission from Nature Publishing Group, from R. S. Piñol and G. Dreyfuss, 1992, “Shuttling of Pre-mRNA Binding Proteins Between Nucleus and Cytoplasm,” *Nature* **355**(6362):730–732; permission conveyed through the Copyright Clearance Center, Inc.]

RNA-Binding Domains

The *RNA recognition motif (RRM)* is the most common RNA-binding domain in hnRNP proteins. This ~80-residue domain, which occurs in many other RNA-binding proteins, contains two highly conserved sequences (RNP1 and RNP2) that are found across organisms ranging from yeast to human — indicating that like many DNA-binding domains, they arose early in eukaryotic evolution.

Structural analyses have shown that the RRM domain consists of a four-stranded β sheet flanked on one side by two α helices (Figure 9-6a). To interact with the negatively charged RNA phosphates, the β sheet forms a positively charged surface. The conserved RNP1 and RNP2 sequences lie side by side on the two central β strands, and their side chains make multiple contacts with a single-stranded region of RNA that lies across the surface of the β sheet. This is seen in the x-ray crystal structure of the two RRM domains of *Drosophila* Sex-lethal protein (Sxl) bound to its specific site in RNA (Figure 9-6b), and the two RRM domains of human polypyrimidine tract-binding protein (PTB) bound to its specific RNA (Figure 9-6c). Sxl and PTB are examples of the strikingly different orientations of RRM domains in different hnRNPs.

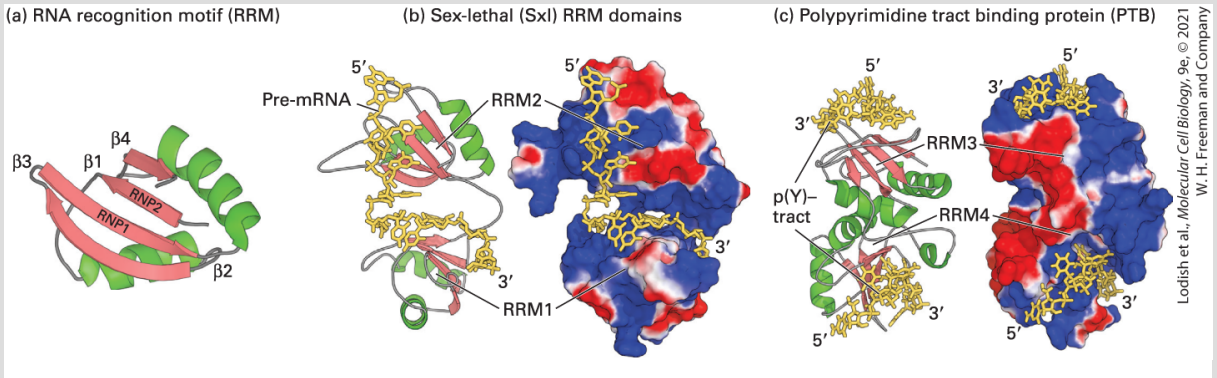


FIGURE 9-6 Structure of the RRM domain and its interaction with RNA. (a) Diagram of the RRM domain showing the two α helices (green) and four β strands (red) that characterize this motif. The conserved RNP1 and RNP2 regions are located in the two central β strands. (b) Surface representation of the two RRM domains in *Drosophila* Sex-lethal (Sxl) protein, which together bind a continuous nine-base sequence in transformer pre-mRNA (yellow). In both (b) and (c), positively charged regions are shown in shades of blue; negatively charged regions, in shades of red; RNA is yellow. The two RRMs in Sxl are oriented like the two parts of an open pair of castanets, with the β sheets of the RRMs facing toward each other. The pre-mRNA is bound to the surfaces of the positively charged β sheets, making most of its contacts with the RNP1 and RNP2 regions of each RRM. (c) Strikingly different orientation of RRM domains in the polypyrimidine tract-binding protein (PTB), illustrating that RRMs are oriented in different relative positions in different hnRNPs. p(Y)-tract is a polypyrimidine tract. In PTB, the two RRMs associate through their α helices, so that the positively charged β sheets face away from each other, upward for RRM3 and downward for RRM4. The structure of CUCUCU single-stranded RNA bound to each of the two RRMs was determined, explaining how PTB can bind to two tracts of six pyrimidines in a single RNA if they are separated by a loop of ~ 15 or more nucleotides. This ability of PTB to form a small loop in a pre-mRNA probably contributes to its ability to function as a splicing repressor at exons where the upstream $3'$ splice site or the downstream $5'$ splice site is flanked by two polypyrimidine tracts. Part (a) See K. Nagai et al., 1995, *Trends Biochem. Sci.* **20**:235.

[Part (b) Data from N. Harada et al., 1999, *Nature* **398**:579. Part (c) Data from F. C. Oberstrass et al., 2006, *Science* **309**:2054.]

The 45-residue *KH domain* is found in the hnRNP K protein and several other RNA-binding proteins. The three-dimensional

structure of representative KH domains is similar to that of the RRM domain but smaller, consisting of a three-stranded β sheet supported from one side by two α helices. Nonetheless, the KH domain interacts with RNA much differently than does the RRM domain. RNA binds to the KH domain by interacting with a hydrophobic surface formed by the α helices and one β strand. The *RGG box,* another RNA-binding motif found in hnRNP proteins, contains five Arg-Gly-Gly (RGG) repeats with several interspersed aromatic amino acids. Its arginine-rich nature is similar to the RNA-binding domains of the HIV Tat protein. KH domains and RGG repeats are often interspersed in two or more sets in a single RNA-binding protein.

Splicing Occurs at Short, Conserved Sequences in Pre-mRNAs via Two Transesterification Reactions

During formation of a mature, functional mRNA, the **introns** are removed and **exons** are spliced together. For short transcription units, RNA splicing usually follows cleavage and polyadenylation of the 3′ end of the primary transcript, as depicted for β-globin in Figure 9-2. However, for long transcription units containing multiple exons, splicing of exons in the nascent RNA begins before transcription of the gene is complete.

Early pioneering research on the nuclear processing of mRNAs revealed that mRNAs are initially transcribed as much longer RNA molecules than the mature mRNAs in the cytoplasm. It was also shown that RNA sequences near the 5′ cap added shortly after transcription initiation are retained in the mature mRNA and that RNA sequences near the polyadenylated 3′ end of mRNA processing intermediates in hnRNA also are retained in the mature mRNA in the cytoplasm. How could an RNA be shorter than the primary transcript, but have the same 5′ and 3′ ends? The solution to this apparent conundrum came from the discovery of introns by electron microscopy of RNA-DNA hybrids of adenovirus DNA and the mRNA-encoding hexon, a major virion capsid protein (Figure 9-7). Other studies revealed nuclear viral RNAs that were colinear with the viral DNA (primary transcripts) and RNAs with one or two of the introns removed (processing intermediates). These results, together with the earlier findings that the 5′ cap and 3′ poly(A) tail at each end of long mRNA precursors are retained in shorter mature cytoplasmic mRNAs, led to the realization that introns are removed from primary transcripts as exons are spliced together.

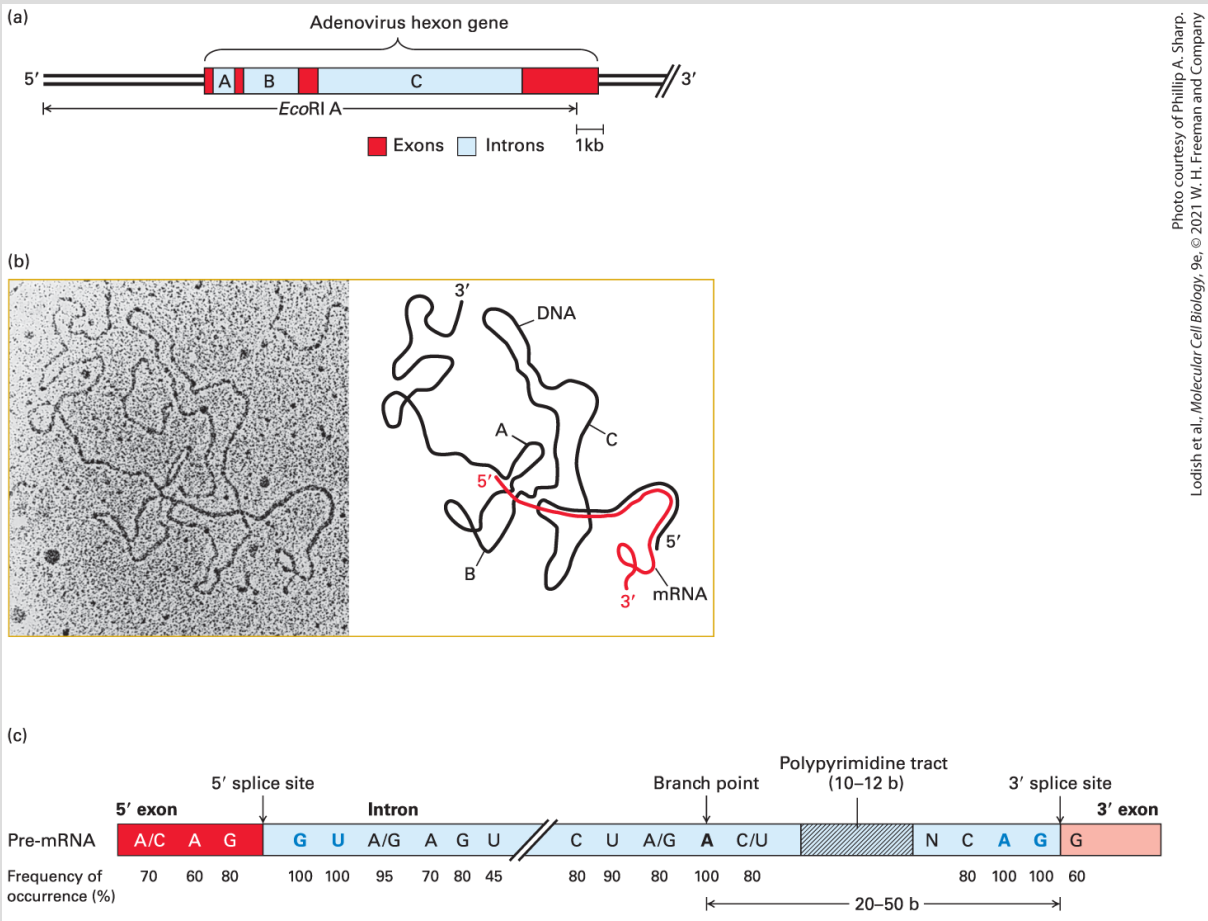


Photo courtesy of Phillip A. Sharp.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 9-7 Electron microscopy of mRNA-template DNA hybrids shows that introns are spliced out during pre-mRNA processing. (a) Diagram of the EcoRI A fragment of adenovirus DNA, which extends from the left end of the genome to just before the end of the final exon of the hexon gene. The gene consists of three short exons and one long (~ 3.5 kb) exon separated by three introns of ~ 1, 2.5, and 9 kb. (b) Electron micrograph *(left)* and schematic drawing *(right)* of a hybrid between an EcoRI A DNA fragment and a hexon mRNA. The loops marked A, B, and C correspond to the introns indicated in (a). Since these intron sequences in the viral genomic DNA are not present in the mature hexon mRNA, they loop out between the exon sequences that hybridize to their complementary sequences in the mRNA. (c) Consensus sequences around splice sites in vertebrate pre-mRNAs. The only nearly invariant bases are the $5'$ GU and the $3'$ AG of the intron (blue), although the flanking bases indicated are found at frequencies higher than expected based on a random distribution. A pyrimidine-rich region (hatch marked) near the $3'$ end of the intron is found in most cases. The branchpoint A, also invariant, usually is 20–50 bases from the $3'$ splice site. The central region of the intron, which may range from 40 bases to 50 kilobases in length, generally is unnecessary for splicing to occur. See R. A. Padgett et al.,

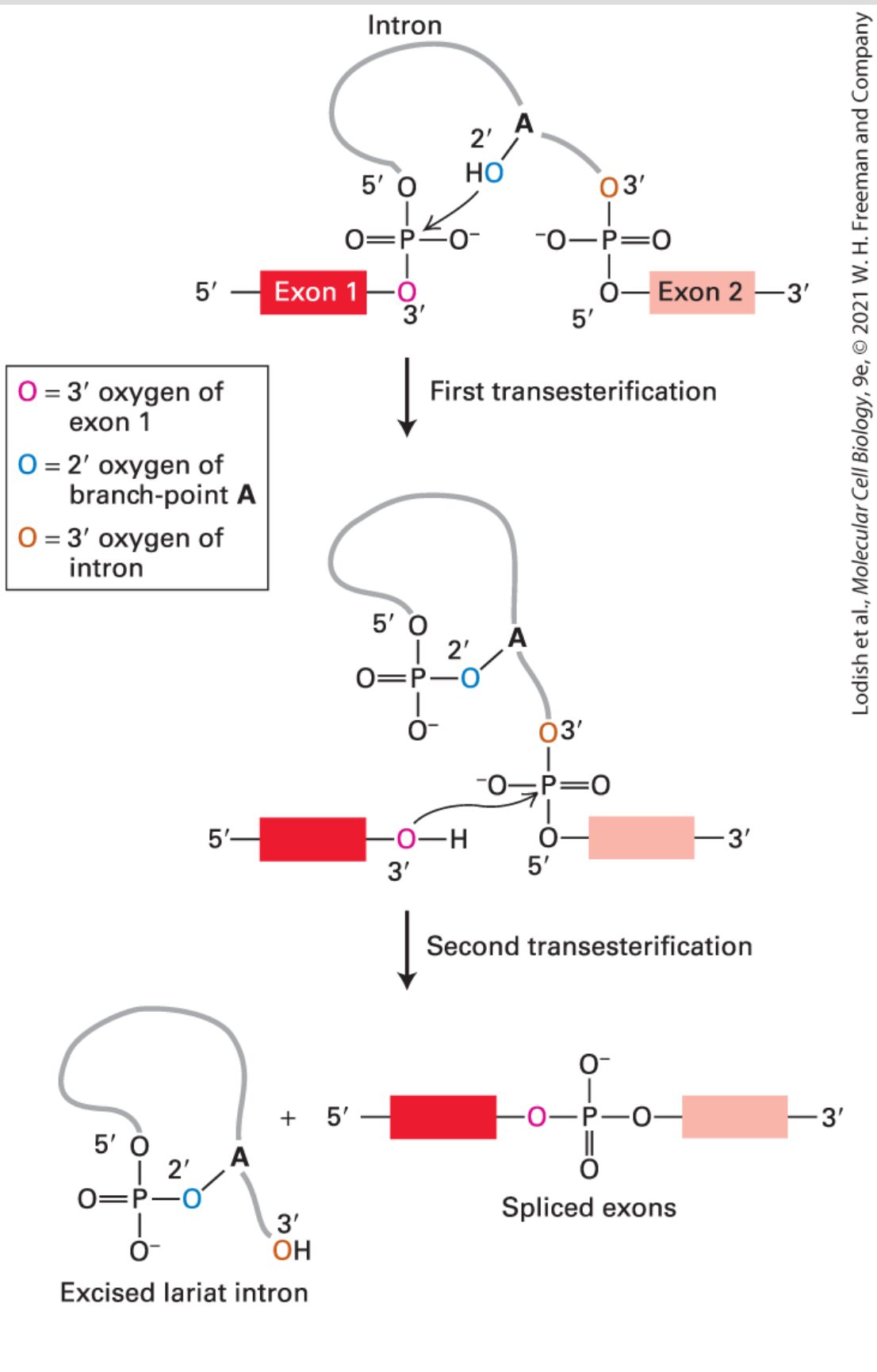
1986, *Ann. Rev. Biochem.* **55**:1119, and E. B. Keller and W. A. Noon, 1984, *Proc. Nat'l. Acad. Sci. USA* **81**:7417.

[Micrograph in (b) Data from S. M. Berget et al., 1977, *Proc. Nat'l. Acad. Sci. USA* **74**:3171;]

The location of *splice sites* — that is, exon-intron junctions — in a pre-mRNA can be determined by comparing the sequence of genomic DNA with that of the cDNA prepared from the corresponding mRNA (see Figure 6-17). Sequences that are present in the genomic DNA but absent from the cDNA represent introns and indicate the positions of splice sites. Such analysis of a large number of different mRNAs revealed moderately conserved, short consensus sequences at the splice sites flanking introns in vertebrate pre-mRNAs; a pyrimidine-rich region just upstream of the 3′ splice site is also common (Figure 9-7c). Studies of mutant genes with deletions introduced into introns showed that much of the central portion of introns can be removed without affecting splicing; generally, only 30–40 nucleotides at each end of an intron are necessary for splicing to occur at normal rates.

Analysis of the intermediates formed during splicing of pre-mRNAs in vitro led to the discovery that splicing of exons proceeds via two sequential *transesterification reactions* (Figure 9-8). Introns are removed as a lariat structure in which the 5′ G of the intron is joined in an unusual 2′,5′-phosphodiester bond to an adenosine near the 3′ end of the intron. This adenosine (A) residue is called the *branch-point A* because it forms an RNA branch in the lariat structure. In

each transesterification reaction, one phosphoester bond is exchanged for another. Since the number of phosphoester bonds in the molecule is not changed in either transesterification reaction, no energy is consumed. The net result of these two reactions is that two exons are ligated and the intervening intron is released as a branched lariat structure.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

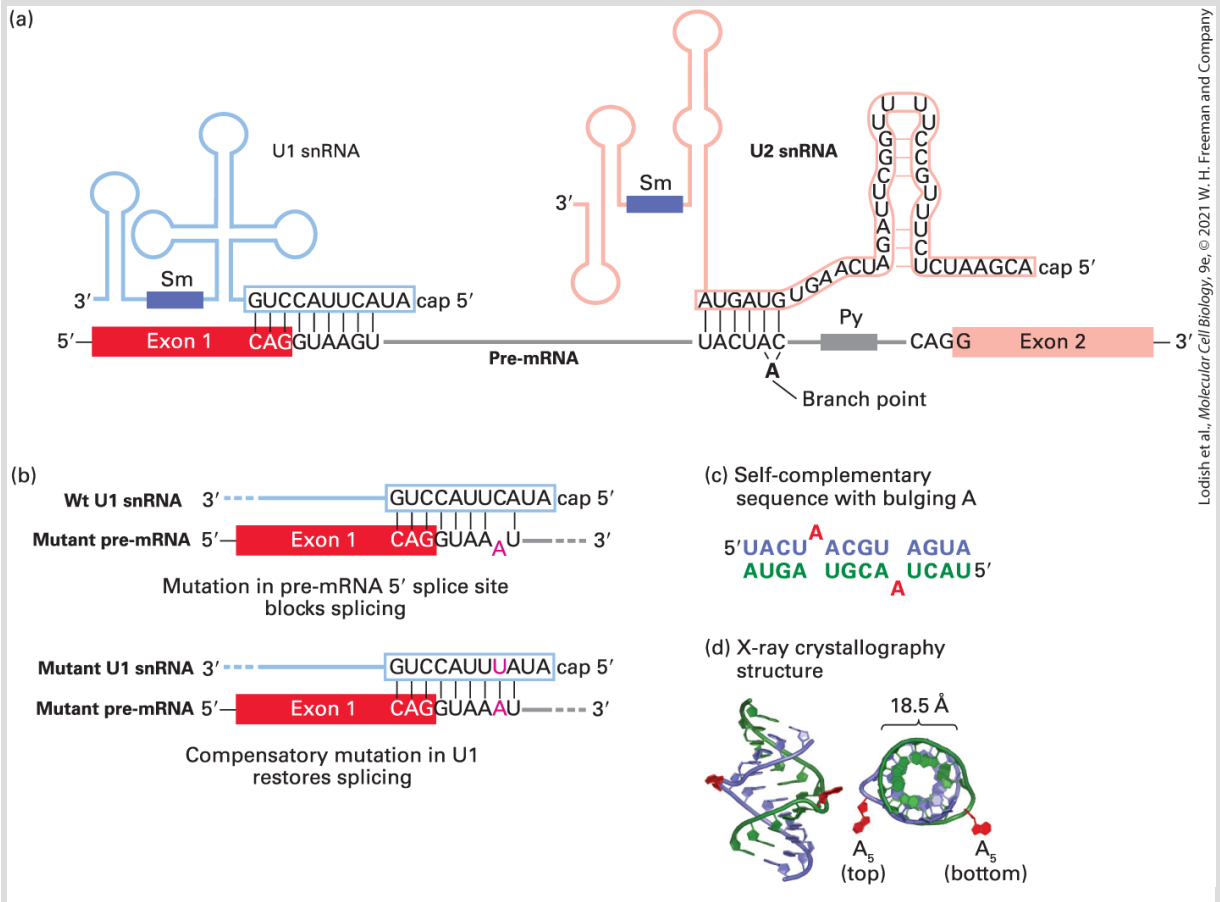
FIGURE 9-8 Two transesterification reactions that result in splicing of exons in pre-mRNA. In the first reaction, the ester bond between the 5′ phosphorus of the intron and the 3′ oxygen (dark red) of exon 1 is exchanged for an ester bond with the 2′ oxygen (blue) of the branch-site A residue. In the second reaction, the ester bond between the 5′ phosphorus of exon 2 and the 3′ oxygen (orange) of the intron is exchanged for an ester bond with the 3′ oxygen of exon 1, releasing the intron as a lariat structure and joining the two exons. Arrows show where activated hydroxyl oxygens react with phosphorus atoms.

During Splicing, snRNAs Base-Pair with Pre-mRNA to Select Splice Sites and Guide the Transesterification Reactions

Splicing requires highly abundant **small nuclear RNAs (snRNAs)**, important for base pairing and other interactions with the pre-mRNA, as well as ∼150 splicing factor proteins. Five U-rich snRNAs restricted to the nucleus, designated *U1, U2, U4, U5*, and *U6*, participate in pre-mRNA splicing. Ranging in length from 107–210 nucleotides in vertebrates, these snRNAs are associated with 5–12 proteins each in the abundant small nuclear ribonucleoprotein particles (snRNPs) in the nuclei of eukaryotic cells (see Figure 9-10).

Definitive evidence for the function of U1 snRNA in splicing came from experiments indicating that base pairing between the 5′ splice site of a pre-mRNA and the 5′ region of U1 snRNA (Figure 9-9a) is

required for RNA splicing (Figure 9-9b). In vitro experiments showed that a synthetic oligonucleotide that hybridizes with the 5′-end region of U1 snRNA blocks RNA splicing. In vivo experiments showed that base pairing–disrupting mutations in the pre-mRNA 5′ splice site (red A in Figure 9-9b, *left*) also block RNA splicing; in this case, however, splicing can be restored by expression of a mutant U1 snRNA with a *compensating mutation* (red U in Figure 9-9b) that restores base pairing to the mutant pre-mRNA 5′ splice site. These results and similar experiments with other mutations in pre-mRNA bases near a 5′ splice site and compensating mutations in U1 snRNA that restore base pairing to the mutant pre-mRNA, showed that, except for the invariant GU at the 5′ end of nearly all introns, base pairing with the U1 5′ region, rather than a specific RNA sequence, is required for RNA splicing.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 9-9 Base pairing between pre-mRNA, U1 snRNA, and U2 snRNA early in the splicing process. (a) In this diagram, secondary structures in the snRNAs that are not altered during splicing are depicted schematically. The rectangles labeled Sm are binding sites for a ring of seven Sm proteins that encircle the snRNA at the positions shown in U1 and U2 snRNAs. The same Sm proteins also form a ring around U4 and U5 snRNAs (not shown). The invariant yeast branch-point sequence UACUAAC is shown here, with the bulged branch-point A indicated. Note that U2 snRNA base-pairs with a sequence that includes the branch-point A, although the branch-point A is not base-paired. (b) Only the 5′ ends of U1 snRNAs and 5′ splice sites in pre-mRNAs are shown. *Left:* A mutation (red A) in a pre-mRNA splice site that interferes with base pairing to the 5′ end of U1 snRNA blocks splicing. *Right:* Expression of a U1 snRNA with a compensating mutation (red U) that restores base pairing also restores splicing of the mutant pre-mRNA. (c, d) Structure of a bulged A in an RNA-RNA helix, as in the pre-mRNA branch-point A base paired to U2 snRNA in (a). The structure of an RNA duplex with the sequence shown (c), containing bulged A residues (red) at position 5 in the RNA helix, was determined by x-ray crystallography (d). The bulged A residues extend from the side of an A-form RNA-RNA helix. The phosphate backbone of one strand is shown in green;

the other strand in blue. The structure on the right is turned 90° for a view down the axis of the helix. See M. J. Moore et al., 1993, in R. Gesteland and J. Atkins, eds., *The RNA World*, Cold Spring Harbor Press, pp. 303–357. See also Y. Zhuang and A. M. Weiner, 1986, *Cell* **46**:827.

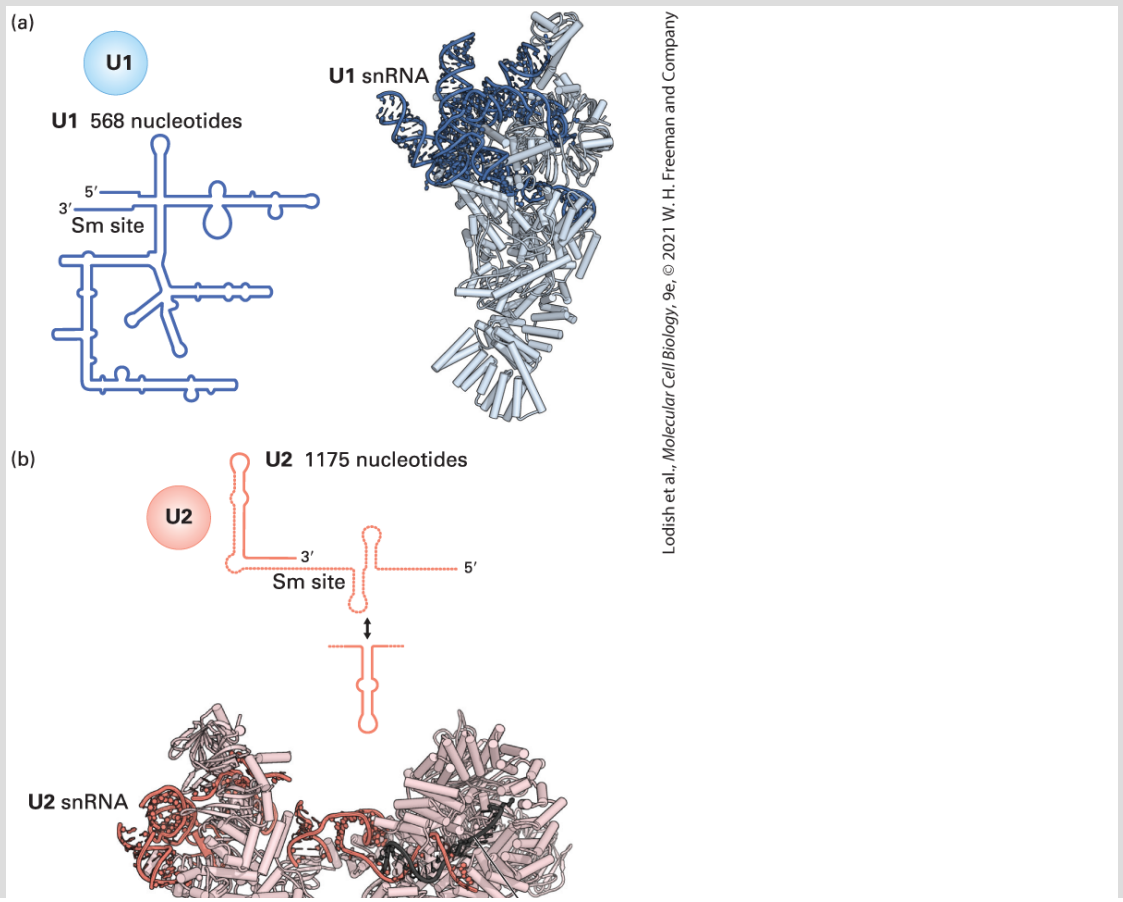
[Parts (c) and (d) Data from J. A. Berglund et al., 2001, *RNA* **7**:682.]

Involvement of U2 snRNA in splicing was initially suspected when it similarly was found to have an internal sequence that is largely complementary to the consensus sequence flanking the branch-point A in pre-mRNAs (see Figure 9-7c). Compensating mutation experiments, similar to those conducted with U1 snRNA and 5′ splice sites, demonstrated that base pairing between U2 snRNA and the branch-point sequence in pre-mRNA (see Figure 9-7c) is also critical to splicing.

Significantly, the branch-point A itself, which is not base-paired to U2 snRNA, bulges out (Figure 9-9a), allowing its 2′ hydroxyl to participate in the first transesterification reaction of RNA splicing (see Figure 9-8). The structure of a model RNA-RNA duplex with two bulged As (Figure 9-9c) was determined by x-ray crystallography, showing the structure of a bulged-out A similar to the bulged-out A generated when U2 snRNA base pairs with the branch-point region of pre-mRNA (Figure 9-9d).

Spliceosomes Catalyze Pre-mRNA Splicing

Much of our understanding of the mechanisms regulating pre-mRNA splicing has come from biochemical studies with extracts of cultured mammalian cells and yeast and from genetic studies in yeast. The two transesterification reactions of RNA splicing are catalyzed by a 1.3 megadalton RNA-protein complex called the **spliceosome**, similar in size to the small ribosomal subunit. The spliceosome is assembled using five smaller ribonucleoprotein complexes called **snRNPs (small ribonuclear particles) U1, U2, U4, U5, and U6** assembled on the snRNAs described above with the same names (Figure 9-10). (U3 RNA, discussed later, is not involved in pre-mRNA splicing; rather, it functions in pre-ribosomal RNA processing; see Figure 9-46.)



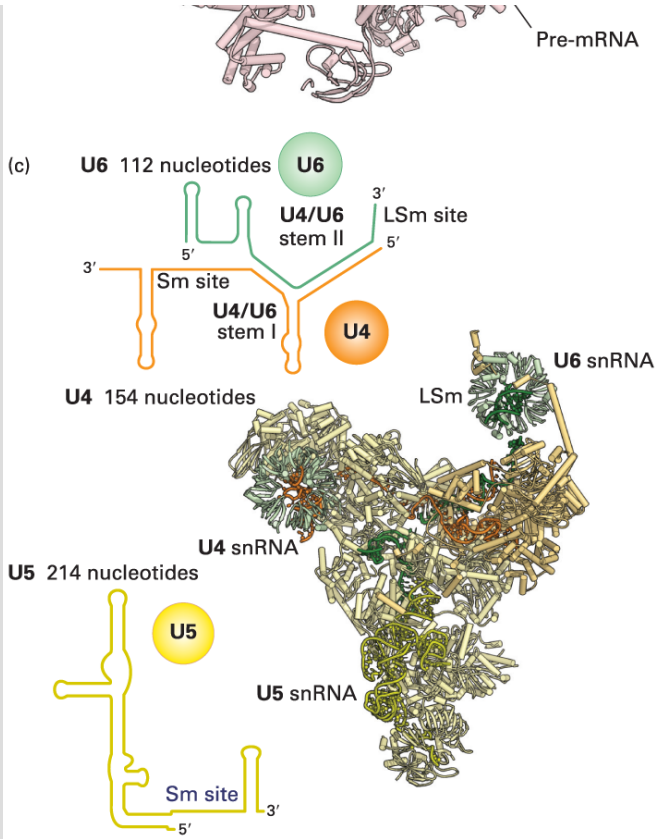


FIGURE 9-10 Structures of *Saccharomyces cerevisiae* snRNPs involved in pre-mRNA splicing. The secondary structures of the snRNAs are diagrammed at the left, and the three-dimensional structures determined by cryo-EM are shown at the right, with the locations of associated proteins indicated by ribbon diagrams. Each of the five small nuclear ribonucleoprotein particles (snRNPs) is shown with shades of distinct colors (U1, blue; U2, pink; U4, orange; U5, yellow; U6, green). "Sm site" in the RNA secondary structure diagrams of the U1, U2, U4, and U5 snRNAs are sites where a ring of seven Sm proteins encircle the snRNA. An alternative seven-protein ring composed of proteins LSM2-8 forms around the 3′ end of the U6 snRNA.

[Data from C. Plaschka, A. J. Newman, and K. Nagai, 2019, *Cold Spring Harb. Perspect. Biol*. **11**:a032391.]

The U1, U2, U4, U5, and U6 snRNPs associate *transiently* with each other and with splice sites in pre-mRNAs to assemble an active spliceosome that includes only the U2, U5, and U6 snRNPs (Figure 9-

11a). The active spliceosome also includes two large multiprotein complexes called *NTR* and *NTC* (Figure 9-11a). As can be seen from Figure 9-11b highlighting the snRNAs, the active spliceosome is composed primarily of protein. The fully assembled, active spliceosome performs the two-step splicing reaction and then dissociates. Studies of yeast temperature-sensitive mutants identified ~150 essential protein-coding genes required for RNA splicing, most of which have clear orthologs in humans. Most of these encode polypeptides that are transiently bound and then dissociate from spliceosome assembly intermediates, as well as factors released in association with the excised lariat intron. The released snRNPs and splicing-protein factors then reassemble onto another pre-mRNA splice site to catalyze another two-transesterification splicing reaction.

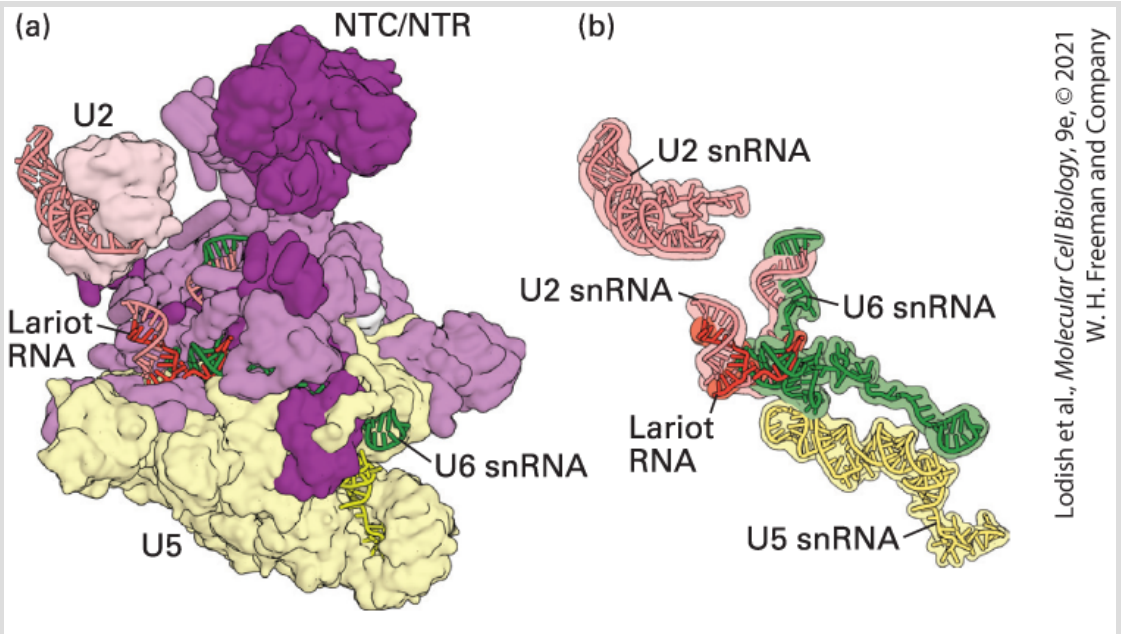


FIGURE 9-11 Structure of the catalytically active yeast spliceosome. (a) U5 snRNP (yellow) forms a scaffold onto which the other three major subcomplexes of the

catalytically active spliceosome (U2 snRNP (pink) and protein complexes NTR (light purple) and NTC (dark purple) dock around the RNA molecules and the catalytic center. The four RNA molecules visible in this space-filling model are colored orange for U2, yellow for U5, green for U6, and red for the intron lariot RNA. (b) Positions of U6, U5, two portions of U2 snRNAs, and the lariat intron RNA (red) in the catalytically active spliceosome, determined at high resolution by cryo-EM.

[Republished with permission from the American Association for the Advancement of Science, from C. Yan et al., "Structure of a Yeast Spliceosome at 3.6-Angstrom Resolution." *Science* **349** (6253):1182–1191; permission conveyed through Copyright Clearance Center, Inc.]

Unlike the ribosome during protein synthesis, the spliceosome is not a stable complex. During assembly of the active spliceosome, some snRNPs and additional splicing factor proteins initially associate with assembly intermediates and then are released. As discussed above, the RNA-splicing reaction results in no change in free energy since there is no change in the number of phosphodiester bonds before and after the reaction. However, during RNA splicing in the cell, the process is driven toward ligation of the 5′ and 3′ exons by RNA-helicase enzymes (proteins), utilizing energy from ATP hydrolysis, as discussed below.

In yeast, the 5′ splice site and branch-point sequences are stringently conserved for almost all introns (Figure 9-12). In contrast, splice sites and branch-point sequences are highly degenerate in mammals, except for the GU and AG at the 5′ and 3′ ends of most introns and the defining A at the branch point (Figure 9-12). This is because compared to yeast, mammals utilize multiple

additional sequence-specific RNA-binding proteins that assist and regulate snRNP U1 and U2 binding to 5′ splice sites and the branch point, respectively. This allows alternative splicing of primary transcripts in mammals by regulating the activities of these RNA-binding proteins. Comparison of the more simplified splicing system in yeast with the more complex system in mammals facilitated studies in both systems and an understanding of how additional nuclear proteins in mammals participate in and regulate alternative RNA splicing.

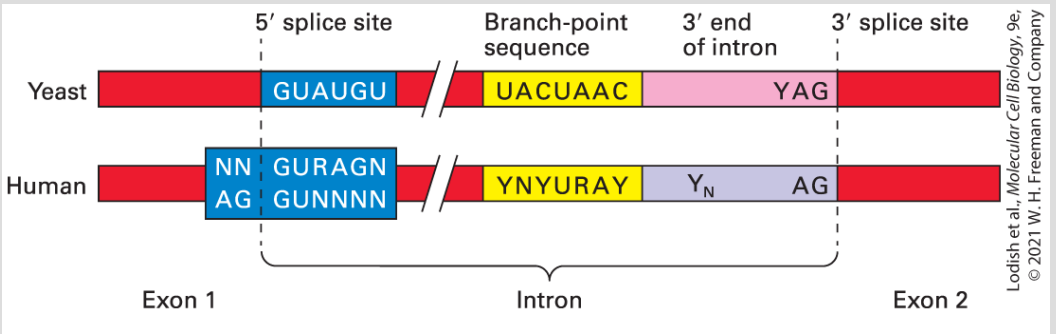


FIGURE 9-12 RNA sequences at the splice sites and branch-point A in yeast and humans. Yeast exhibit more conserved splice site and branch-point sequences than mammalian cells [note the high number of N's (any base) in the mammalian sequence].

Pre-mRNA Splicing Cycle

In yeast, 10 complexes of snRNPs and protein splicing factors have been distinguished during a splicing cycle (Figure 9-13). As discussed above (Figure 9-9a), spliceosome assembly initiates through base pairing interactions between the U1 snRNA in a U1

snRNP and a 5′ splice site in a pre-mRNA (Figure 9-9a, *left*). Next, in mammalian cells, a heterodimeric splicing factor protein called *U2 associated factor (U2AF)*, together with another protein factor, splicing factor 1 (SF1), promote association of a U2 snRNP with a branch-point sequence near the 3′ end of the intron (Figure 9-9a, *right*). The small U2AF subunit binds the 3′ splice site AG, and the large U2AF subunit binds the polypyrimidine tract in metazoans. SF1 interacts transiently with the branch-point A, and then is released as U2 snRNA base pairs to the branch-point region (Figure 9-9a). Together, these interactions specify the location of a 3′ splice site used in a splicing reaction. Other protein splicing factors not shown associate with and stabilize these U1 and U2 snRNA interactions with the pre-mRNA.

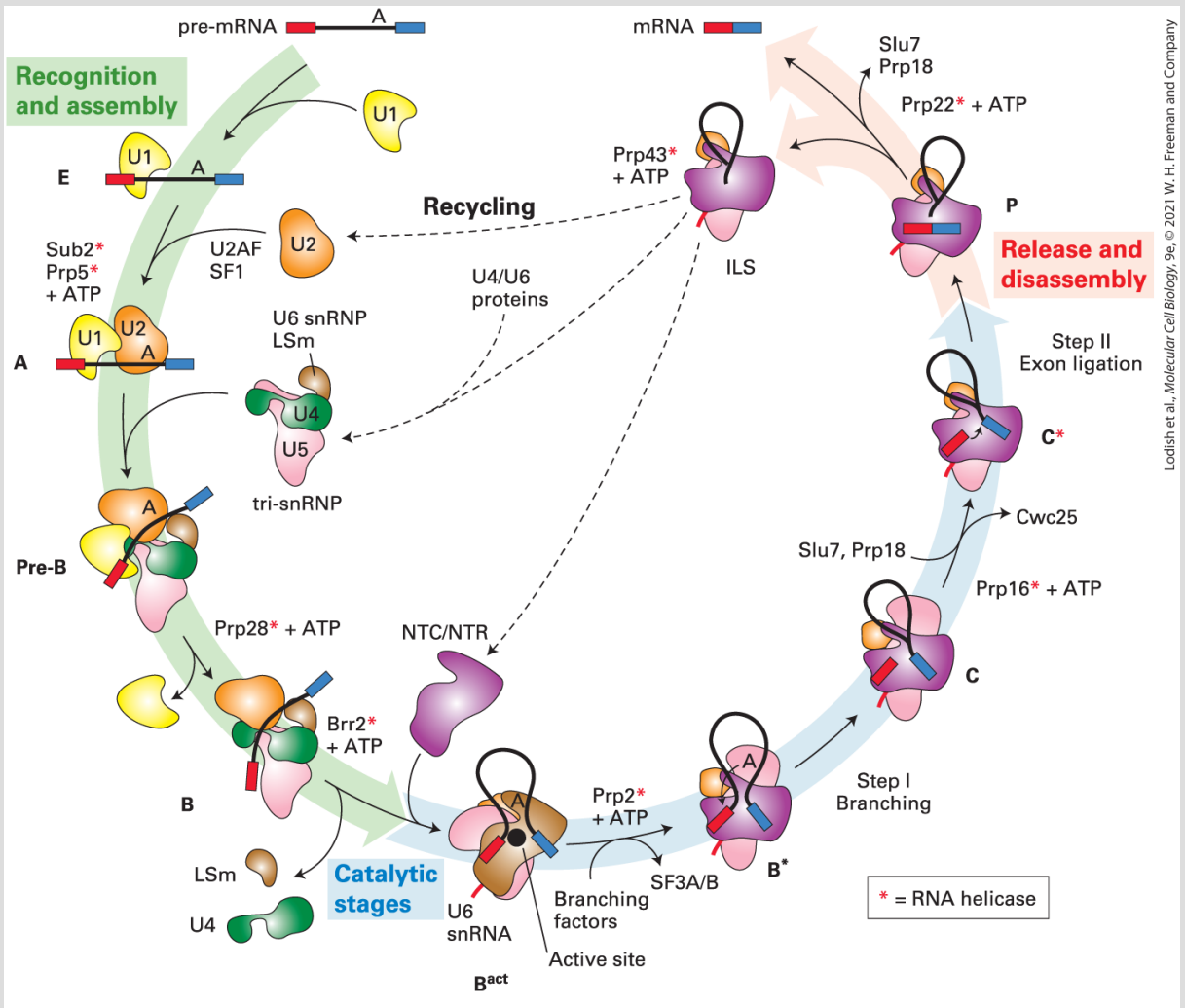


FIGURE 9-13 Ten complexes of pre-mRNAs with snRNPs and protein splicing factors have been distinguished during the splicing cycle in yeast. The names of each complex are indicated around the circumference of the diagram. See text for explanation.

[Data from C. Plaschka, A. J. Newman, and K. Nagai, 2019, *Cold Spring Harb. Perspect. Biol.* **11**:a032391.]

U2 snRNP association is followed by large conformational rearrangements of the snRNPs, including association of a pre-assembled *tri-snRNP complex* of U4, U5, and U6 snRNPs (Figure 9-10) during the transition from the A to pre-B complex (Figure 9-13). This

is followed by dissociation of the U1 snRNP to form the B-complex. The U4 snRNP then dissociates as the large multisubunit protein complexes NTC (composed of seven polypeptides) and NTR (composed of six polypeptides) associate, generating the B^{act} complex. The snRNP rearrangements that occur during the transition from the B complex to the B^{act} complex generate new interactions between U6 and U2 and between U6 and the 5′ splice site (Figure 9-14a) that are required for catalysis of the two splicing-transesterification reactions.

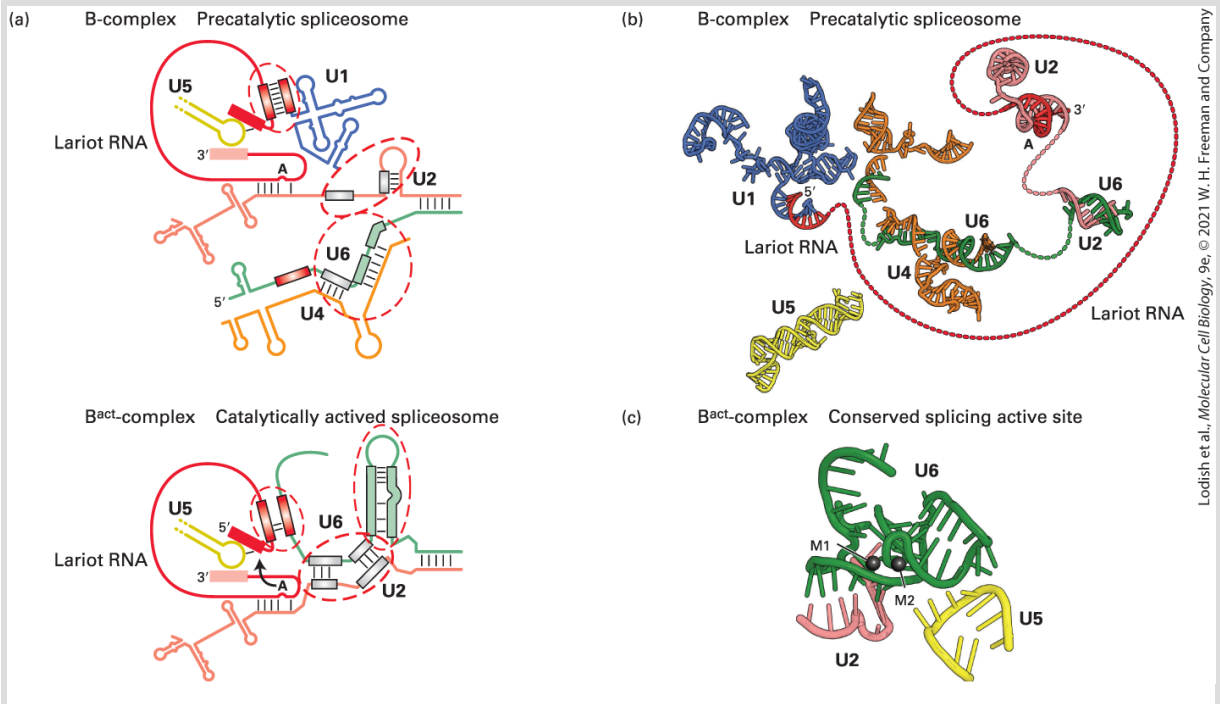


FIGURE 9-14 Rearrangements of snRNPs generate an active catalytic site. (a) The network of RNA interactions in the precatalytic B-complex (*top*) and the catalytically active B^{act} complex (*bottom*). During activation, base pairing interactions in the B complex indicated by dashed red ovals including regions in U4, U6 and U2 undergo major rearrangements. U1 and U4 snRNPs dissociate from the spliceosome, allowing the 5′ end of the U6 snRNA to base pair through its highly conserved ACAGAG motif to the 5′ splice site (dashed red oval). Other critical base-pairing interactions in the Bact complex are indicated (dashed red ovals). (b)

Structures of the RNAs in the precatalytic B-complex visualized at high resolution by cryo-EM. (c) RNA elements at the splicing active site of the Bact complex. Mg^{+2} ions are held by U6 snRNA in the correct geometry relative to the $5'$ and $3'$ splice sites and the branchpoint A and elements of U2 and U5 snRNAs to catalyze the two splicing transesterification reactions.

[Part (a) Data from M. C. Wahl, C. L. Will, and R. Luhrmann, 2009, *Cell* **136**:701–718. Part (b) Data from C. Plaschka, A. J. Newman, and K. Nagai, 2019, *Cold Spring Harb. Perspect. Biol.* **11**:a032391. Part (c) Data from C. Yan, R. Wan, and Y. Shi, 2019, *Cold Spring Harb. Perspect. Biol.* **11**:a032409.]

These rearrangements of the snRNPs and protein splicing factors generate a catalytic active site (Figure 9-14b) where the first transesterification is catalyzed, followed by the second transesterification joining the $5'$ and $3'$ exons. The two reactions take place without significant rearrangement of the active site other than the positions of the phosphoester bonds at the splice sites (distinguishing complexes B*, C, and C*; Figure 9-13). Unlike the ribosome, in which RNA accounts for $\sim$50 percent of the molecular mass, the three snRNAs (U2, U5, and U6) of the catalytically active human spliceosome represent no more than 5 percent of its total molecular mass (Figure 9-11). This seems consistent with the important functions of proteins at many steps in the splicing process. Spliceosomal protein components organize the splicing active site, deliver the reactive chemical groups into the active site, drive the reaction toward exon joining, and regulate the splicing process.

The splicing process is completed by disassembly of the spliceosome, which requires the NTC and NTR protein complexes

(Figure 9-13, complex P → ILS). This results in release of the product spliced RNA with the 5′ exon ligated to the 3′ exon by a standard 5′ to 3′ phosphodiester bond, and release of the intron lariat RNA associated with U2, U5, and U6 snRNPs, and the NTC and NTR protein complexes (labeled *ILS* in Figure 9-13 for *i*ntron *l*ariat *s*pliceosome). Degradation of the intron lariat by a debranching enzyme that cleaves the 2′–5′ phosphodiester bond with the branch-site A and by 5′ and 3′ exonucleases releases the U2, U5, and U6 snRNPs and splicing factors for use in another cycle of splicing (Figure 9-13, Recycling).

Several RNA helicases are required in the splicing process (indicated by * in Figure 9-13). For example, the RNA helicase Brr2 is required for the critical transition from the catalytically inactive B complex that contains all the snRNPs required for splicing (but in inactive conformations), to the catalytically active B^{act} complex. Many RNA helicases are proposed to break RNA-RNA base pairs by binding single-stranded RNA together with ATP next to an RNA-RNA duplex. ATP hydrolysis is then coupled to a change in the helicase conformation that pulls the single-stranded RNA through the active site of the helicase in a 3′ direction. This dislodges a complementary single strand or a protein bound to the RNA. Brr2 is postulated to bind the single-stranded U4 snRNA in the U4/U6/U5 tri-snRNP next to the U4/U6 duplex (Figure 9-10, U4/U6 stem 1) and then pull on the RNA in the 3′ direction, causing unwinding of the U4/U6 duplex and dissociation of the associated proteins. This allows U4 to be released

from the complex and U6 to change conformation in order to make interactions with the 5′ splice site and to bind two Mg^{2+} ions in the proper geometry relative to the splice sites and branch-point A to catalyze the exchange of phosphodiester bonds (Figure 9-14c). RNA is central to the catalytic mechanism. The transesterification reactions are catalyzed by the two Mg^{2+} ions bound by snRNA U6. The critical branch-site A 2′-OH is exposed through RNA-RNA interactions. However, proteins stabilize and guide these RNA-RNA interactions, and RNA helicases use energy from ATP hydrolysis to push the reaction cycle in favor of exon joining. Consequently, the spliceosome is considered to be a *protein-directed metalloribozyme*.

Exon Junction Complexes in Vertebrates

Following RNA splicing in vertebrates, a specific set of hnRNP proteins remain bound to the spliced RNA approximately 20 nucleotides 5′ to each exon-exon junction, thus forming an *exon-junction complex.* One of the hnRNP proteins associated with the exon-junction complex is the *R*NA *e*xport *f*actor (REF), which functions in the export of fully processed mRNPs from the nucleus to the cytoplasm, as will be discussed in Section 9.3. Other proteins associated with the exon-junction complex function in a quality-control mechanism that leads to the degradation of improperly spliced mRNAs, known as *nonsense-mediated decay* (Section 9.4).

Rare 5′-AU … AC-3′ Introns

A small fraction of pre-mRNAs (∼1 percent in humans) contain introns whose splice sites do not conform to the standard consensus sequence. This class of introns begins with AU and ends with AC rather than following the usual "GU-AG rule" (see Figure 9-7c). Splicing of this special class of introns occurs via a splicing cycle analogous to that shown in Figure 9-13, except that four novel, low-abundance snRNPs, together with the standard U5 snRNP, are involved.

Trans-Splicing

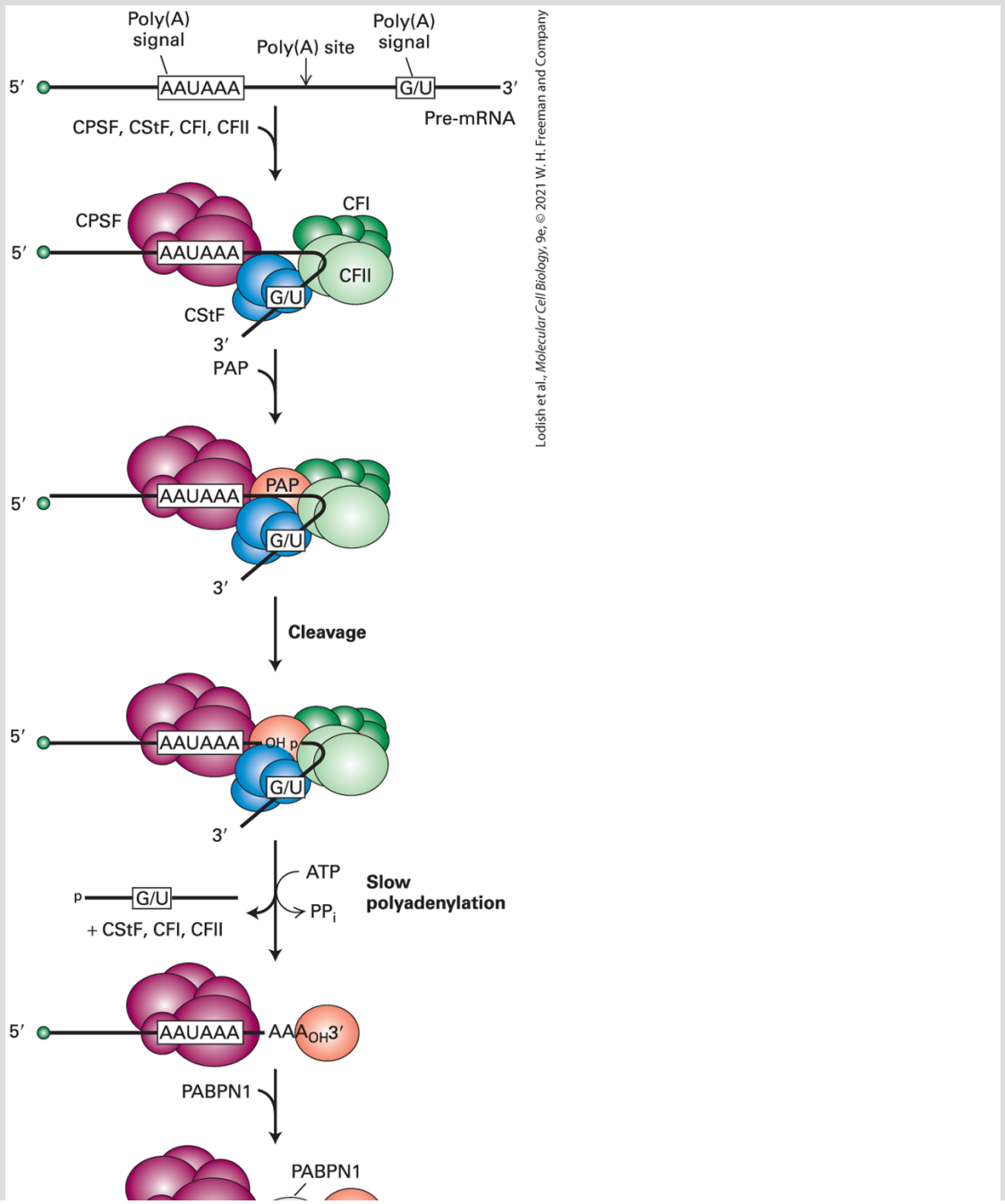
Nearly all functional mRNAs in vertebrate, insect, and plant cells are derived from a single molecule of the corresponding pre-mRNA by removal of internal introns and splicing of exons. However, in two types of protozoans — trypanosomes and euglenoids — mRNAs are constructed by splicing together separate RNA molecules. This process, referred to as *trans-splicing,* is also used in the synthesis of 10–15 percent of the mRNAs in the nematode (roundworm) *Caenorhabditis elegans,* an important model organism for studying embryonic development. Trans-splicing is carried out by snRNPs and splicing factors by a process similar to the splicing of exons in a single pre-mRNA (Figure 9-8), except that there is a break in the loop representing the intron.

3′ Cleavage and Polyadenylation of Pre-mRNAs Are Tightly Coupled

In eukaryotic cells, all mRNAs, except histone mRNAs, have a 3′ **poly(A) tail**. Early studies of pulse-labeled adenovirus and SV40 RNA demonstrated that the viral primary transcripts extend beyond the sequence that is polyadenylated. These results suggested that A residues are added to a 3′ hydroxyl generated by endonucleolytic cleavage of a longer transcript. But the predicted downstream RNA fragments were never detected in vivo, presumably because of their rapid degradation. However, detection of both predicted cleavage products was observed in the in vitro processing reactions performed with nuclear extracts of cultured human cells. The cleavage/polyadenylation process and degradation of the RNA downstream of the cleavage site occurs much more slowly in these in vitro reactions, simplifying detection of the downstream cleavage product.

Early sequencing of cDNA clones from animal cells showed that nearly all mRNAs contain the sequence AAUAAA 10–35 nucleotides *upstream* from the poly(A) tail (Figure 9-15). Polyadenylation of RNA transcripts is virtually eliminated when the corresponding sequence in the template DNA is mutated to any other sequence, except the closely related sequence (AUUAAA). The unprocessed RNA transcripts produced from such mutant templates do not accumulate in nuclei but are rapidly degraded. Further mutagenesis

studies revealed that a second signal downstream from the cleavage site is required for efficient cleavage and polyadenylation of most pre-mRNAs in animal cells. This downstream signal is not a specific sequence but rather a GU-rich or simply a U-rich region within ~20 nucleotides of the cleavage site.



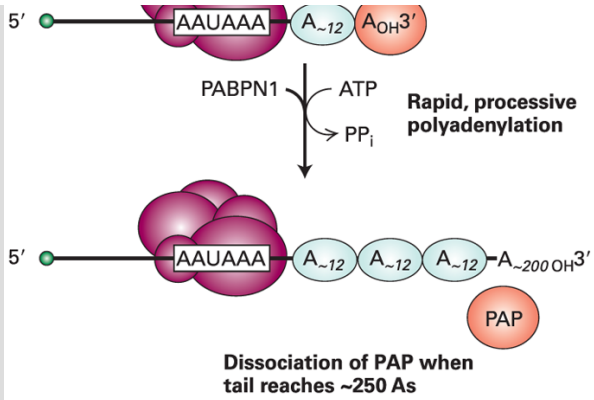


FIGURE 9-15 Model for cleavage and polyadenylation of pre-mRNAs in mammalian cells. Cleavage and polyadenylation specificity factor (CPSF) binds to the upstream AAUAAA poly(A) signal. CStF interacts with a downstream GU- or U-rich sequence and with bound CPSF, forming a loop in the RNA; binding of CFI and CFII helps stabilize the complex. Binding of poly(A) polymerase (PAP) then stimulates cleavage at a poly(A) site, which usually is 10–35 nucleotides $3'$ of the upstream poly(A) signal. The cleavage factors are released, as is the downstream RNA cleavage product, which is rapidly degraded. Bound PAP then adds ~ 12 A residues at a slow rate to the $3'$-hydroxyl group generated by the cleavage reaction. Binding of poly(A)-binding protein (PABPNI) to the initial short poly(A) tail accelerates the rate of addition by PAP. After 200–250 A residues have been added, PABPN1 signals PAP to stop polymerization.

Purification of the proteins required for cleavage and polyadenylation of pre-mRNA precursors in vitro led to the initial identification of proteins involved in the process. This was followed by isolation and sequencing of the genes encoding them, purification of the recombinant proteins expressed from these genes, and then further identification of associated proteins by binding to the overexpressed proteins and by co-purification of epitope-tagged versions of the proteins expressed in vivo. Studies

with these purified proteins led to the model of cleavage and polyadenylation in Figure 9-15.

In *S. cerevisiae*, 14 or 15 polypeptides assemble into an ∼1 million Dalton complex with three modules that perform cleavage and polyadenylation at the poly(A) site. One module binds the AAUAAA RNA sequence in the pre-mRNA and cleaves the RNA 10–35 bases downstream. The variation in distance may be due to different secondary structures in the RNA between the AAUAAA and the cleavage site for each pre-mRNA. The second module of the cleavage polyadenylation complex includes two phosphatases that appear to regulate elongation by RNA polymerase II and polyadenylation. The third module includes the poly(A) polymerase enzyme that adds A's to the 3′ end of the RNA generated when the first module cleaves downstream of AAUAAA sequence.

Following cleavage at the poly(A) site, polyadenylation proceeds in two phases. Addition of the first 12 or so A residues occurs slowly, followed by rapid addition of up to 200–250 more A residues. The rapid phase requires the binding of multiple copies of a *poly(A)-binding protein* containing an RRM motif. This protein is designated *PABPNI* to distinguish it from the poly(A)-binding proteins present in the cytoplasm (*PABPC1,2,3*). PABPN1 binds to the short A tail initially added by PAP, stimulating the rate of polymerization of additional A residues by PAP, resulting in a much faster phase of polyadenylation. PABPN1 is also responsible for signaling poly(A)

polymerase to terminate polymerization when the poly(A) tail reaches a length of 200–250 residues, although the mechanism for controlling the length of the tail is not yet understood. Binding of PABPN1 to the poly(A) tail is essential for mRNA export into the cytoplasm where it is replaced by a PABC1, 2, or 3 and shuttled back into the nucleus.

Alternative Polyadenylation Sites

In addition to alternative splicing, mRNA regulation is also mediated by *alternative polyadenylation,* which occurs for ~50 percent of human mRNAs. Alternative polyadenylation results from the use of two or more alternative cleavage/polyadenylation signals in alternative cell types. In some cases, this appears to be due to different concentrations of the polyadenylation factors in alternative cell types coupled with alternative poly(A) sites that have higher or lower affinity for the cleavage polyadenylation complex that binds the downstream G/U-rich portion of the cleavage/polyadenylation signal (Figure 9-15). In these cases, when the concentration of the factor is low, only the highest affinity cleavage/polyadenylation sites are used. But in alternative cell types where the factor concentration is higher, an upstream low-affinity site is used preferentially because it is bound by the cleavage/polyadenylation factor at the higher concentration, and once the pre-mRNA is cleaved, the downstream site cannot be used. In other cases, sequence specific RNA-binding proteins may block or enhance binding of the

cleavage/polyadenylation factors, similarly to the way they regulate the binding of splicing repressors and activators.

Alternative sites of polyadenylation can also be coupled to alternative splicing of the final exon in an mRNA. As a consequence, protein isoforms can be expressed that have alternative C-terminal amino acid sequences. This is observed in the expression of alternative immunoglobulin molecules during B-lymphocyte development (see Figure 24-19). Initially, an antibody is produced with a transmembrane domain that anchors the antibody in the plasma membrane, and a cytoplasmic domain that signals when the antigen-binding extracellular domain encounters antigen — the molecule bound by an antibody. When that occurs, processing of the pre-mRNA is modified so that an alternative 3′ exon is included in the mRNA. The resulting antibody molecules translated from this alternatively processed mRNA lack the membrane-spanning domain and, as a consequence, are secreted into the extracellular space where they can neutralize pathogens (see Chapters 14 and 24).

RNA Polymerase II Transcription Termination

Pol II generally terminates transcription at any one of several alternative sites within ~2 kb of the poly(A) site. Experiments with SV40 and adenovirus, both DNA viruses, showed that when the polyadenylation signal is mutated, RNA polymerase does not terminate transcription but continues transcription until the next

poly(A) site in the viral genome is encountered. Similar results were observed when a wild-type copy of the β-globin gene was inserted downstream of a viral promoter in the genome of a DNA virus. These experiments showed that transcription termination by RNA polymerase II is coupled to cleavage and polyadenylation of the transcript, and that the RNA fragment downstream of the cleavage site is rapidly degraded. This is thought to be due to the de-protection of the $5'$ end of the nascent transcript. Because no cap is present on the $5'$ end of the cleaved RNA, it is susceptible to digestion by the major $5' \rightarrow 3'$ exoribonuclease in the nucleus, *XRN1*. It is thought that when XRN1 reaches the still transcribing polymerase, it triggers termination, either by pulling the $3'$ end of the nascent RNA out of the polymerase active site, or by introducing a conformational change in the polymerase that causes transcription termination. Once the nascent RNA is removed from the elongating polymerase, the contacts between the RNA polymerase II clamp and the RNA-DNA hybrid within the polymerase (see Figure 8-9b) are lost, allowing the clamp to swing open and release the polymerase from the DNA template.

KEY CONCEPTS OF SECTION 9.1

Processing of Eukaryotic Pre-mRNA

- In the nucleus of eukaryotic cells, pre-mRNAs are associated with hnRNP proteins and processed by $5'$ capping, $3'$ cleavage/polyadenylation, and RNA splicing to remove internal introns before being transported to the cytoplasm (see Figure 9-2).

- In multicellular organisms, splicing of exons in long genes with multiple exons usually begins as the pre-mRNA is still being transcribed. Cleavage and polyadenylation to form the $3'$ end of the mRNA occur after the poly(A) site is transcribed.
- Transcription and RNA processing are coupled via the carboxy-terminal domain (CTD) of RNA polymerase II. $5'$-capping enzymes interact with the CTD phosphorylated at Ser-5 by TFIIH during transcription initiation. The $5'$ cap protects the nascent pre-mRNA from $5'$exonucleases.
- Factors involved in RNA splicing and $3'$ cleavage and polyadenylation associate with the Pol II CTD phosphorylated on Ser-2 by cyclin T-CDK9 during release of the paused polymerase early in gene transcription (Figures 8-13a, 8-16). These interactions with the phosphorylated CTD allow the processing factors to interact with the nascent pre-mRNA shortly after it emerges from the surface of the polymerase.
- Splicing occurs at short, conserved *splice sites* in pre-mRNAs (Figure 9-7c), via two transesterification reactions (Figure 9-8). The first of these generates a *lariat intron* with an unusual $2'$-$5'$ phosphoester bond at the *branch-point A*.
- Five different abundant *small nuclear RNAs (snRNAs)* are found in snRNA-protein complexes (*snRNPs*). These interact via base pairing of their snRNAs with one another and with pre-mRNA to form the 1.3 MDa spliceosome (see Figure 9-11).
- Multiple specific proteins are also associated with the catalytically active spliceosome, comprising 95 percent of its mass (Figure 9-11b).
- The spliceosome is not a stable complex. During assembly of the active spliceosome, some snRNPs and additional splicing factor proteins initially associate with assembly intermediates, and then are released.
- Ten complexes of snRNPs and protein splicing factors have been distinguished during a splicing cycle (Figure 9-13).
- Several RNA helicases are required in the splicing process. They utilize energy from ATP hydrolysis to drive the transesterification reactions in the direction of exon ligation.
- Since the basic catalytic mechanism of the spliceosome is performed by two Mg^{2+} ions bound by U6 (Figure 9-14c), and the substrate splice sites are delivered to the RNA-based active site by interactions with both snRNAs and proteins, the spliceosome is considered to be a *protein-directed metalloribozyme.*
- Following RNA splicing in vertebrates, a specific set of hnRNP proteins remain bound to the spliced RNA approximately 20 nucleotides $5'$ to each exon-exon junction forming an *exon-junction complex (EJC)*. As discussed in later sections, EJCs participate in mRNA export to the cytoplasm and in quality-control surveillance mechanisms that eliminate improperly processed mRNAs.
- The $3'$ end of an mRNA is generated by cleavage of the pre-mRNA at a poly(A) site, an AAUAAA 10-35 bases upstream of the cleavage site and a GU- or U-rich sequence ~ 20

nucleotides downstream from the cleavage site. A poly(A) polymerase in a multisubunit cleavage-polyadenylation complex that performs the cleavage reactions adds a poly(A) tail of ~ 250 nucleotides in vertebrates, ~ 100 nucleotides in yeast, to the $3'$ end. The poly(A) tail is then immediately bound by nuclear poly(A) binding protein (PABPN1), protecting the $3'$ end from digestion by an exosome.

9.2 Regulation of Pre-mRNA Processing

Now that we've seen how pre-mRNAs are processed into mature, functional mRNAs, we consider how regulation of this process can contribute to control of gene expression. Recall from Figure 7-3 that multicellular eukaryotes contain both simple and complex transcription units encoded in their DNA. The primary transcripts produced from the former contain one poly(A) site and exhibit only one pattern of RNA splicing, even if multiple introns are present; thus simple transcription units encode a single mRNA. In contrast, the primary transcripts produced from complex transcription units (~95 percent of all human transcription units) can be processed in alternative ways to yield different mRNAs that encode distinct proteins (see Figure 7-3). This process of **alternative RNA splicing** greatly expands the number of functionally distinct proteins encoded by the ~19,000 human genes (where one gene equals one transcription unit for a primary transcript). The increase in complexity of mRNAs resulting from alternative splicing in metazoans reaches its zenith in the central nervous system, where regulation of complex patterns of alternative RNA splicing is required for generating synaptic connections between neurons, and consequently, appropriate behavior (see Chapter 23).

Additional Nuclear Proteins Contribute to Splice-Site Selection in the Long Pre-mRNAs of Humans and Other Vertebrates

How are splice sites recognized during processing of eukaryotic pre-mRNAs so that exons can be joined precisely to generate long open reading frames? As mentioned above, almost all 5′ splice sites in yeast conform to the sequence/GUAUGU, which is perfectly complementary to the sequence at the 5′ end of U1 snRNA (Figure 9-9). As first shown in the compensatory base pairing experiment described in Figure 9-9b, Watson-Crick base pairing between U1 snRNA and the first six bases of the intron are required for pre-mRNA splicing. In yeast, perfect complementarity between a six-base sequence near the middle of U2 snRNA and the branch-point A sequence near the 3′ end of the intron (Figures 9-9a and 9-12) is also essential for splicing. Only ∼5 percent of *S. cerevisiae* genes contain an intron, and most of these contain a single intron. In this organism, pre-mRNA splice sites are selected primarily by base-pairing to U1 snRNA for the 5′ splice site at the beginning of an intron, and to U2 snRNA at the branch-point sequence just upstream of the 3′ splice site (Figure 9-9a). However, in humans, RNA sequences at 5′ and 3′ splice sites and the branch site are far more degenerate (Figure 9-12). Most introns begin with 5′ GU and end with 3′ AG preceded by a polypyrimidine tract of 10–12 bases, and a

very short, degenerate six-base sequence encompassing the branch-point A. These consensus splice-site sequences are so degenerate that they cannot be used to predict splice sites in the human genome sequence; the vast majority of sequences that match these degenerate sequences are *not* used as splice sites. Consequently, other aspects of human pre-mRNA sequences must contribute to splice site selection.

Extensive mutational analysis of sequences within ~200 bp of splice sites of a number of specific mammalian genes identified sequences in some pre-mRNAs required for splicing in addition to the short, degenerate 5′ splice sites, branch point, and 3′ splice sites (Figure 9-12, *bottom*). These were different specific sequences in different genes within ~200 bases of the 5′ and 3′ splice sites. Characterization of nuclear RNA-binding proteins that bind to these regulatory sequences in RNA, and cDNA cloning and sequencing of related genes, have led to the discovery of at least 1500 human RNA-binding proteins that bind to short 3–4 base RNA sequences. These are hnRNP proteins discussed earlier, and, consequently, consist of one or more RNA-binding domains (RRMs, Figure 9-6) fused to other functional domains. These additional functional domains often either stimulate or inhibit splicing at a splice site within ~200 bases of the binding site for the regulatory hnRNP. This allows relatively short exons (average length ~150 bases in humans) to be recognized in long primary transcripts that can have introns of tens of kb. This is because the splice sites are determined by the interactions of

these RNA-binding proteins with specific sequences in the pre-mRNA in addition to the interactions of the snRNAs with 5′ splice sites and branch points, and of the U2AF protein splicing factor with 3′ splice sites. The RNA-sequences bound by RNA-binding proteins that stimulate RNA splicing are referred to as **splicing enhancers**, and the RNA-sequences bound by RNA-binding proteins that inhibit RNA splicing are referred to as **splicing silencers**.

In humans and other metazoans, the information for defining some splice sites is encoded by splicing enhancers within the sequence of an exon, consequently called *exonic splicing enhancers*. A subset of hnRNP proteins called **SR proteins**, bind exonic splicing enhancers. SR proteins contain one or more RRM RNA-binding domain. They also contain intrinsically disordered regions of polypeptide rich in arginine (R) and serine (S) residues, called *RS domains*, which can interact with RS domains in other RNA-binding proteins. When bound to exonic splicing enhancers, SR proteins mediate the cooperative binding of U1 snRNP to a 5′ splice site and U2 snRNP to a branch point, through a network of protein-protein interactions that span the exon (Figure 9-16). The complex of SR proteins, snRNPs, and other splicing factors (e.g., U2AF) that assemble across an exon, called a **cross-exon recognition complex**, permits precise specification of some exons in long pre-mRNAs. Many other exons are properly spliced without the assistance of SR proteins because they are demarcated by *constitutive splice sites* with good

complementarity to U1 and U2 snRNAs and polypyrimidine tracts and 3′ splice site RNA sequences with high affinity for U2AF.

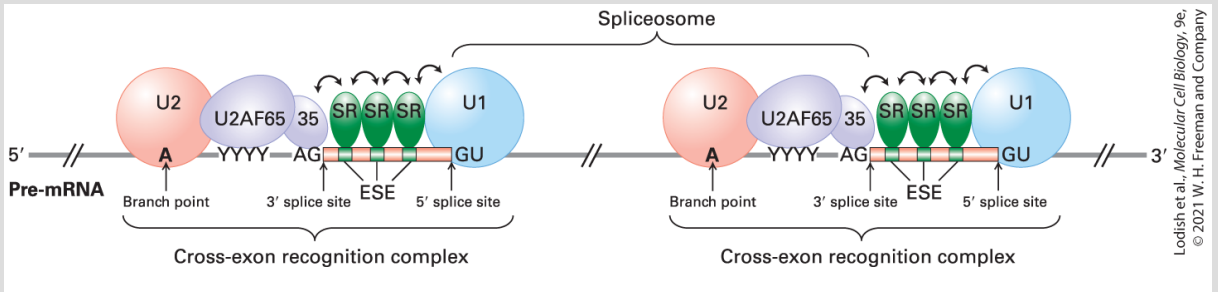


FIGURE 9-16 Exon recognition through cooperative binding of SR proteins and splicing factors to pre-mRNA. The correct 5′-GU and 3′-AG splice sites are recognized by splicing factors on the basis of their proximity to exons. The exons contain exonic splicing enhancers (ESEs) that are binding sites for SR proteins. When bound to ESEs, the SR proteins interact with one another and promote the cooperative binding of the U1 snRNP to the 5′ splice site of the downstream intron, SF1 and then the U2 snRNP to the branch point of the upstream intron, the 65- and 35-kD subunits of U2AF to the pyrimidine-rich region and AG 3′ splice site, respectively, of the upstream intron, and other splicing factors (not shown). The resulting RNA-protein cross-exon recognition complex spans an exon and activates the correct splice sites for RNA splicing. Note that the U1 and U2 snRNPs in this unit do not become part of the same spliceosome. The U2 snRNP on the right forms a spliceosome with the U1 snRNP bound to the 5′ end of the same intron. The U1 snRNP shown on the right forms a spliceosome with the U2 snRNP bound to the branch point of the downstream intron (not shown), and the U2 snRNP on the left forms a spliceosome with a U1 snRNP bound to the 5′ splice site of the upstream intron (not shown). Double-headed arrows indicate protein-protein interactions. See T. Maniatis, 2002, *Nature* **418**:236; see also S. M. Berget, 1995, *J. Biol. Chem.* **270**:2411.

This participation of regulatory proteins in the selection of 5′ and 3′ splice sites allows metazoans to regulate the use of [alternative splice sites]. The sequences of the first cDNAs cloned, including mRNAs for alpha and beta globins, antibody light chains, and

chicken ovalbumin, revealed only a single type of spliced mRNA expressed from the genes encoding them. However, this is a very unusual situation in vertebrates that applies to these unusually abundant mRNAs expressed in highly specialized cell types. Extensive cDNA cloning and sequencing and comparison to genomic DNA sequence has revealed that ~95 percent of mammalian mRNAs are processed from primary transcripts that are alternatively spliced into more than one type of mRNA, often in a cell-type-specific manner, as for fibronectin (Figure 5-28). This greatly increases the total number of distinct proteins encoded in metazoans. Furthermore, by regulating the expression and activities of these RNA-binding proteins that regulate RNA splicing, different mammalian cell types can express different mRNA isoforms from the same transcription unit, and hence different protein isoforms, contributing to the distinct functions of different cell types.

Expression and Function of Related K^+-Channel Protein Isoforms in Vertebrate Inner Ear Hair Cells

As mentioned earlier, alternative RNA splicing is most extensive in the vertebrate nervous system. A good example of how alternative RNA splicing regulates the function of a protein central to neuron function is found in the inner ear of vertebrates. Individual ciliated neurons called *hair cells* respond most strongly to a specific

frequency of sound. Cells tuned to low frequency ($\sim 50\,\mathrm{Hz}$) are found at one end of the tubular cochlea that makes up the inner ear; cells responding to high frequency ($\sim 5000\,\mathrm{Hz}$) are found at the other end (Figure 9-17a). Cells in between the ends respond to a gradient of frequencies between these extremes. One component in the tuning of hair cells in reptiles and birds is the opening of K^+ ion channels in response to increased intracellular Ca^{2+} concentrations. The Ca^{2+} concentration at which the channel opens determines the frequency with which the membrane potential oscillates and hence the frequency to which the cell is tuned.

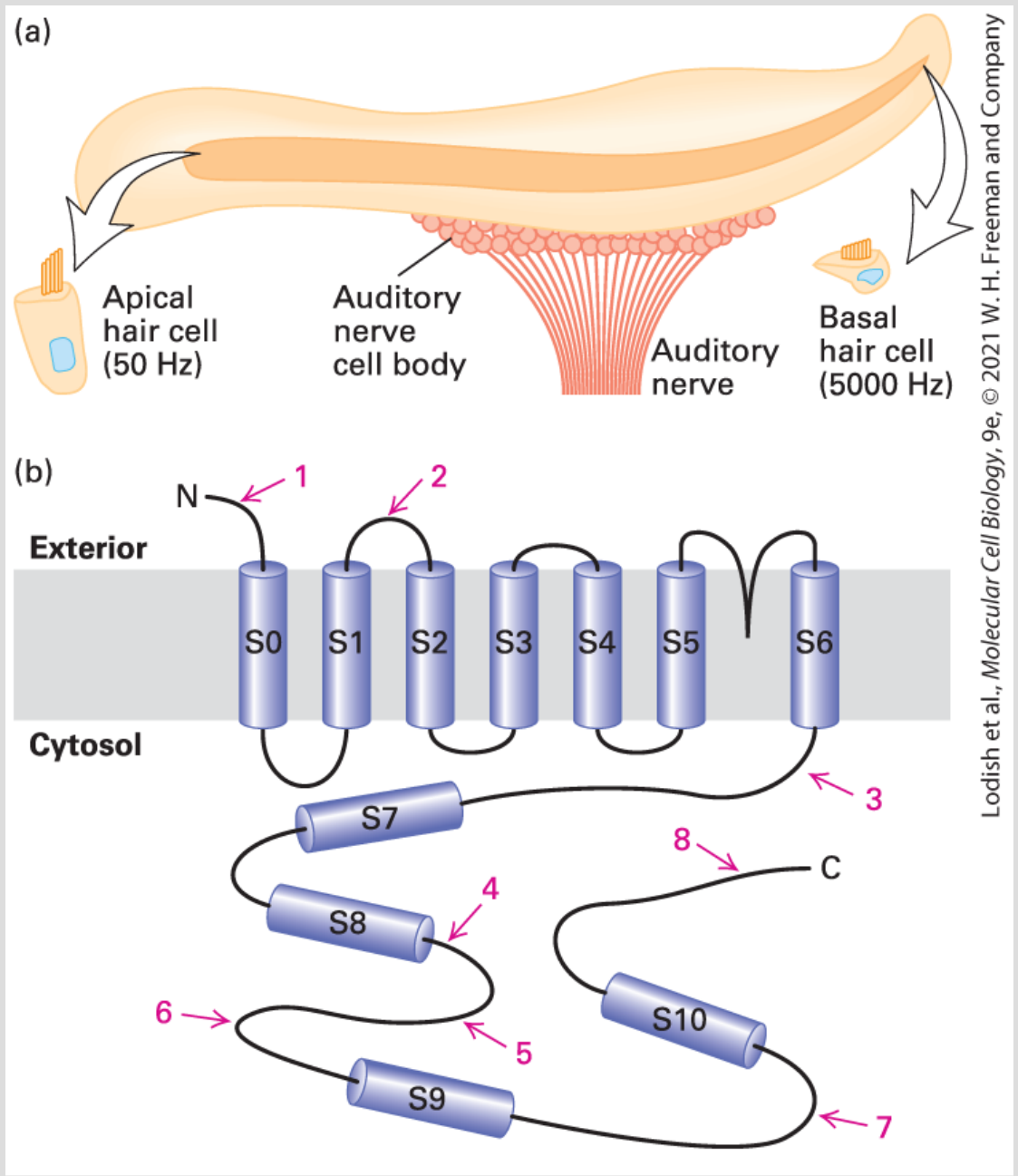


FIGURE 9-17 Function of alternative splicing in the perception of sounds of different frequency. (a) The chicken cochlea, a 5-mm-long tube, contains an epithelium of auditory hair cells that are tuned to a gradient of vibrational frequencies from 50 Hz at the apical end (*left*) to 5000 Hz at the basal end (*right*). (b) The Ca^{2+}-activated K1 channel contains seven transmembrane a helices (S0–S6), which associate to form the channel. The cytosolic domain, which includes four hydrophobic regions (S7–S10), regulates opening of the channel in response to Ca^{2+}. Isoforms of the channel, encoded by alternatively spliced mRNAs produced

from the same primary transcript, open at different Ca^{2+} concentrations and thus respond to different frequencies. Red numbers refer to regions where alternative splicing produces different amino acid sequences in the various isoforms. See K. P. Rosenblatt et al., 1997, *Neuron* **19**:1061.

The gene encoding this Ca^{2+}-activated K^+ channel is expressed as multiple, alternatively spliced mRNAs. The various proteins encoded by these alternative mRNAs open at different Ca^+ concentrations. Hair cells with different response frequencies express different isoforms of the channel protein depending on their position along the length of the cochlea. The sequence variation in the protein is very complex: there are at least eight regions in the mRNA where alternative exons are utilized, permitting the expression of 576 possible isoforms (Figure 9-17b). PCR analysis of mRNAs from individual hair cells has shown that each hair cell expresses a mixture of different alternative Ca^{2+}-activated K^+-channel mRNAs, with different forms predominating in different cells according to their position along the cochlea. This remarkable arrangement suggests that splicing of the Ca^{2+}-activated K^+-channel pre-mRNA is regulated in response to extracellular signals that inform the cell of its position along the cochlea.

Other studies demonstrated that splicing at one of the alternative splice sites in the Ca^{2+}-activated K^+-channel pre-mRNA in the rat is suppressed when a specific protein kinase is activated by neuron depolarization in response to synaptic activity from interacting neurons. This observation raises the possibility that a splicing

repressor specific for this site may be activated when it is phosphorylated by this protein kinase, whose activity in turn is regulated by synaptic activity. Since hnRNP and SR proteins are extensively modified by phosphorylation and other post-translational modifications, it seems likely that complex regulation of alternative RNA splicing through post-translational modifications of splicing factors has a significant function in modulating neuron function.

Regulation of RNA Splicing Through Splicing Enhancers and Silencers Controls *Drosophila* Sexual Differentiation

One of the earliest and best understood examples of regulated alternative splicing of pre-mRNA came from studies of sexual differentiation in *Drosophila*. Genes required for normal *Drosophila* sexual differentiation were first characterized by isolating *Drosophila* mutants defective in the process. When the proteins encoded by the wild-type genes were characterized biochemically, two of them were found to regulate a cascade of alternative RNA splicing in *Drosophila* embryos. Subsequent research provided insight into how these proteins regulate RNA processing and ultimately lead to the expression of one or the other of the two sex-specific transcriptional

repressors that suppress the development of characteristics of the opposite sex.

The Sxl protein, encoded by the *sex-lethal* gene, is the first protein to act in the cascade (Figure 9-18). The Sxl protein is present only in female embryos. Early in development, the *Sxl* gene is transcribed from a promoter that functions only in early female embryos. Later in development, this female-specific promoter is shut off and another promoter for *sex-lethal* becomes active in both male and female embryos. However, in the absence of early Sxl protein, the *sex-lethal* pre-mRNA in male embryos is spliced to produce an mRNA that contains a stop codon early in the sequence (Figure 9-18a, blue shading). The net result is that male embryos produce no functional Sxl protein either early or later in development.

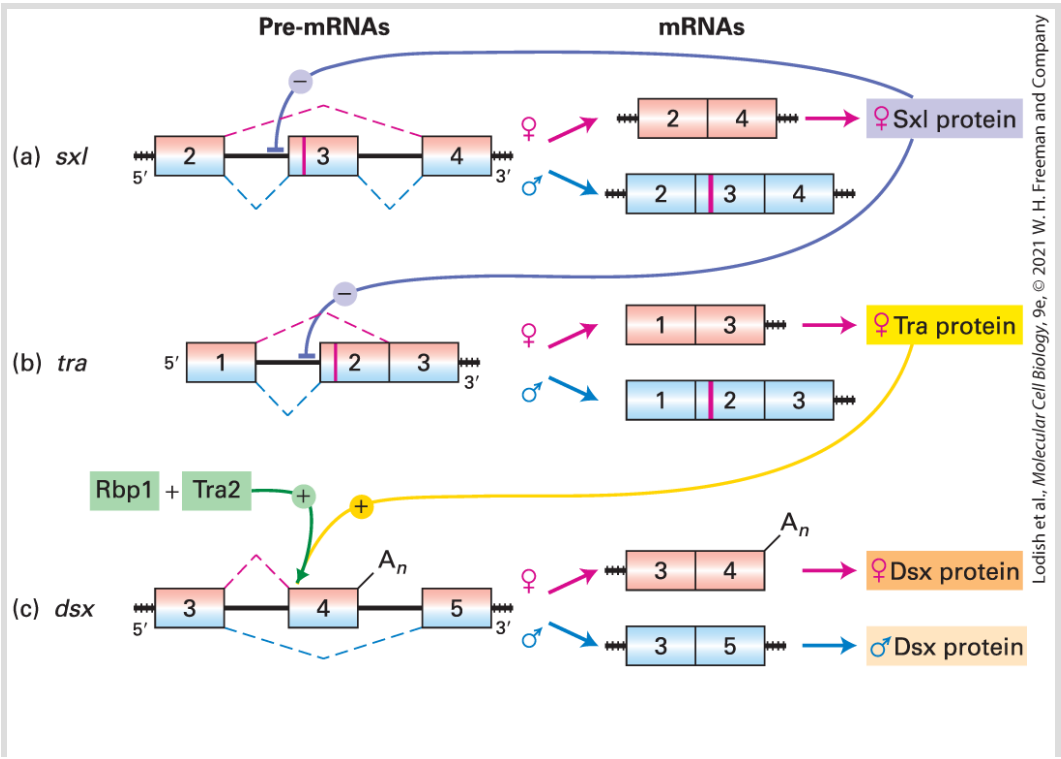


FIGURE 9-18 Cascade of regulated splicing that controls sex determination in *Drosophila* embryos. For clarity, only the exons (boxes) and introns (black lines) where regulated splicing occurs are shown. Splicing is indicated by red dashed lines above (female) and blue dashed lines below (male) the pre-mRNAs. Vertical red lines in exons indicate in-frame stop codons, which prevent synthesis of functional protein. Only female embryos produce functional Sxl protein, which represses splicing between exons 2 and 3 in sxl pre-mRNA (a) and between exons 1 and 2 in tra pre-mRNA (b). (c) In contrast, the cooperative binding of Tra protein and two SR proteins, Rbp1 and Tra2, activates splicing between exons 3 and 4 and cleavage/polyadenylation (A_n) at the 3′ end of exon 4 in dsx pre-mRNA in female embryos. In male embryos, which lack functional Tra, the SR proteins do not bind to exon 4, and consequently exon 3 is spliced to exon 5. The distinct Dsx proteins produced in female and male embryos as the result of this cascade of regulated splicing repress transcription of genes required for sexual differentiation of the opposite sex. See M. J. Moore et al., 1993, in R. Gesteland and J. Atkins, eds., *The RNA World*, Cold Spring Harbor Press, pp. 303–357.

In contrast, the Sxl protein expressed in early female embryos directs splicing of the *sex-lethal* pre-mRNA so that a functional *sex-lethal* mRNA is produced (Figure 9-18a). Sxl accomplishes this by binding to a sequence in the pre-mRNA near the 3′ end of the intron between exon 2 and exon 3, thereby blocking the proper association of U2AF and U2 snRNP with the branch-point A and the 3′ AG (Figure 9-16). As a consequence, the U1 snRNP bound to the 3′ end of exon 2 assembles into a spliceosome with U2 snRNP bound to the branch point at the 3′ end of the intron between exons 3 and 4, leading to splicing of exon 2 to exon 4 and skipping of exon 3 (Figure 9-18a). The binding site for Sxl in the Sxl pre-mRNA is called an

intronic splicing silencer because of its location in an intron and its function in blocking, or "silencing," use of a splice site. The resulting female-specific *sex-lethal* mRNA is translated into functional Sxl protein, which reinforces its own expression in female embryos by continuing to cause skipping of exon 3. The absence of Sxl protein in male embryos allows the inclusion of exon 3 and, consequently, of the stop codon that prevents translation of functional Sxl protein.

Sxl protein also regulates alternative RNA splicing of the *transformer* gene pre-mRNA, the second gene in the regulatory cascade (Figure 9-18b). In male embryos, where no Sxl is expressed, *tra* exon 1 is spliced to *tra* exon 2, which contains a stop codon that prevents synthesis of a functional *transformer* protein, very similar to the inhibition of Sxl protein in male cells (Figure 9-18a). In female embryos, however, binding of Sxl protein to an intronic splicing silencer at the 3′ end of the intron between exons 1 and 2 blocks binding of U2AF at this site. The interaction of Sxl with *transformer* pre-mRNA is mediated by two RRM domains in the Sxl protein (see Figure 9-6b). When Sxl is bound, U2AF binds to a lower affinity site farther 3′ in the pre-mRNA. As a result, exon 1 is spliced to this alternative 3′ splice site, eliminating exon 2 with its stop codon. The resulting female-specific *transformer* mRNA, which contains additional constitutively spliced exons, is translated into functional Transformer (Tra) protein.

In the last step in the regulatory cascade, Tra protein regulates the alternative processing of pre-mRNA transcribed from the *double-sex* gene (Figure 9-18c). In female embryos, a complex of Tra and two constitutively expressed proteins, Rbp1 and Tra2, directs splicing of exon 3 to exon 4 and also promotes cleavage/polyadenylation at the alternative poly(A) site at the 3′ end of exon 4 — leading to a short, female-specific version of the Dsx protein. In male embryos, which produce no Tra protein, exon 4 is skipped, so that exon 3 is spliced to exon 5. Exon 5 is constitutively spliced to exon 6, which is polyadenylated at its 3′ end — leading to a longer, male-specific version of the Dsx protein. The RNA sequence to which Tra binds in exon 4 is called an *exonic splicing enhancer* since it enhances splicing at a nearby splice site.

As a result of this cascade of regulated RNA processing depicted in Figure 9-18, different Dsx proteins are expressed in male and female embryos. The male Dsx protein is a transcriptional repressor that inhibits the expression of genes required for female development. Conversely, the female Dsx protein represses transcription of genes required for male development. In wild-type *Drosophila*, no Sxl protein is expressed in cells of male embryos, whereas it is expressed in female embryos. This is an example of "on-off switch" regulation. The initial expression of Sxl in early female embryos positively regulates its own expression. Other examples of nuclear RNA-binding proteins that generate switch-like control of alternative

RNA splicing in neurons have been identified and termed *master splicing factors.*

Figure 9-19 illustrates how the Tra/Tra2/Rbp1 complex is thought to interact with *double-sex (dsx)* pre-mRNA. Rbp1 and Tra2 are SR proteins, but they do not interact with exon 4 in the absence of the Tra protein. Tra protein interacts with Rbp1 and Tra2, resulting in the cooperative binding of all three proteins to six exonic splicing enhancers in exon 4. The bound Tra2 and Rbp1 proteins then promote the binding of U2AF and U2 snRNP to the 3′ end of the intron between exons 3 and 4 (see Figure 9-17). The Tra/Tra2/Rbp1 complexes may also enhance binding of the cleavage/polyadenylation complex (Figure 9-15) to the 3′ end of exon 4.

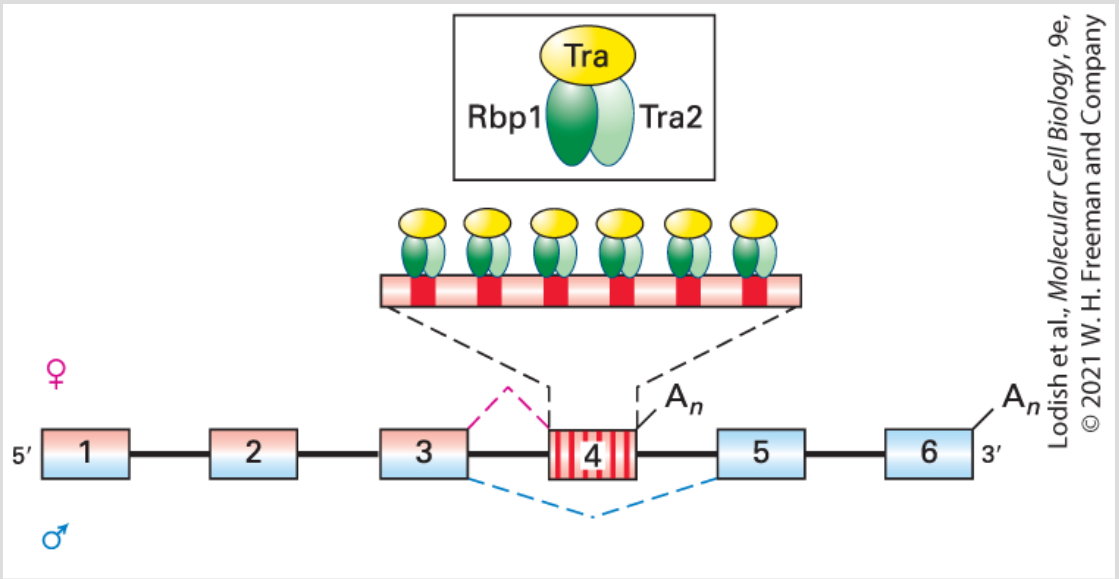


FIGURE 9-19 Model of splicing activation by Tra protein and the SR proteins Rbp1 and Tra2. In female *Drosophila* embryos, splicing of exons 3 and 4 in dsx pre-mRNA is activated by binding of Tra/Tra2/Rbp1 complexes to six sites in exon 4. Because Rbp1

and Tra2 cannot bind to the pre-mRNA in the absence of Tra, exon 4 is skipped in male embryos. See the text for discussion. A_n represents $3'$ polyadenylation. See T. Maniatis and B. Tasic, 2002, *Nature* **418**:236.

Splicing Repressors and Activators Control Splicing at Alternative Sites

As is evident from Figure 9-18, the *Drosophila* Sxl protein and Tra protein have opposite effects: Sxl prevents splicing, causing exons to be skipped, whereas Tra promotes splicing. The action of similar proteins may explain the cell-type-specific expression of fibronectin isoforms in humans. For instance, an Sxl-like splicing repressor expressed in hepatocytes might bind to splice sites for the EIIIA and EIIIB exons in the fibronectin pre-mRNA, causing them to be skipped during RNA splicing (Figure 5-28). Alternatively, a Tra-like splicing activator expressed in fibroblasts might activate the splice sites associated with the fibronectin EIIIA and EIIIB exons, leading to inclusion of these exons in the mature mRNA. Experimental examination in some systems has revealed that inclusion of an exon in some cell types versus skipping of the same exon in other cell types results from the combined influence of several splicing repressors and enhancers. RNA binding sites for repressors can also occur in exons, where they are called *exonic splicing silencers.* And binding sites for splicing activators, usually SR proteins, can also occur in introns, where they are called *intronic splicing enhancers.*

Expression of *Dscam* Isoforms in *Drosophila* Retinal Neurons

The most extreme example of regulated alternative RNA processing yet uncovered occurs in expression of the *Dscam* gene in *Drosophila*. Mutations in this gene interfere with the normal synaptic connections made between axons and dendrites during fly development. Analysis of the *Dscam* gene showed that it contains 95 alternatively spliced exons that could be spliced to generate 38,016 possible isoforms (Figure 9-20)! One of 12 possible exons encoding domain Ig2 of the protein is included in each Dscam mRNA. One of 48 possible exons encoding domain Ig3 is included in each Dscam mRNA, and one of 33 possible exons encoding domain Ig7 and one of 2 possible exons encoding the Dscam transmembrane domain are included in each Dscam RNA. Exons diagrammed at the top as vertical black lines are constitutively spliced into all Dscam mRNAs. As a result, the Dscam pre-mRNA is processed into one of $12 \times 48 \times 33 \times 2 = 38{,}016$ possible isoforms. *Drosophila* mutants with a version of the gene that can be spliced in only about 22,000 different ways have specific defects in connectivity between neurons. These results indicate that expression of most of the possible *Dscam* isoforms through regulated RNA splicing helps to specify the tens of millions of different specific synaptic connections between neurons in the *Drosophila* brain. In other words, the correct wiring of neurons in the brain requires regulated RNA splicing.

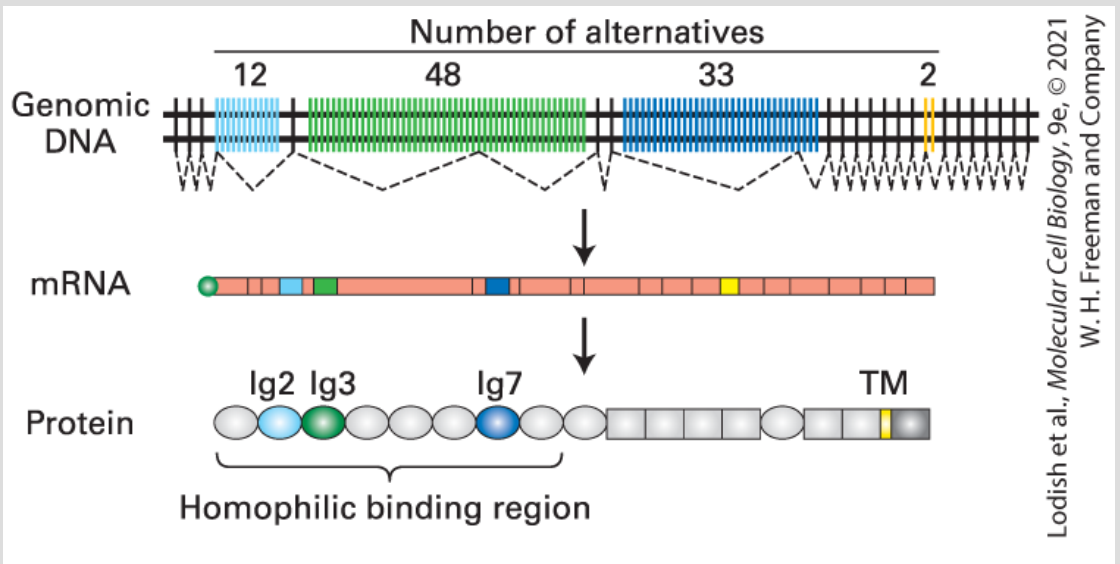


FIGURE 9-20 The *Drosophila Dscam* gene is processed into a vast number of alternative isoforms. *Dscam* encodes a cell surface protein on neurons. The protein (*bottom*) is composed of 10 different immunoglobulin domains (ovals), six different fibronectin type III domains (rectangles), one transmembrane domain (yellow), and a C-terminal cytoplasmic domain. One of 12 possible Ig2 domains, one of 48 possible Ig3 domains, one of 33 possible Ig7 domains, and one of two possible transmembrane domains, each encoded by one exon, are included in the protein because of the alternative RNA splicing diagrammed at the top. The fully processed mRNA contains one of the alternative exons encoding an Ig2 domain diagrammed in light blue in genomic DNA, one of the alternative exons encoding Ig3 shown in green in genomic DNA, and one of the exons shown in yellow. The exons shown in black are constitutively spliced into each of the messages, generating $12 \times 48 \times 33 \times 2 = 38{,}016$ possible isoforms. See M. R. Sawaya et al., 2008, *Cell* **134**:1007.

Abnormal RNA Splicing and Disease

Multiple examples of human diseases result from abnormalities in RNA splicing. Here we look at examples of mutations that affect exon definition and defects in microexon splicing.

Mutations Affecting Exon Definition Cause Disease

Mutations that affect exon definition are responsible for approximately 15 percent of the single-base-pair mutations that cause human genetic diseases. Some single-base-pair mutations occur in 5′ or 3′ splice sites, often resulting in the use of nearby alternative *cryptic* splice sites present in the normal gene sequence. In the absence of the normal splice site, the cross-exon recognition complex (Figure 9-16) recognizes these alternative sites, which generally are a poorer match to consensus splice-site sequences than the splice site in the wild-type gene that was mutated. Other mutations that cause abnormal splicing result in a new 5′- or 3′-consensus sequence in a pre-mRNA that becomes recognized in place of the normal splice site. This generally results in processing of an abnormal mRNA that does not encode a functional protein. Finally, some mutations can interfere with binding of a pre-mRNA to a specific SR protein. These mutations interfere with the assembly of a cross-exon recognition complex on the mutated exon, and thus lead to exon skipping.

Understanding the mechanisms of exon definition has led to strategies to treat Duschenne Muscular Dystrophy (DMD). Mutations in the sex-linked DMD gene encoded on the X-chromosome are responsible for muscular dystrophy in approximately one in 5000 males, making it one of the most common human genetic diseases. The gene encodes dystrophin, a protein that participates in linking the actin-cytoskeleton to the muscle cell plasma membrane (see

Figure 20-39). When dystrophin is defective, the connection is weakened and muscle cell contraction by movement of actin fibers damages the plasma membrane leading to apoptosis in the mutant muscle cells. DMD mutations in many patients result from a frame shift or stop codon in a DMD exon that evolved from exon duplication (e.g., the tan-colored exons in the fibronectin gene in Figure 5-28). Synthetic oligonucleotides that hybridize to the 5′ splice site of the exon with the mutation prevent U1 snRNP from binding to the 5′ splice site and cause skipping of the mutant exon and in-frame splicing of the preceding exon to the exon following the mutant exon. The resulting mRNA lacks any chain terminating mutations. It encodes a protein lacking one copy of the repeated protein domain, but that protein functions almost as well as the wild-type protein. Current work to develop more efficient membrane-permeant synthetic oligonucleotide derivatives and delivery methods are an active area of research for treatment of DMD due to mutations in repeated exons arising from exon duplication. ■

Spinal Muscular Atrophy is one of the most common genetic causes of childhood mortality. This disease results from mutations in a region of the genome containing two closely related genes, *SMN1* and *SMN2*, that arose by gene duplication during human evolution. *SMN2* encodes a protein identical with *SMN1*; however, correctly spliced *SMN2* mRNA is expressed at a much lower level than *SMN1* mRNA because of a mutation in an exonic splicing

enhancer in the duplicated *SMN2* gene. This mutation resulted in a synonymous codon that does not change the encoded amino acid sequence but does inhibit binding of an SR protein to this region of the *SMN2* pre-mRNA. This leads to skipping of the exon during processing of most *SMN2* primary transcripts because a cross-exon-recognition complex (Figure 9-16) cannot assemble across the affected exon, and consequently cannot stimulate association of U1 snRNP and U2 snRNP with 5′ and 3′ splice sites for this exon in the *SMN2* pre-mRNA. This has little consequence in unaffected individuals because most Smn protein is expressed from the *SMN1* gene.

Genetic studies in the mouse indicate that expression of some Smn protein is required for normal mammalian embryonic development. In the mouse, there is only one *Smn* gene. Mouse embryos with inactivating mutations in both the maternal and paternal copies of the *Smn* gene die early during embryogenesis. However, heterozygous mutants that retain one functional copy of the *Smn* gene develop normally because the lower level of Smn protein expressed from the single copy of the functional gene is sufficient for viability. In other words, inactivating mutations in the mouse *Smn* gene are recessive, as is the case for the majority of protein-coding genes.

The human disease spinal muscular atrophy results when a patient inherits inactivating mutations in both the maternal and paternal

alleles of *SMN1,* the gene that expresses most Smn protein. In these patients, the low level of Smn protein translated from the small fraction of *SMN2* mRNA that is correctly spliced is sufficient to maintain cell viability during embryogenesis and fetal development. However, it is not sufficient to maintain viability of spinal cord motor neurons in childhood, resulting in their death and the associated disease. ■

Microexons and Their Association with Autism Spectrum Disorder

Several types of alternative pre-mRNA splicing have been observed in vertebrate cells (Figure 9-21a). The most common and extensively conserved type of alternatively spliced exons in the central nervous system (CNS) are called **microexons** because they are unusually short: 3 to 27 bases long in multiples of three so that they maintain the reading frames of up- and downstream exons of more typical longer lengths (average ∼150 bases in humans). The discovery of these microexons was surprising because their very short length was expected to prevent the previously understood standard cross-exon recognition complex interactions needed for exon definition (Figure 9-16). The short stretches of amino acids encoded by microexons (1–9 amino acids) generally reside on protein surfaces that participate in protein-protein interactions, and are significantly enriched in genes with critical functions in synaptic biology (Figure 9-21b). Of critical importance here, ∼1/3 of patients with autism

spectrum disorder (ASD) exhibit abnormally low levels of microexon splicing in the CNS.

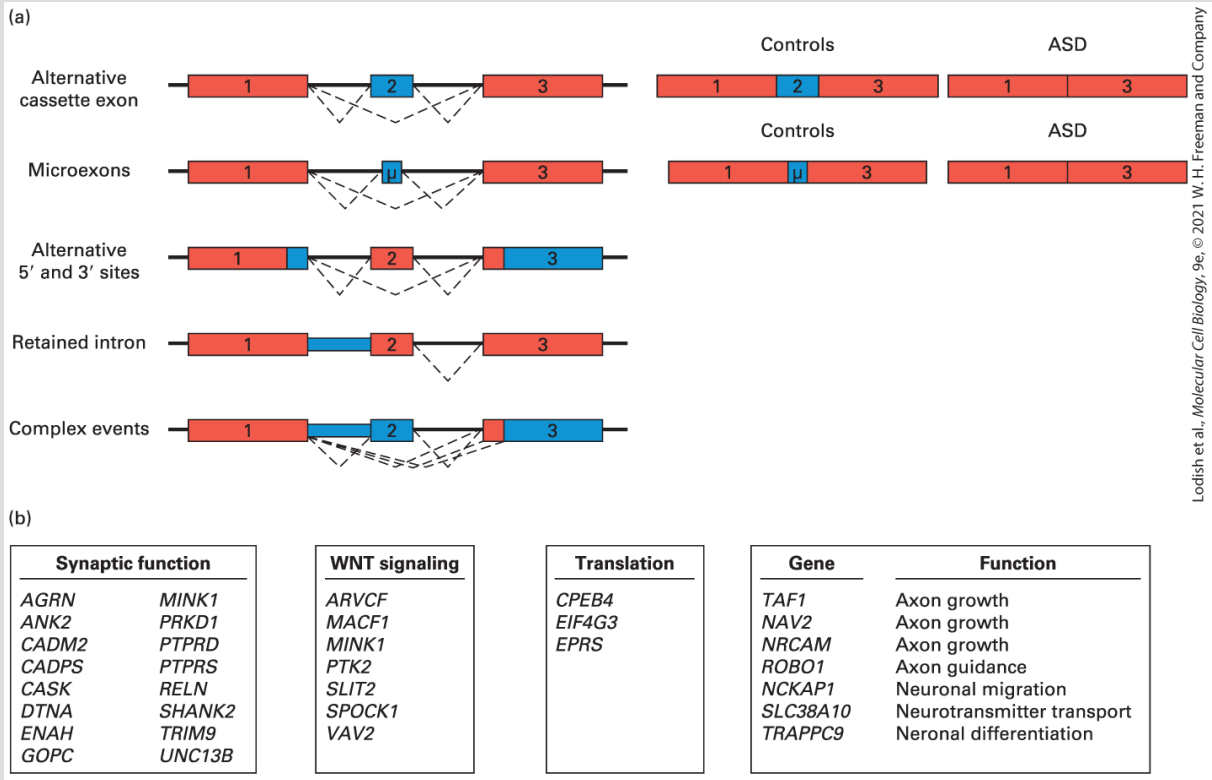


FIGURE 9-21 Splicing changes in autism spectrum disorder (ASD). (a) Types of alternative splicing. Thin black lines indicate splicing between the indicated splice sites. RNA sequencing analyses (RNA-seq) revealed that cassette exons and microexons (3–27 nucleotides) are much more common forms of alternative RNA splicing in the normal human central nervous system than alternative $5'$ and $3'$ splice sites, retained introns, or complex events. In one-third of patients with autism spectrum disorder (ASD), more skipping of cassette exons and microexons are observed than in normals. Abnormalities in the frequency of splicing of alternative $5'$ and $3'$ splice sites are observed only rarely in ASD. Complex splicing events, while present, occur at too low an incidence to determine if their incidence is altered significantly in ASD patients. (b) Genes containing microexons with abnormal inclusion in ASD patients are listed in the context of molecular hubs of ASD pathogenesis (colored boxes) or their functions in neuronal biology (grey box).

[Data from M. Quesnel-Vallières et al., 2019, *Nat. Rev. Genet.* **20**:51.]

Further research on the mechanism underlying this defect in microexon splicing focused on a particular, neuron-specific SR-protein, SRRM4. Human SRRM4 mRNA was expressed at low level compared to normal in the same $\sim 1/3$ of ASD patients with abnormally low levels of included microexons. Further, the number of microexons that were skipped in these patients correlated with the severity of the decrease in SRRM4 expression. The importance of SRRM4 in some forms of ASD was further supported by genetic experiments in mice. Homozygous knock-out of the *Srrm4* gene in mice is lethal in most pups due to postnatal defects in respiration linked to abnormal innervation of the diaphragm and other neurological defects. Heterozygous *Srrm4* +/- mice are viable, but they express approximately one-half the level of Srrm4 mRNA than *Srrm4* +/+ littermates. This is associated with a significant decrease in microexon inclusion in the CNS and with multiple autistic-like features, including altered social behaviors, altered synaptic density, and signaling. These results provide evidence that misregulation of human *SRRM4* is causally linked to a substantial fraction of autism cases.

Mouse Srrm4 protein has been found to stimulate the inclusion of microexons by binding to RNA sequences found upstream from microexons. From this position, Srrm4 is able to stimulate snRNP U1 association with the microexon $5'$ splice site in much the same way that SR-proteins bound to exonic splicing enhancers at the $3'$ end of an exon stimulate snRNP U1 association with $5'$ splice sites in exons

of typical length (see Figure 9-16). Once U1 snRNP is associated with the microexon 5′ splice site, it can interact with U2 snRNP associated with the branch point preceding the next 3′ splice site in the pre-mRNA and participate in a splicing reaction that joins the microexon to the downstream exon.

Myelodysplastic syndromes (MDS) are a group of blood cell diseases characterized by deregulated production of dysplastic (abnormal) myelocytes (phagocytic white blood cells). A high percentage of MDS patients progress to acute myeloid leukemia (AML), which is terminal if not controlled. Remarkably, sequencing of all exons (the exome) from MDS patients revealed that mutations in genes encoding a few specific RNA splicing factors occur in ∼60–70 percent of all MDS patients. The vast majority of these mutations occur in genes for the splicing factors U2AF35, SRSF2, ZRSR2, or SF3B1 (Figure 9-22). U2AF35 is the small subunit of the heterodimeric U2AF splicing factor required for association of the U2 snRNP with the branch-point A near the 3′ end of an intron (Figure 9-13). SRSF2 and ZRSR2 are SR proteins that associate with exonic splicing enhancers in many genes. SF3B1 is a protein component of human U2 snRNP. Specific mutations in other splicing factors (U2AF65, SF1, SF3A1 and PRPF40B) also occur in MDS patients, but less frequently. U2AF65 is the large subunit of U2AF that associates with the pyrimidine tract just upstream of the 3′ splice site. SF1 is required along with U2AF for association of U2 snRNP with the branchpoint A. SF3A1 is another protein component

of human U2 snRNP. PRP40B is an essential splicing factor, but its precise function is not yet understood.

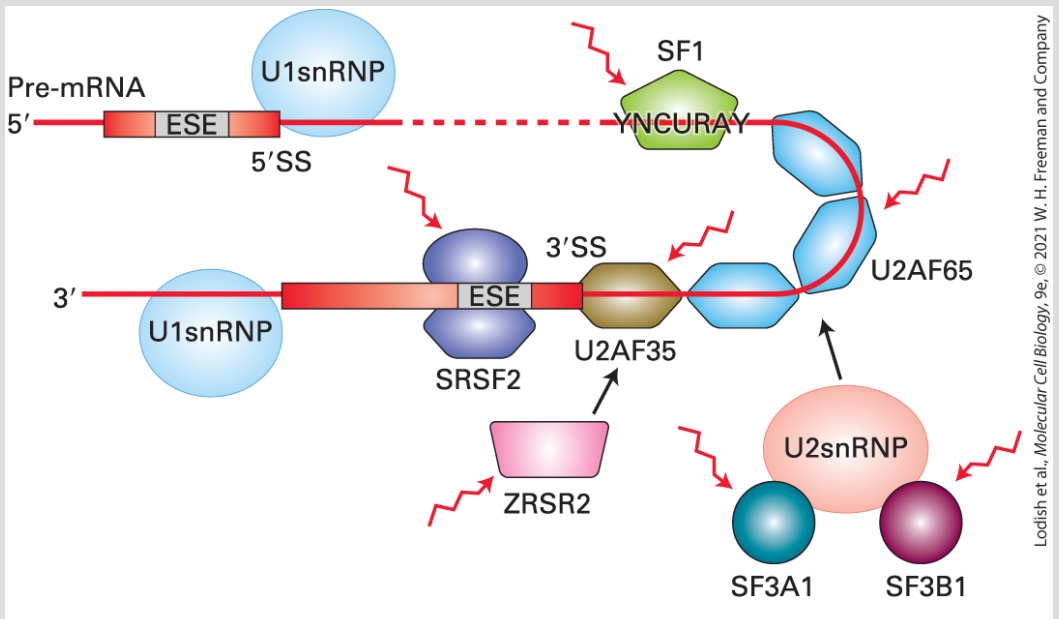


FIGURE 9-22 Specific amino acid changes in some proteins required for U2 snRNP binding to the branch-point A are associated with myelodysplastic syndrome and some leukemias. Spliceosome components mutated in myelodysplasic syndromes, related diseases causing reduced cell replication and abnormal development of myelocytes (phagocytic white blood cells). RNA splicing is initiated by the recruitment of U1 snRNP to a $5'$ splice site. SF1 binds the branch-point sequence, and the larger subunit of the U2 auxiliary factor (U2AF65) binds the downstream polypyrimidine tract. The smaller subunit of U2AF (U2AF35) binds to the AG dinucleotide of the $3'$ splice site. U2 snRNP association with the branch-point A also requires two protein splicing factors, SF3A1 and SF3B1. U2AF interactions with an SR-protein bound near the $5'$ end of the downstream exon are also required (see Figure 9-15). SRSF2 and ZRSR2 are SR-proteins that fulfill this function at a large number of $3'$ splice sites. Specific amino acid changes in one of these splicing factors indicated by a lightning bolt is found in the abnormal myelocytes in 60–70 percent of patients with myelodysplastic syndrome. These mutations lead to abnormal inclusion of alternatively spliced exons and introns in a large number of pre-mRNAs. The mutations are lethal when homozygous, but somehow interfere with normal

replication and development when half of the splicing factors have these specific amino acid changes in heterozygous cells. See text.

[Data from K. Yoshida et al., 2011, *Nature* **478**:64.]

Only one of these mutations in a splicing factor gene is observed in the abnormal myeloctyes from any one MDS patient, and the MDS cells always retain a wild-type allele of the mutant gene. That is, the myelodysplastic cells are always heterozygous for the splicing factor mutation. *Most remarkable of all: the same specific mutations recur in different patients.* For example, all MDS-associated mutations in SRSF2 change proline 95 into another amino acid, either histidine, leucine, or arginine. MDS-associated mutations in the small subunit of U2AF, which interacts with the AG at the 3′ end of most introns, either change serine 34 to a phenylalanine or a tyrosine, or change glutamine 157 to an arginine or a proline.

Genetic techniques described in Chapter 6 permitted researchers to substitute MDS-associated mutations for the wild-type splicing factor genes in cultured cells. Cells with homozygous MDS-associated mutations were not viable, indicating that co-expression of the wild-type splicing factor is essential. This is presumably to generate a sufficient number of normally spliced mRNAs for the heterozygous cells to survive. Analysis of RNA splicing in the heterozygous cells showed global abnormalities, as in the abnormal myeloid cells from MDS-patients. The very specific mutations recovered from MDS patients appear to generate a change of

splicing factor function rather than loss of function. For example, loss of SRSF2 results in failure of cassette exon splicing (Figure 9-22), while expression of MDS-associated SRSF2 mutants result in a change in cassette exon splicing in a sequence-specific manner based on the sequence of exonic splicing enhancer (ESE) motifs present in the abnormally spliced exon. In contrast, mutations in U2AF1, which interacts with the pyrimidine tract upstream from the 3′ splice site, alter splicing by promoting the recognition of abnormal 3′ splice sites. Mutations in SF3B1 promote usage of cryptic 3′ splice sites.

How these abnormalities in pre-mRNA splicing influence the phenotype of myelodysplastic cells is not known. In particular, it is not known if the abnormal phenotype is due to (1) expression of one or a few abnormally spliced mRNAs, influencing the function of one or a small number of proteins that regulate myeloid blood cell replication and differentiation, or (2) whether the phenotype results from the pleiotropic effects of abnormal expression of a large number of protein isoforms. ■

Self-Splicing Group II Introns Provide Clues to the Evolution of snRNAs

Under certain nonphysiological in vitro conditions, pure preparations of some RNA transcripts slowly splice out introns in

the absence of any protein. This observation led to the recognition that some introns are *self-splicing*. Two types of self-splicing introns have been discovered: **group I introns**, present in nuclear rRNA genes of protozoans, and **group II introns**, present in protein-coding genes and some rRNA and tRNA genes in mitochondria and chloroplasts of plants and fungi. Discovery of the catalytic activity of self-splicing introns revolutionized concepts about the functions of RNA. As discussed in Chapter 4, RNA is now known to catalyze peptide-bond formation during protein synthesis in ribosomes. Here we discuss the probable role of group II introns, now found only in mitochondrial and chloroplast DNA, in the evolution of snRNAs; the functioning of group I introns is considered in the later section on rRNA processing.

Even though their precise sequences are not highly conserved, all group II introns fold into a conserved, complex secondary structure containing numerous stem loops (Figure 9-23a). Self-splicing by a group II intron occurs via two transesterification reactions, involving intermediates and products analogous to those found in nuclear pre-mRNA splicing. The mechanistic similarities between group II intron self-splicing and spliceosomal splicing led to the hypothesis that snRNAs function analogously to the stem loops in the secondary structure of group II introns. According to this hypothesis, snRNAs interact with 5′ and 3′ splice sites of pre-mRNAs and with each other to produce a three-dimensional RNA structure

functionally analogous to that of group II self-splicing introns (Figure 9-23b).

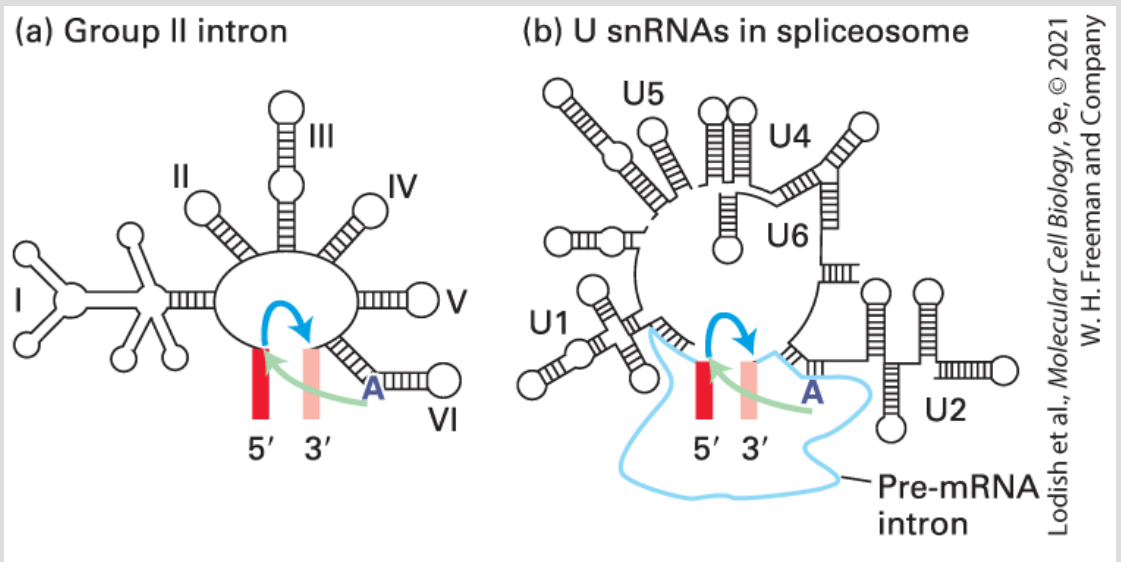


FIGURE 9-23 Comparison of group II self-splicing introns and the spliceosome. The schematic diagrams compare the secondary structures of (a) group II self-splicing introns and (b) U snRNAs present in the spliceosome. The first transesterification reaction is indicated by light green arrows; the second reaction, by blue arrows. The branch-point A is boldfaced. The similarity in these structures suggests that the spliceosomal snRNAs evolved from group II introns, with the trans-acting snRNAs being functionally analogous to the corresponding domains in group II introns. The colored bars flanking the introns in (a) and (b) represent exons. See P. A. Sharp, 1991, *Science* **254**:663.

An extension of this hypothesis is that introns in ancient pre-mRNAs evolved from group II self-splicing introns through the progressive loss of internal RNA structures, which concurrently evolved into trans-acting snRNAs that perform the same functions. Support for this type of evolutionary model comes from experiments with group II intron mutants in which domain V and part of domain I are

deleted (Figure 9-23a). RNA transcripts containing such mutant introns are defective in self-splicing, but when RNA molecules equivalent to the deleted regions are added to the in vitro reaction, self-splicing occurs. This finding demonstrates that these domains in group II introns can be trans-acting, like snRNAs.

Further evidence that spliceosomal splicing evolved from group II introns comes from recent cryo-EM structures of group II introns showing that the RNA interactions that occur during group II intron splicing are very similar to the interactions at the spliceosome active site (see Figure 9-14b, c). Although group II introns can self-splice in vitro at elevated temperatures and Mg^{2+} concentrations, under in vivo conditions proteins called *maturases,* which bind to group II intron RNA, are required for rapid splicing. Maturases are thought to stabilize the precise three-dimensional interactions of the intron RNA required to catalyze the two splicing transesterification reactions. By analogy, we can now see from the cryo-EM structures of catalytically active spliceosomes (see Figure 9-11) that proteins associated with spliceosomes also stabilize the precise geometry of snRNAs and intron nucleotides required to catalyze pre-mRNA splicing.

The evolution of snRNAs may have been an important step in the rapid evolution of multicellular eukaryotes. As internal intron sequences were lost and their functions in RNA splicing supplanted by trans-acting snRNAs, the remaining intron sequences would be

free to diverge. This in turn likely facilitated the evolution of new genes through **exon shuffling** since there are few constraints on the sequence of new introns generated in the process (see Figures 7-18 and 7-19). It also permitted the increase in protein diversity that results from alternative RNA splicing and an additional level of gene control resulting from regulation of alternative RNA splicing.

Nuclear Exonucleases and the Exosome Degrade RNA That Is Processed out of Pre-mRNAs

Because the human genome contains long introns, only ~5 percent of the nucleotides that are polymerized by RNA polymerase II during transcription are retained in mature, processed mRNAs. Although this process appears inefficient, it probably evolved in multicellular organisms because the process of exon shuffling facilitated the evolution of new genes in organisms with long introns (Chapter 7). The introns that are spliced out and the region downstream from the cleavage and polyadenylation site are degraded by nuclear exoribonucleases that hydrolyze one base at a time from either the 5′ or 3′ end of an RNA strand.

As mentioned earlier, the 2′,5′-phosphodiester bond in excised introns (see Figure 9-13, ILS) is hydrolyzed by a debranching

enzyme, yielding a linear molecule with unprotected ends that can be attacked by exonucleases.

The predominant nuclear decay pathway for RNAs removed from pre-mRNAs by RNA splicing, and the transcribed RNA downstream from a cleavage/polyadenylation site, is $3' \rightarrow 5'$ hydrolysis by a large protein complex called the **exosome** (Figure 9-24). The exosome consists of a nine-subunit barrel referred to as *Exo-9*, comprised of nine similar polypeptides (gray in Figure 9-24). Although these have sequences homologous to bacterial RNA phosphorylases, the Exo-9 barrel has no ribonuclease activity. Rather, it forms a channel through which the RNA substrate is threaded $3'$ end first to finally reach the tenth subunit in the exosome, RRP44, which has the only ribonuclease activity of any subunit of the complex. The major site of RNA digestion is in the exonuclease active site at the end of the channel (Figure 9-24b, marked "Exo.") Since this active site is within the exosome, RNAs outside of the exosome are protected from digestion by the highly active RRP44 exonuclease. The long channel through the large exosome is thought to hold the RNA in optimal conformation for removal of one nucleotide at a time by the rapid exonuclease activity of the complex, without releasing the substrate RNA. This results in what is known as *processive* degradation; the RNA substrate is digested completely (to a tetranucleotide) without being released from the enzyme.

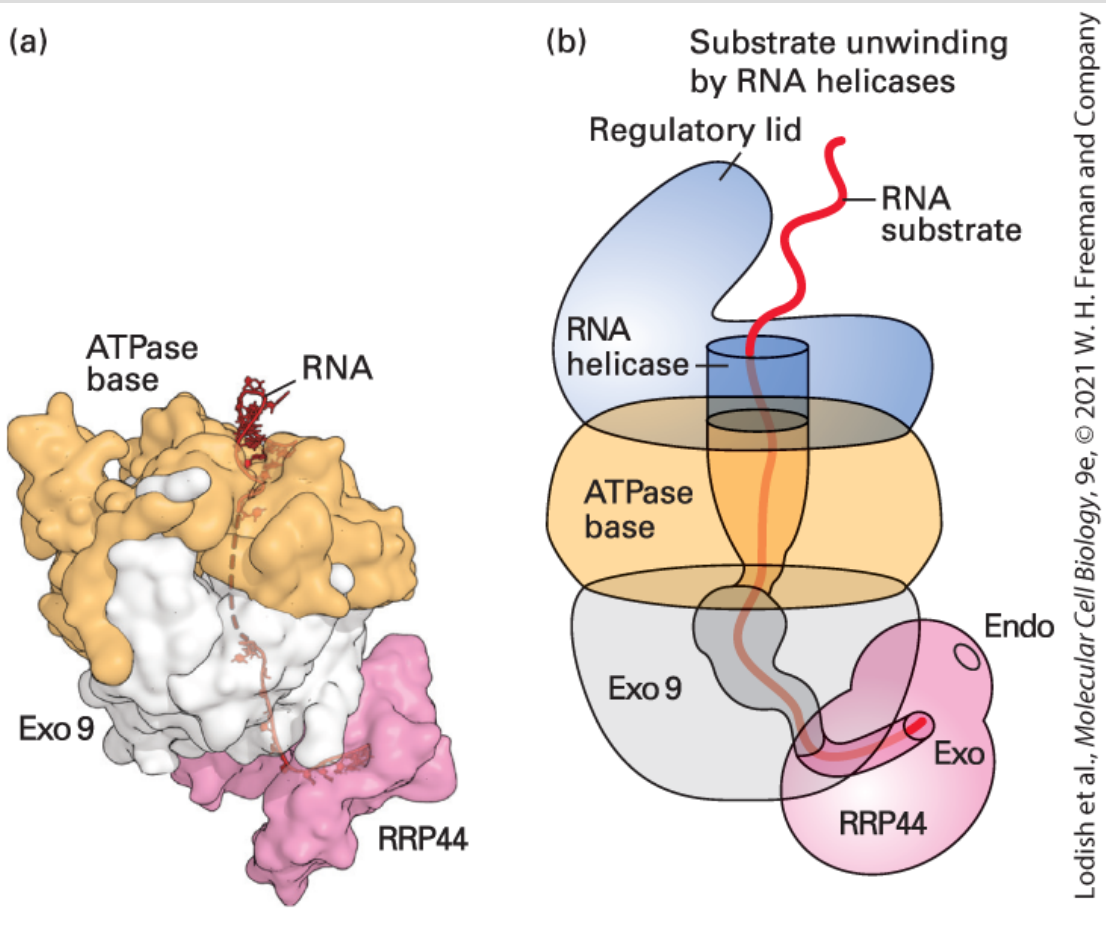


FIGURE 9-24 Structure of the exosome. (a) Structure of the 420 kDa exosome core complex, termed *Exo 10*. It consists of the catalytically inactive Exo 9 complex composed of nine closely related subunits (gray) and the S1 and KH ring (orange). The Exo 9 complex associates with RRP44 (violet), which contains sites of endonuclease (endo) and exonuclease (exo) activity. (b) Schematic representation showing the multisubunit exosome lid, which includes RNA helicases that push RNA substrates through a narrow channel in the lid (blue) that continues into the Exo 10 complex until the 3′ end of the substrate RNA reaches the processive catalytic exonuclease site in RRP44.

[Data from D. Makino, F. Halbach, and E. Conti, 2013, *Nat. Rev. Mol. Cell Biol.* **14**:655.]

In addition to an exonuclease site, RRP44 has an endonuclease site on its surface. This is thought to be utilized when RNA degradation

at the RRP44 exonuclease site is inhibited because of a highly stable RNA structure or chemical damage such as oxidation at the 3′ end of the RNA. In this case, the RRP44 endonuclease site can cleave the substrate RNA 5′ to the block to exonuclease digestion when the blocked substrate RNA diffuses out of the exosome channel. This generates a new 3′ end susceptible to rapid digestion that can be threaded into the exosome, resulting in complete digestion of the RNA.

Since the exonuclease active site is located at the end of a long tunnel in the exosome with dimensions just adequate to accommodate an extended RNA chain, the highly active enzymatic activity does not have access to other nuclear or cytoplasmic RNAs. Rather, substrates must be recognized and loaded into the pore at the top of the exosome before RNA helicases in the lid push the 3′ end of the RNA down the channel until it reaches the exonuclease active site in RRP44 (Figure 9-24b). This regulatory lid on nuclear exosomes is called the *TRAMP complex* and contains the RNA helicase, *Mtr4*. Exosomes in the cytoplasm are associated with a regulatory lid called the *SKI complex* containing the Ski2 RNA helicase. Note that basic design principles of the exosome are similar to those of the major protein degrading multiprotein complex in eukaryotic cells, the proteasome (see Figure 3-32). In both degradative complexes, degradation takes place inside a protein barrel that excludes other potential cellular substrates. The selection of substrates to be rapidly degraded is made by proteins in

a regulatory cap structure that binds one end of the substrate polymer and then, using energy from ATP hydrolysis, pushes it into the digestive cavity as an extended polymer highly susceptible to the degradative enzymes.

In addition to degrading introns, nuclear exosomes appear to degrade pre-mRNAs that have not been properly spliced or polyadenylated, thus providing a quality-control mechanism. It is not yet clear how the exosome recognizes improperly processed pre-mRNAs. But in yeast cells with temperature-sensitive mutant poly(A) polymerase (see Figure 9-15), pre-mRNAs are retained at their sites of transcription in the nucleus at the nonpermissive temperature. These abnormally processed pre-mRNAs are released in cells with a second mutation in a subunit of the exosome found only in nuclear and not in cytoplasmic exosomes (EXOSC10; 100 kD in humans). Also, exosomes are found concentrated at sites of transcription in *Drosophila* polytene chromosomes, where they are associated with RNA polymerase II elongation factors. These results suggest that the exosome participates in an as yet poorly understood quality-control mechanism that recognizes aberrantly processed pre-mRNAs, preventing their export to the cytoplasm and ultimately leading to their degradation.

To avoid being degraded by nuclear exonucleases, nascent transcripts, pre-mRNA processing intermediates, and mature mRNAs in the nucleus must have their ends protected. As discussed

above, the 5′ end of a nascent transcript is protected by addition of the 5′ cap structure as soon as the 5′ end emerges from the polymerase. The 5′ cap is protected because it is bound by a *nuclear cap-binding complex,* which protects it from 5′ exonucleases and also functions in export of mRNA to the cytoplasm. The 3′ end of a nascent transcript lies within the RNA polymerase and thus is inaccessible to exonucleases (see Figure 8-9b). As discussed previously, the free 3′ end generated by cleavage of a pre-mRNA downstream from the poly(A) signal is rapidly polyadenylated by the poly(A) polymerase in a 900 kDa cleavage and polyadenylation complex, and the resulting poly(A) tail is bound by PABPN1 (see Figure 9-15). This tight coupling of cleavage and polyadenylation protects the 3′ end from exonuclease attack.

RNA Processing Solves the Problem of Pervasive Transcription of the Genome in Mammalian Cells

As discussed in Chapter 8, analysis of the location of transcribing RNA polymerase II in metazoan cells revealed the surprising result that the polymerases transcribe in both the downstream direction into coding regions and in the upstream direction away from coding regions at nearly equal frequency from most promoters. This finding was also confirmed by deep sequencing of small RNAs isolated from mammalian cells that revealed low levels of short,

capped RNAs transcribed from both the sense and antisense strands at CpG-island promoters ($\sim$70 percent of mammalian promoters). Indeed, deep sequencing of all cellular RNAs showed that both strands of nearly the entire genome are transcribed, although much of this RNA is present at extremely low concentrations of less than one molecule per cell. This raised the question of how the cell deals with such **pervasive transcription** of the genome.

Sequence analysis of these low-abundance, short RNAs indicates that they are likely prevented from reaching high concentrations by RNA processing and nuclear surveillance of abnormally processed RNAs. Sequencing of RNAs from several cell types revealed that the antisense RNAs have a higher frequency of AAUAAA polyadenylation site sequences than transcripts transcribed in the sense direction into gene coding regions. This is probably because the high AT composition of mammalian DNA ($\sim$60 percent in humans) results in occasional transcription of AAUAAA in antisense transcripts ($\sim 0.6^6 = 4.7$ percent of all six base sequences with an average of 60 percent AT). However, AAUAAA sequences in protein-coding genes occur at much lower frequency because the genes have evolved to have one or a small number of poly(A) sites. Consequently, the enrichment of poly(A) sites in antisense transcripts leads to cleavage of most antisense transcripts within a few kb of their transcription start sites by cleavage/polyadenylation factors, followed by termination of transcription. Most cleaved antisense transcripts are degraded by nuclear exosomes; evidence

for this comes from experiments in which temperature-sensitive mutations in exosome subunits cause a rapid increase in short antisense promoter-proximal RNAs when the mutant cells are shifted to the nonpermissive temperature. As a result, even though a large number of RNA polymerase II molecules transcribe in the "wrong" direction, most of the transcripts generated are degraded before they reach a length of a few kb.

RNA Editing Alters the Sequences of Some Pre-mRNAs

In the mid-1980s, sequencing of numerous cDNA clones and corresponding genomic DNAs from multiple organisms led to the unexpected discovery of another type of pre-mRNA processing. In this type of processing, called **RNA editing**, the sequence of a pre-mRNA is altered; as a result, the sequence of the corresponding mature mRNA differs from the exons encoding it in genomic DNA.

RNA editing is widespread in the mitochondria of protozoans and plants and also in chloroplasts. In the mitochondria of certain pathogenic trypanosomes, more than half the sequence of some mRNAs is altered from the sequence of the corresponding primary transcripts. Additions and deletions of specific numbers of U's follow templates provided by base-paired short "guide" RNAs. These RNAs are encoded by thousands of mini-mitochondrial DNA circles

catenated to many fewer large mitochondrial DNA molecules. The reason for this baroque mechanism for encoding mitochondrial proteins in such protozoans is not clear. But this system does represent a potential target for drugs to inhibit the complex processing enzymes essential to the microbe that do not exist in the cells of their human or other vertebrate hosts.

In multicellular eukaryotes, RNA editing is much rarer, and, thus far, only single-base changes have been observed in mammals. Such minor editing, however, turns out to have significant functional consequences in some cases. An important example of RNA editing in mammals involves the *apoB* gene. This gene encodes two alternative forms of a serum protein central to the uptake and transport of cholesterol. Consequently, it is important in the pathogenic processes that lead to *atherosclerosis,* the arterial disease that is the major cause of death in the developed world. The *apoB* gene expresses both the serum protein apolipoprotein B-100 (apoB-100) in hepatocytes, the major cell type in the liver, and apoB-48, expressed in intestinal epithelial cells. The ~240-kDa apoB-48 corresponds to the N-terminal region of the ~500-kDa apoB-100. As we detail in Chapter 14, both apoB proteins are components of large lipoprotein complexes that transport lipids in the serum. However, only low-density lipoprotein (LDL) complexes, which contain apoB-100 on their surface, deliver cholesterol to body tissues by binding to the LDL receptor present on all cells.

The cell-type-specific expression of the two forms of apoB results from editing of *apoB* pre-mRNA so as to change the nucleotide at position 6666 in the sequence from a C to a U. This alteration, which occurs only in intestinal cells, converts a CAA codon for glutamine to a UAA stop codon, leading to synthesis of the shorter apoB-48 (Figure 9-25). Studies with the partially purified enzyme that performs the post-transcriptional deamination of C_{6666} to U shows that it can recognize and edit an RNA as short as 26 nucleotides with the sequence surrounding C_{6666} in the *apoB* primary transcript, but it does not deaminate C's in RNAs with other sequences.

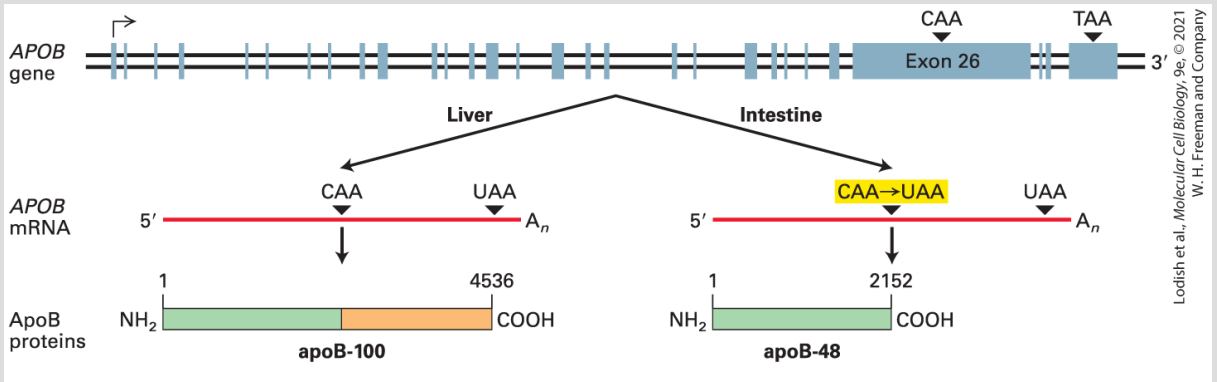


FIGURE 9-25 RNA editing of apoB pre-mRNA. The apoB mRNA produced in the liver has the same sequence as the exons in the primary transcript. This mRNA is translated into apoB-100, which has two functional domains: an N-terminal domain (green) that associates with lipids and a C-terminal domain (orange) that binds to LDL receptors on cell membranes. In the apo-B mRNA produced in the intestine, the CAA codon in exon 26 is edited to a UAA stop codon. As a result, intestinal cells produce apoB-48, which corresponds to the N-terminal domain of apoB-100. See P. Hodges and J. Scott, 1992, *Trends Biochem. Sci.* **17**:77.

KEY CONCEPTS OF SECTION 9.2

Regulation of Pre-mRNA Processing

- Because of alternative splicing of primary transcripts, the use of alternative promoters, and cleavage at different poly(A) sites, different mRNAs may be expressed from the same gene in different cell types or at different developmental stages (see Figure 9-18). This vastly expands the number of distinct proteins encoded in the genomes of multicellular animals and plants compared to eukaryotes that exhibit little alternative RNA splicing such as *Saccharomyces cerevisiae.*
- SR proteins that bind to exonic splicing enhancer sequences present in some exons are critical for defining the splice sites of these exons in the large pre-mRNAs of vertebrates. A network of interactions between SR proteins, snRNPs, and splicing factors forms a cross-exon recognition complex over the short vertebrate exon that specifies correct localization of splice sites (see Figure 9-16).
- Alternative splicing can be regulated by RNA-binding proteins that bind to specific sequences near regulated splice sites. Splicing repressors may sterically block the binding of splicing factors to specific sites in pre-mRNAs or inhibit their function. Splicing activators enhance splicing by interacting with splicing factors, thus promoting their association with a regulated splice site. The RNA sequences bound by splicing repressors are called *intronic* or *exonic splicing silencers*, depending on their location in an intron or exon. RNA sequences bound by splicing activators are called *intronic* or *exonic splicing enhancers*.
- Excised introns and RNA downstream from the cleavage/polyadenylation site are degraded primarily by exosomes, multiprotein complexes that contain an internal $3' \rightarrow 5'$ exonuclease. Exosomes also degrade improperly processed pre-mRNAs.
- Multiple human genetic diseases result from abnormal splicing. A base pair change in a pre-mRNA can create a consensus $5'$ or $3'$ splice site sequence, leading to splicing of an abnormal mRNA that does not encode a functional protein. A base pair change in an exonic splicing enhancer can cause skipping of the exon. Mutations causing specific amino acid changes in a subset of splicing factors cause myelodysplastic syndromes that are precursors to acute myeloid leukemia (AML).
- Group II self-splicing introns in mitochondria and chloroplasts of protozoa, plants, and fungi may represent evolutionary precursors of snRNAs and their ability to act in trans at the short splice site sequences of nuclear pre-mRNAs.
- In RNA editing, the nucleotide sequence of a pre-mRNA is altered in the nucleus. In vertebrates, this process is relatively rare, and only single-base C to U changes have been observed, but those changes can have important consequences by altering the amino acid encoded by an edited codon (see Figure 9-25). In pathogenic trypanosome mitochondria (called *kinetoplasts*) RNA editing using separately transcribed guide

RNAs leads to insertions and base changes comprising more than half of the length of some kinetoplast mRNAs.

9.3 Transport of mRNA Across the Nuclear Envelope

Before an mRNA can be translated into its encoded protein, it must be exported out of the nucleus into the cytoplasm. The nucleus is separated from the cytoplasm by the **nuclear envelope**, a double membrane that is continuous with the endoplasmic reticulum (ER) (see Figure 1-16). Like the plasma membrane surrounding cells, the inner and outer nuclear membranes consist of water-impermeable phospholipid bilayers and multiple associated proteins. Fully processed mRNAs in the nucleus remain bound by hnRNP proteins in complexes referred to as *nuclear mRNPs.* mRNPs and other macromolecules including tRNAs and ribosomal subunits traverse the nuclear envelope through *nuclear pores.* This section will focus on the export of mRNPs through the nuclear pore and the mechanisms that allow some level of regulation of this step. Transport of other cargoes across the nuclear pore is discussed in Chapter 13.

Embedded in the nuclear envelope, **nuclear pore complexes (NPCs)** are cylindrical in shape with a diameter of ~30 nm (see Figure 13-32). Proteins and RNPs larger than ~40–60 kDa must be selectively transported across NPCs with the assistance of soluble transporter proteins. These nuclear transporters bind cargo to be transported across an NPC and also interact reversibly with short hydrophobic FG-repeat regions in random coil polypeptides extending from the

NPC walls into the central channel. As a consequence of these reversible interactions with FG-domains, the transporter and its stably bound cargo can be passed from FG-domain to FG-domain, allowing the complex of transporter and cargo to diffuse down a concentration gradient: the concentration of the transporter-cargo protein complex is high in the nucleus where the complexes form, and low in the cytoplasm where the transporter-cargo complexes dissociate.

Most types of mRNPs are transported through the NPC by the **mRNP exporter**, a heterodimer consisting of a large subunit, called *n*uclear e*x*port *f*actor *1* (NXF1), and a small subunit, *n*uclear e*x*port *t*ransporter *1* (NXT1). NXF1 binds nuclear mRNPs through associations with both RNA and other proteins in the mRNP complex. One of the most important of these is REF (*R*NA *e*xport *f*actor), a component of the exon-junction complexes discussed earlier, which is bound approximately 20 nucleotides 5′ to each exon-exon junction. The NXF1/NXT1 mRNP exporter also associates with SR proteins bound to exonic splicing enhancers. Thus SR proteins associated with exons function to direct both the splicing of pre-mRNAs and the export of fully processed mRNAs through NPCs to the cytoplasm. mRNPs are probably bound along their length by multiple NXF1/NXT1 mRNP exporters, which interact with the FG-domains of FG-nucleoporins to facilitate export of mRNPs through the NPC central channel.

Protein filaments extend from the NPC core scaffold into the nucleoplasm, forming a *nuclear basket* (see Figure 13-32b). Protein filaments also extend into the cytoplasm. These filaments assist in mRNP export. Gle2, an adapter protein that reversibly binds both NXF1 and a protein in the nuclear basket, brings nuclear mRNPs to the pore in preparation for export. A protein in the cytoplasmic filaments of the NPC binds an RNA helicase (Dbp5) that functions in the dissociation of NXF1/NXT1 and other hnRNP proteins from the mRNP as it reaches the cytoplasm.

In a process called *mRNP remodeling,* the proteins associated with an mRNA in the nuclear mRNP complex are exchanged for a different set of proteins as the mRNP is transported through the NPC (Figure 9-26). Some nuclear mRNP proteins, like hnRNPC (see Figure 9-5) dissociate early in transport, remaining in the nucleus to bind to newly synthesized nascent pre-mRNA. Other nuclear mRNP proteins remain with the mRNP complex as it traverses the pore and do not dissociate from the mRNP until the complex reaches the cytoplasm, like hnRNPA1 (see Figure 9-5). Proteins in this category include the NXF1/NXT1 mRNP exporter, cap-binding complex (CBC) bound to the 5′ cap, and PABPN1 bound to the poly(A) tail. They dissociate from the mRNP on the cytoplasmic side of the NPC through the action of the Dbp5 RNA helicase that associates with cytoplasmic NPC filaments, as discussed above. These proteins are then imported back into the nucleus where they can function in the export of another mRNP. In the cytoplasm, the cap-binding

translation initiation factor eIF4E replaces CBC bound to the 5′ cap of nuclear mRNPs. In humans, the nuclear poly(A)-binding protein PABPN1 is replaced with the cytoplasmic poly(A)-binding proteins PABPC1, PABPC2, or PABPC3. Only a single PABP is found in budding yeast, in both the nucleus and the cytoplasm.

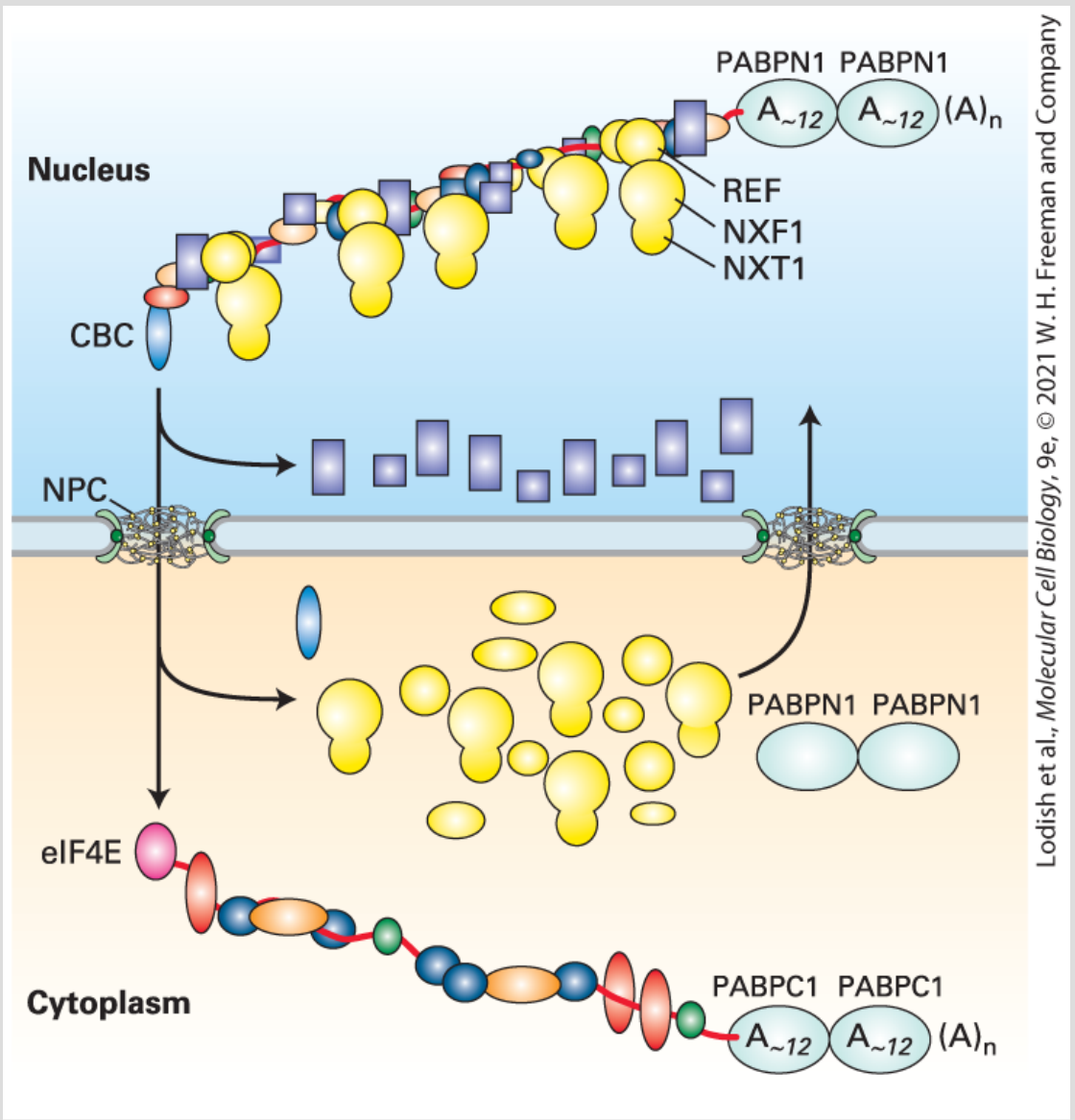


FIGURE 9-26 Remodeling of mRNPs during nuclear export. Some mRNP proteins (purple) dissociate from nuclear mRNP complexes before export through an NPC. Some (yellow) are exported through the NPC associated with the mRNP but

dissociate in the cytoplasm and are shuttled back into the nucleus through an NPC. In the cytoplasm, translation initiation factor eIF4E replaces CBC bound to the 5′ cap and PABPCI, 2, or 3 replaces PABPNI.

SR Proteins Mediate Nuclear Export of mRNA

Studies of *S. cerevisiae* indicate that the direction of mRNP export from the nucleus into the cytoplasm is controlled by phosphorylation and dephosphorylation of mRNP adapter proteins such as REF that assist in the binding of the NXF1/NXT1 exporter to mRNPs. For example, a yeast SR protein (Npl3) functions as an adapter protein that promotes the binding of the yeast mRNP exporter (Figure 9-27). The SR protein initially binds to nascent pre-mRNAs in its phosphorylated form. When 3′ cleavage and polyadenylation are completed, the adapter protein is dephosphorylated by a specific nuclear protein phosphatase essential for mRNP export. Only the dephosphorylated adapter protein can bind the mRNP exporter, thereby coupling mRNP export to correct polyadenylation. This is one form of mRNA *quality control*. If the nascent mRNP is not correctly processed, it is not recognized by the phosphatase that dephosphorylates Npl3. Consequently, it is not bound by the mRNA exporter and not exported from the nucleus. Instead, it is degraded by exosomes, the multiprotein

complexes that degrade unprotected RNAs in the nucleus and cytoplasm (see Figure 9-24).

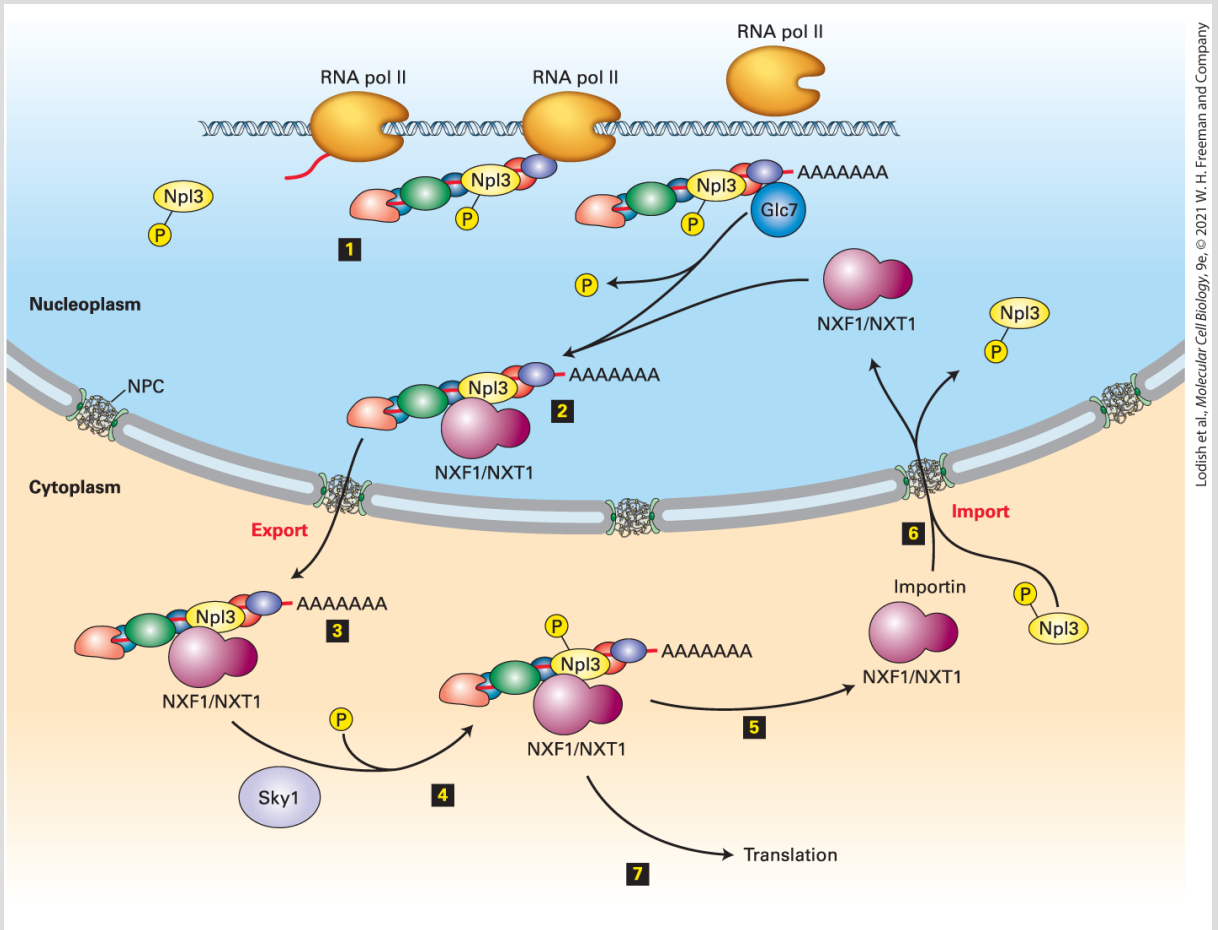


FIGURE 9-27 Reversible phosphorylation of mRNP proteins controls the direction of mRNP nuclear export. Step 1: The yeast SR protein Npl3 binds nascent pre-mRNAs in its phosphorylated form. Step 2: When polyadenylation has occurred successfully, the Glc7 nuclear phosphatase essential for mRNP export dephosphorylates Npl3, promoting the binding of the yeast mRNP exporter, NXF1/NXT1. Step 3: The mRNP exporter allows diffusion of the mRNP complex through the central channel of the nuclear pore complex (NPC). Steps 4 and 5: The cytoplasmic protein kinase Sky1 phosphorylates Npl3 in the cytoplasm, causing dissociation of the mRNP exporter and phosphorylated Npl3, probably through the action of an RNA helicase associated with NPC cytoplasmic filaments. Step 6: The mRNA transporter and phosphorylated Npl3 are transported back into the nucleus through NPCs. Step 7: Transported mRNA is available for translation in the cytoplasm. See

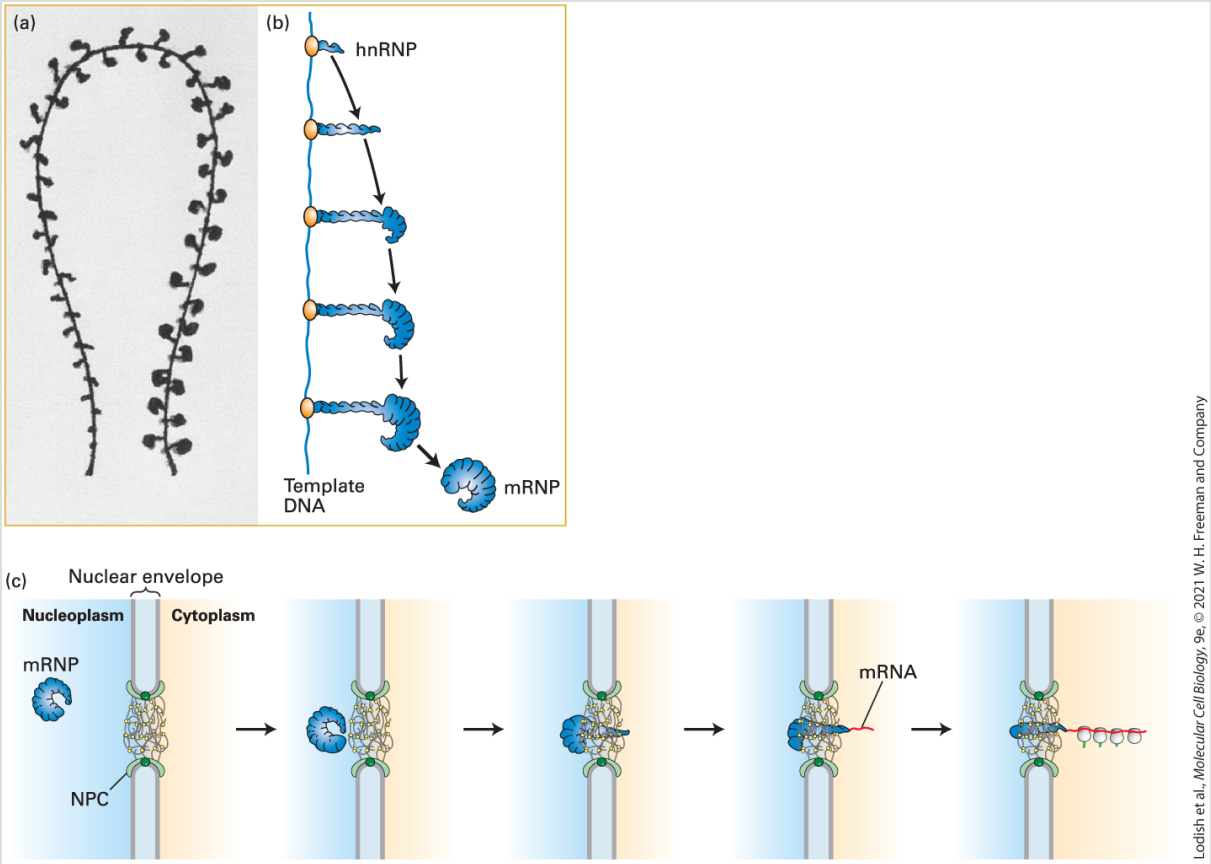
E. Izaurralde, 2004, *Nat. Struct. Mol. Biol.* **11**:210–212. Data from W. Gilbert and C. Guthrie, 2004, *Mol. Cell* **13**:201–212.

Following export to the cytoplasm, the Npl3 SR protein is phosphorylated by a specific protein kinase localized in the cytoplasm. This causes it to dissociate from the mRNP, along with the mRNP exporter. In this way, dephosphorylation of adapter mRNP proteins in the nucleus once RNA processing is complete, and their phosphorylation and resulting dissociation in the cytoplasm drives mRNP nuclear export. This results in a higher concentration of mRNP exporter–mRNP complexes in the nucleus, where they form, and a lower concentration of these complexes in the cytoplasm, where they dissociate. Consequently, the direction of mRNP export may be driven by simple diffusion down a concentration gradient of the transport-competent mRNP exporter–mRNP complex across the NPC, from high concentration in the nucleus to low concentration in the cytoplasm.

Nuclear Export of Balbiani Ring mRNPs

The larval salivary glands of the insect *Chironomous tentans* provide a good model system where assembly of an mRNP and its transport across an NPC can be visualized by electron microscopy. In these larvae, genes in large chromosomal puffs called *Balbiani rings* are abundantly transcribed into nascent pre-mRNAs that associate with hnRNP proteins and are processed into coiled mRNPs with a final

mRNA length of ~75 kb (Figure 9-28a, b). These giant mRNAs encode large glue proteins that adhere the developing larvae to a leaf. After processing of the pre-mRNA in Balbiani ring hnRNPs, the resulting mRNPs move through nuclear pores to the cytoplasm. Electron micrographs of sections of these cells show mRNPs that appear to uncoil during their passage through nuclear pores and then bind to ribosomes as they enter the cytoplasm. This uncoiling is probably a consequence of the remodeling of mRNPs as the result of phosphorylation of mRNP proteins by cytoplasmic kinases and the action of the RNA helicase associated with NPC cytoplasmic filaments, as discussed in the previous section. The observation that mRNPs become associated with ribosomes during transport indicates that the 5′ end leads the way through the nuclear pore complex. Detailed electron microscopic studies of the transport of Balbiani ring mRNPs through nuclear pore complexes led to the model depicted in Figure 9-28c.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 9-28 Formation of heterogeneous ribonucleoprotein particles (hnRNPs) and export of mRNPs through NPCs in *Chironomus tentans* salivary gland cells during metamorphosis. (a) Model of a single chromatin transcription loop and assembly of Balbiani ring (BR) mRNP in *Chironomus tentans*. Nascent RNA transcripts produced from the template DNA rapidly associate with proteins, forming hnRNPs. The gradual increase in size of the hnRNPs reflects the increasing length of RNA transcripts at greater distances from the transcription start site. The model was reconstructed from electron micrographs of serial thin sections of salivary gland cells. (b) Model of the morphology of the BR hnRNP as it grows in length during its transcription. Following processing and cleavage/polyadenylation of the pre-mRNA, the resulting ribonucleoprotein particle, referred to as an *mRNP*, appeared like a microscopic croissant by EM. (c) Model for the transport of BR mRNPs through the nuclear pore complex (NPC) based on electron microscopic studies. Note that the curved mRNPs appear to uncoil as they pass through nuclear pores. As the mRNA enters the cytoplasm, it rapidly associates with ribosomes, indicating that the 5′ end passes through the NPC first. Parts (b) and (c) see B. Daneholt, 1997, *Cell* **88**:585. See also B. Daneholt, 2001, *Proc. Nat'l. Acad. Sci. USA* **98**:7012.

[Part (a) Republished with permission from Elsevier, from C. Erricson et al., 1989, "The Ultrastructure of Upstream and Downstream Regions of an Active Balbiani Ring Gene," *Cell* **56**(4):631–639; permission conveyed through Copyright Clearance Center, Inc.]

Pre-mRNAs Associated with Spliceosomes Are Not Exported from the Nucleus

It is critical that only fully processed, mature mRNAs be exported from the nucleus because translation of incompletely processed pre-mRNAs containing introns would produce defective proteins that might interfere with normal cellular functions. To prevent this, pre-mRNAs associated with snRNPs in spliceosomes are usually prevented from being transported to the cytoplasm.

In one type of experiment demonstrating this restriction, levels of cytoplasmic and nuclear RNAs were determined for RNAs expressed from a derivative of the rabbit β–globin gene containing only the second intron (see Figure 9-2). This modified β–globin gene expressed high levels of fully processed β–globin mRNA in the cytoplasm and only low levels of the unspliced hnRNA in the nucleus, presumably the initial transcript before being subjected to RNA splicing and export to the cytoplasm. However, when the invariant GT at the 5′ end of the globin intron, *or* the invariant AG at the 3′ end of the intron were mutated, splicing of this pre-mRNA was

blocked, and it was retained in the nucleus. Remarkably, when both the 5′ and 3′ splice sites in the pre-mRNA were mutated, splicing did not occur, as expected, but the unspliced mutant β–globin RNA was exported to the cytoplasm. Further studies revealed that transcripts with only one intact splice site were bound by snRNPs to form intermediates in spliceosome assembly that could not complete the splicing process because of the absence of the second splice site. However, RNA expressed from the mutant gene in which both the 5′- and 3′-splice site–invariant bases were mutated was not assembled into spliceosome assembly intermediates with U1 or U2 snRNPs, but as mentioned above, these RNAs were exported. Further studies in yeast have shown that a nuclear protein that associates with a protein in the NPC nuclear basket is required to retain pre-mRNAs in the nucleus that are associated with spliceosome assembly intermediates. If the gene encoding this protein or the nuclear pore basket protein to which it binds is deleted, unspliced pre-mRNAs are exported. Thus intermediates in spliceosome assembly are actively prevented from being exported to the cytoplasm. Another important observation made in these experiments was that the mutant hnRNAs retaining only one 5′ or 3′ splice site that are retained in the nucleus do not build up to high concentration. Consequently, some mechanism exists that causes degradation of the hnRNAs that cannot be spliced. Understanding how such RNAs are selected for degradation is a goal of current research.

Many cases of thalassemia, an inherited disease that results in abnormally low levels of adult globin proteins, result from mutations in adult globin-gene splice sites that decrease the efficiency of splicing and cause the mutant pre-mRNA to be retained in the nuclei of developing red blood cells in the bone marrow (erythrocyte precursors). The resulting unspliced nuclear globin pre-mRNAs are recognized by mechanisms still being elucidated, and degraded by nuclear exosomes (see Figure 9-24). ■

HIV Rev Protein Regulates the Transport of Unspliced Viral mRNAs

As discussed earlier, transport of fully processed nuclear mRNPs containing mature, functional mRNAs from the nucleus to the cytoplasm entails a complex mechanism that is crucial to gene expression (see Figures 9-26 and 9-27). Regulation of this transport could theoretically provide another means of gene control, although it appears to be relatively rare. The only known examples of regulated mRNA export occur during the cellular response to conditions that cause protein denaturation (e.g., heat shock) or during viral infection when virus-induced alterations in nuclear transport maximize viral replication.

One example of the regulation of mRNP export comes from a protein encoded by human immunodeficiency virus (HIV). A

retrovirus, HIV integrates a DNA copy of its RNA genome into the host-cell DNA (see Figure 5-43). The integrated viral DNA, or provirus, contains a single transcription unit, which is transcribed into a single primary transcript by cellular RNA polymerase II. The HIV transcript can be spliced in alternative ways to yield three classes of mRNAs: a 9-kb unspliced mRNA, ~4-kb mRNAs formed by removal of one intron, and ~2-kb mRNAs formed by removal of two or more introns (Figure 9-29). After their synthesis in the host-cell nucleus, all three classes of HIV mRNAs are transported to the cytoplasm and translated into viral proteins; some of the 9-kb unspliced RNA is used as the viral genome in progeny virions that bud from the cell surface.

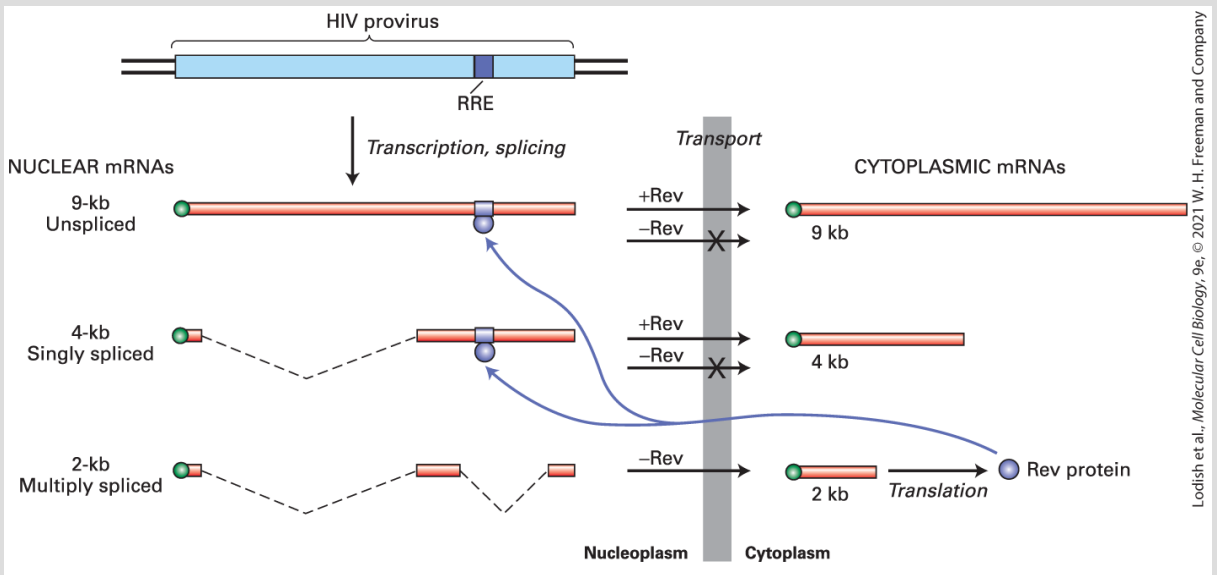


FIGURE 9-29 Transport of HIV mRNAs from the nucleus to the cytoplasm. The HIV genome, which contains several coding regions, is transcribed into a single 9-kb primary transcript. Several ~4-kb mRNAs result from alternative splicing out of any one of several introns (dashed lines) and several ~2-kb mRNAs from splicing out of two or more alternative introns. After transport to the cytoplasm, the various RNA species are translated into

different viral proteins. Rev protein, encoded by a 2-kb mRNA, interacts with the Rev-response element (RRE) in the unspliced and singly spliced mRNAs, stimulating their transport to the cytoplasm. See B. R. Cullen and M. H. Malim, 1991, *Trends Biochem. Sci.* **16**:346.

Since the 9-kb and 4-kb HIV mRNAs contain splice sites, they can be viewed as incompletely spliced pre-RNAs. Association of such incompletely spliced mRNAs with snRNPs in spliceosome assembly intermediates normally blocks their export from the nucleus. Thus HIV, as well as other retroviruses, must have some mechanism for overcoming this block, permitting export of the longer viral mRNAs.

Studies with HIV mutants showed that transport of unspliced 9-kb and singly spliced 4-kb viral mRNAs from the nucleus to the cytoplasm requires the virus-encoded Rev protein. Biochemical experiments show that Rev binds to a specific Rev-response element (RRE) present in HIV RNA. Early in an infection, before any Rev protein is synthesized, only the multiply spliced 2-kb mRNAs can be exported. One of these 2-kb mRNAs encodes Rev, which contains a leucine-rich nuclear export signal that interacts with the transporter Exportin 1. After translation and nuclear import of Rev, Rev binds to the RRE sequence in unspliced and singly spliced HIV RNAs. Rev then associates with Exportin 1, resulting in export of the larger unspliced and singly spliced HIV mRNAs through the nuclear pore complex without the assistance of NXF1/NXT1, the major mRNP exporter for cellular mRNPs. In cells infected with HIV mutants lacking the RRE, unspliced and singly spliced viral mRNAs remain

in the nucleus, confirming that the RRE is required for Rev-mediated stimulation of nuclear export.

KEY CONCEPTS OF SECTION 9.3

Transport of mRNA Across the Nuclear Envelope

- Most mRNPs are exported from the nucleus by a heterodimeric mRNP exporter that interacts with the central channel of the nuclear pore complex (NPC).
- The direction of transport (nucleus to cytoplasm) is controlled by phosphorylation of mRNP adaptor proteins. For example, the yeast Npl3 adapter protein enters nuclei only when it is phorphorylated and associates with nascent pre-mRNAs as they are transcribed (see Figure 9-27). When polyadenylation of the pre-mRNA is complete, a nuclear protein phosphatase (Glc7) removes the phosphates from Npl3, allowing the pre-mRNA to interact with the NXF1/NXT1 mRNP exporter, causing export of the pre-mRNP to the cytoplasm. When the mRNP reaches the cytoplasm, the Sky1 protein kinase restricted to the cytoplasm phosphorylates Npl3, causing dissociation of the mRNP transporter, completing delivery of the mRNP to the cytoplasm.
- The mRNP exporter binds to most mRNAs cooperatively with SR proteins bound to exonic splicing enhancers and with REF associated with exon-junction complexes as well as with additional mRNP proteins.
- Pre-mRNAs bound by a spliceosome normally are not exported from the nucleus, ensuring that only fully processed, functional mRNAs reach the cytoplasm for translation.

9.4 Cytoplasmic Mechanisms of Post-Transcriptional Control

Before proceeding, let's quickly review the steps in gene expression at which control is exerted. We saw in the previous chapter that regulation of transcription initiation and transcription elongation in the promoter-proximal region are the initial mechanisms for controlling the expression of genes in the gene expression pathway of DNA → RNA → protein. In preceding sections of this chapter, we learned that the expression of protein isoforms is controlled by regulating alternative RNA splicing and cleavage and polyadenylation at alternative poly(A) sites. Although nuclear export of fully and correctly processed mRNPs to the cytoplasm is rarely regulated, the export of improperly processed or aberrantly remodeled pre-mRNPs is prevented, and such abnormal transcripts are degraded by the exosome. However, retroviruses, including HIV, have evolved mechanisms that permit pre-mRNAs that retain splice sites to be exported and translated.

In this section, we consider other mechanisms of post-transcriptional control that contribute to regulating the expression of many genes. Most of these mechanisms operate in the cytoplasm, controlling the stability or localization of mRNA or its translation into protein.

The Concentration of an mRNA in the Cytoplasm Is Determined by Its Rate of Synthesis and Its Rate of Degradation

The most stable mRNAs encode proteins required in large numbers, such as the ribosomal proteins; these can accumulate to very high copy number per cell. In contrast, highly unstable mRNAs encoding proteins expressed in short bursts — such as cytokines, secreted proteins that regulate the immune response — rarely achieve such high concentrations even when transcribed, processed, and exported from the nucleus at high rates. We begin by discussing the major pathways that degrade mRNAs. Next we discuss two related mechanisms of gene control that provide powerful new techniques for manipulating the expression of specific genes for experimental and therapeutic purposes.

These two mechanisms are controlled by short, ~22-nucleotide, single-stranded RNAs called **microRNAs (miRNAs)** and **short interfering RNAs (siRNAs)**. Both base-pair with specific target mRNAs, causing their rapid degradation (siRNAs) or inhibiting their translation and inducing a slower form of degradation (miRNAs). Approximately 1900 human miRNAs have been characterized. Most of these are expressed in specific cell types at particular times during embryogenesis and after birth. Many miRNAs can target more than one mRNA. Consequently, these mechanisms contribute significantly to the regulation of gene expression. siRNAs, involved

in the process called *RNA interference*, are an important cellular defense against viral infection and excessive transposition by retrotransposons. We also discuss mechanisms that control the overall rate of protein synthesis, and highly specific regulation of the translation and stability of particular mRNAs by proteins that bind them with high specificity. Finally, we discuss mechanisms that control the localization of mRNAs in the cytoplasm of asymmetric cells so that the encoded protein is translated at sites in the cell where it is needed.

Degradation of mRNAs in the Cytoplasm Occurs by Several Mechanisms

The concentration of an mRNA is a function of both its rate of synthesis and its rate of degradation. For this reason, if two genes are transcribed at the same rate, the steady-state concentration of the corresponding mRNA that is more stable will be higher than the concentration of the other. The stability of an mRNA also determines how rapidly the synthesis of the encoded protein can be shut down. For a stable mRNA, synthesis of the encoded protein persists long after transcription of the gene is repressed. Most bacterial mRNAs are unstable, decaying exponentially with a typical half-life of a few minutes. For this reason, a bacterial cell can rapidly adjust the synthesis of proteins to accommodate changes to the

cellular environment. Most cells in multicellular organisms, on the other hand, exist in a fairly constant environment and carry out a specific set of functions over periods of days to months or even the lifetime of the organism (nerve cells, for example). Accordingly, most mRNAs of multicellular eukaryotes have half-lives of many hours.

Some proteins in eukaryotic cells are required only for short periods and must be expressed in bursts. For example, certain signaling molecules called *cytokines*, which are involved in regulating the immune response of mammals, are synthesized and secreted in short bursts (see Chapter 24). Similarly, many of the transcription factors that regulate the onset of the S phase of the cell cycle, such as Fos and Jun, are synthesized for only brief periods (see Chapter 19). Expression of such proteins occurs in short bursts because transcription of their genes can be rapidly turned on and off, and their mRNAs have unusually short half-lives, on the order of 30 minutes or less, so that concentration of the mRNA falls rapidly when its transcription is repressed.

Cytoplasmic mRNAs are degraded by one of the three pathways shown in Figure 9-30. For most mammalian mRNAs, the *deadenylation-dependent pathway* is followed (Figure 9-30a). The length of the poly(A) tail gradually decreases with time through the action of a deadenylating nuclease complex in the cytoplasm. When it is shortened sufficiently, PABPC molecules can no longer bind and

stabilize the interaction of the 5′ cap and translation initiation factors (see Figure 5-36). The exposed cap then is removed by a heterodymeric decapping enzyme (Dcp1/Dcp2), leaving the unprotected mRNA susceptible to degradation by a 5′ → 3′ exonuclease (Xrn1) (Figure 9-30 step **1**). Removal of the poly(A) tail also makes mRNAs susceptible to degradation by cytoplasmic exosomes containing 3′ → 5′ exonuclease activity (Figure 9-30 step **2**). The 5′ → 3′ exonuclease pathway step **1** predominates in yeast, and the 3′ → 5′ exosome pathway step **2** predominates in mammalian cells.

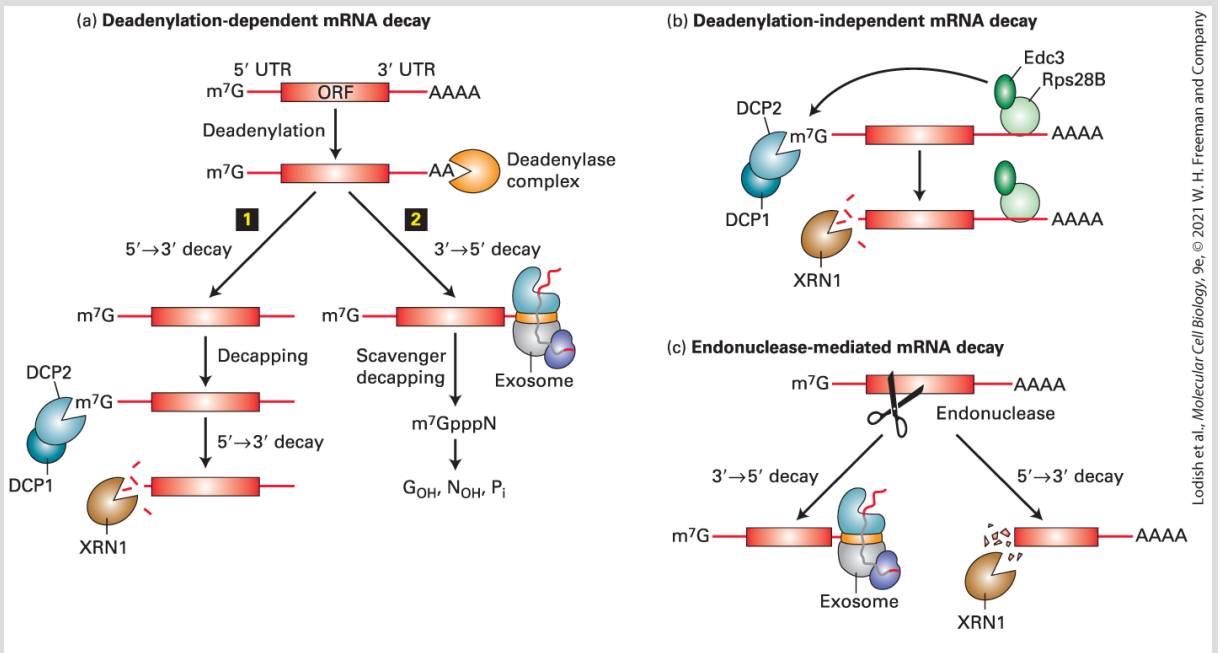


FIGURE 9-30 Pathways for degradation of eukaryotic mRNAs. (a) In the most common pathway of mRNA degradation, the deadenylation-dependent pathway, the poly(A) tail is progressively shortened by a deadenylase complex in the cytoplasm (orange) until it reaches a length of 20 or fewer A residues, at which point the interaction with PABPC1 and the remaining poly(A) is destabilized. This leads to weakened interactions between the 5′ cap and translation initiation factors (see Figures 5-36, step **3**, and Figure 5-39). The

deadenylated mRNA then may either **1** be decapped by the DCP1/DCP2 deadenylation complex and degraded by XRN1, a $5' \rightarrow 3'$ exonuclease, or **2** be degraded by $3' \rightarrow 5'$ exonucleases in cytoplasmic exosomes. (b) Other mRNAs are decapped before they are deadenylated and then degraded by the XRN1 $5' \rightarrow 3'$ exonuclease. (c) Some mRNAs are cleaved internally by an endonuclease and the fragments degraded by a cytoplasmic exosome and the XRN1 exonuclease. See N. L. Garneau, J. Wilusz, and C. J. Wilusz, 2007, *Nat. Rev. Mol. Cell Biol.* **8**:113.

The decapping enzymes and $5' \rightarrow 3'$ exonuclease are concentrated in cytoplasmic **P bodies** (processing bodies), regions of the cytoplasm with unusually high concentrations of RNPs. These regions of highly concentrated proteins are not surrounded by a lipid bilayer membrane. Their formation is dependent on multivalent, low-affinity interactions between simple short repeats of amino acid sequence within intrinsically disordered regions (IDRs) of cytoplasmic RNA-binding proteins associated with some RNAs. These are similar to the transcriptional condensates discussed in Chapter 8 (see Figure 8-38). The process that generates such protein concentrates in both the nucleus and cytoplasm has been called a "liquid-liquid phase separation." This process was first studied for RNA-binding proteins in the cytoplasm that generate P bodies and large cytoplasmic mRNP complexes that are transported to specific regions of the cell by myosin motor proteins that "walk" on actin filaments (see Figure 9-40, and Chapter 17) or by kinesin and dynein motor proteins that walk on microtubules (Chapter 18).

The rate of mRNA deadenylation varies inversely with the frequency of translation initiation for an mRNA: the higher the frequency of

initiation, the slower the rate of deadenylation. This relation probably is due to the reciprocal interactions between translation initiation factors bound at the 5′ cap and PABPC bound to the poly(A) tail. For an mRNA that is translated at a high rate, initiation factors are bound to the cap much of the time, stabilizing the binding of PABPC and thereby protecting the poly(A) tail from deadenylation exonuclease complexes.

AU-Rich Elements

Many short-lived mRNAs in mammalian cells encoding proteins — such as cytokines and transcription factors whose concentrations are changed rapidly — contain multiple, sometimes overlapping copies of the sequence AUUUA in their 3′ untranslated region. These are known as *AU-rich elements*. Specific RNA-binding proteins have been found that both bind to these 3′-AU-rich sequences and also interact with a deadenylating enzyme and with the exosome. This causes rapid deadenylation and subsequent degradation of these mRNAs (Figure 9-30a).

In this mechanism, the rate of mRNA degradation is uncoupled from the frequency of translation. Thus mRNAs containing AU-rich elements can be translated at high frequency yet also be degraded rapidly, allowing the encoded proteins to be expressed in short bursts.

Some mRNAs are degraded primarily by a *deadenylation-independent decapping pathway* (see Figure 9-30b). This is because certain sequences at the 5′ end of an mRNA make the cap sensitive to the decapping enzyme. For these mRNAs, the rate at which they are decapped controls the rate at which they are degraded because once the 5′ cap is removed, the RNA is rapidly hydrolyzed by the $5' \rightarrow 3'$ exonuclease Xrn1.

Finally, as shown in Figure 9-30c, some mRNAs are degraded by an *endonucleolytic pathway* that does not involve decapping or significant deadenylation. One example of this type of pathway is the RNAi pathway discussed below. Each siRNA-RISC complex can degrade multiple targeted RNA molecules. The fragments generated by internal cleavage then are degraded by exonucleases.

MicroRNAs Repress Translation and Induce Degradation of Specific mRNAs

MicroRNAs (miRNAs) were first discovered during analysis of mutations in the *lin-4* and *let-7* genes of the nematode *C. elegans,* which influence development of the organism. Cloning and analysis of wild-type *lin-4* and *let-7* revealed that they do not encode protein but rather RNAs only 21 and 22 nucleotides long, respectively. The RNAs hybridize to the 3′ untranslated regions of specific target mRNAs. For example, the *lin-4* miRNA, which is expressed early in

larval development, hybridizes to the 3′ untranslated regions of both the *lin-14* and *lin-28* mRNAs in the cytoplasm, thereby destabilizing these mRNAs and repressing their translation. Expression of *let-7* miRNA occurs at comparable times during embryogenesis of all bilaterally symmetric animals, including insects and mammals.

miRNA-mediated regulation appears to be widespread in all multicellular plants and animals. In the past few years, small RNAs of 20–26 nucleotides have been isolated, cloned, and sequenced from various tissues of multiple-model organisms. Recent estimates suggest the expression of ~60% of all human genes is regulated by one or more of the 556 well-validated human miRNAs that have been isolated from various tissues. The potential for regulation of multiple mRNAs by one miRNA is high because base pairing between the miRNA and the sequence in the 3′ ends of mRNAs that they regulate need not be perfect (Figure 9-31a). In fact, considerable experimentation with synthetic miRNAs has shown that complementarity between bases two to seven at the 5′ end of an miRNA (called the *seed* sequence) and its target mRNA 3′ untranslated region are most critical for target mRNA selection.

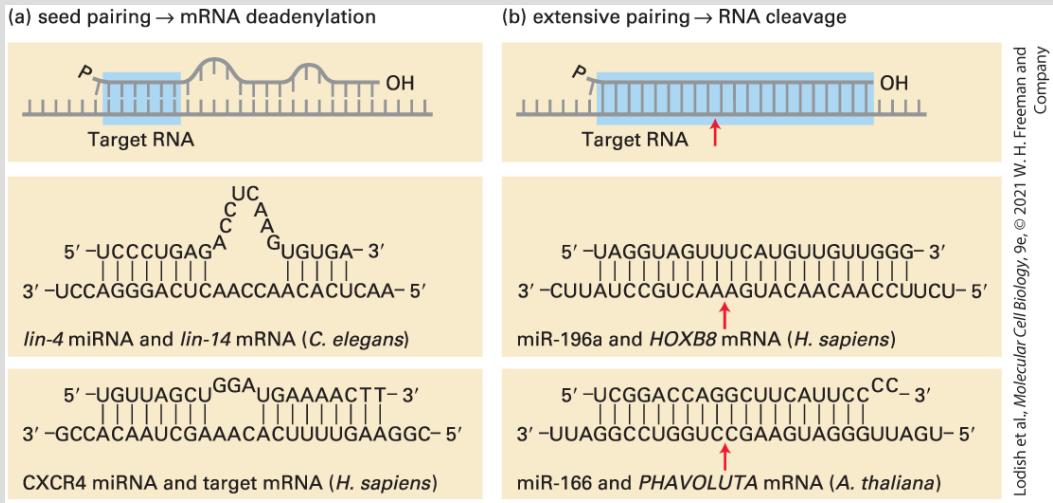
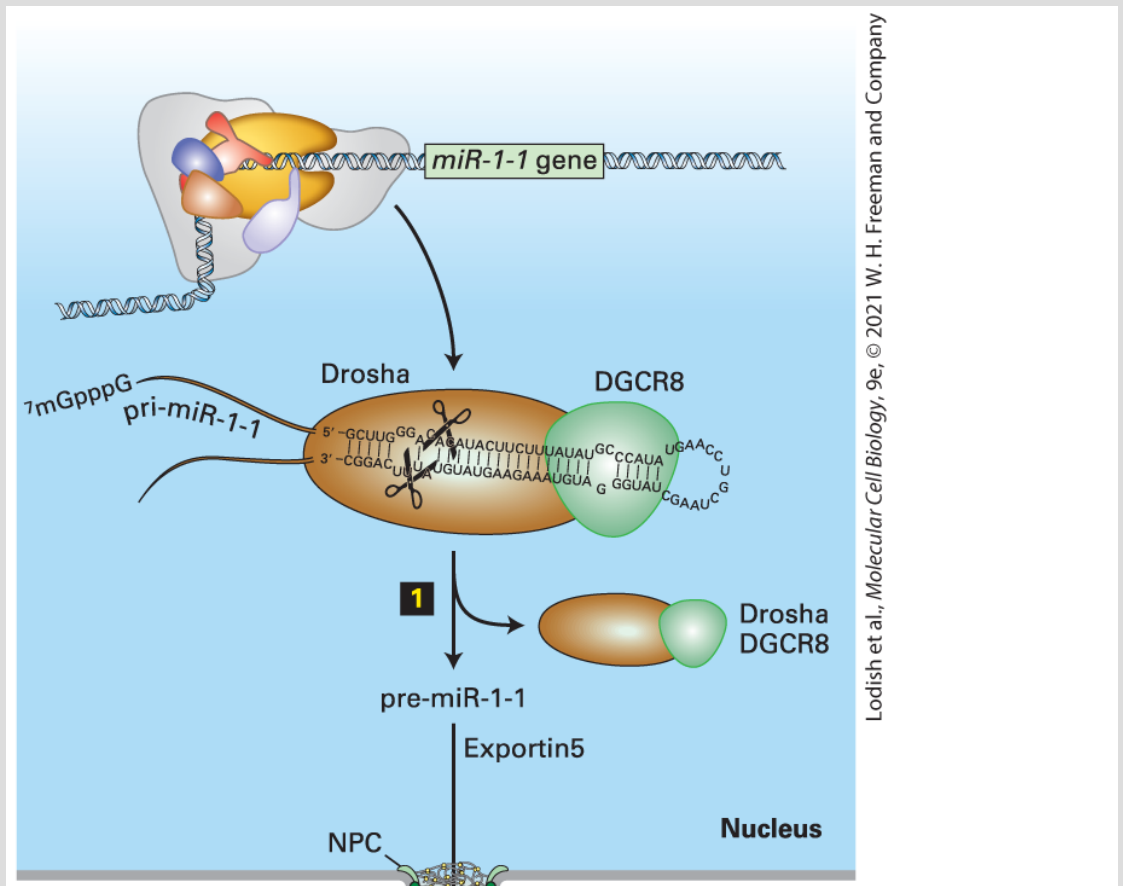


FIGURE 9-31 The extent of base pairing with the miRNA determines the fate of the target RNA. (a) When miRNAs hybridize imperfectly with their target mRNAs, deadenylation is induced and the mRNA is degraded by deadenylation-dependent decay (Figure 9-30a). Nucleotides 2 to 7 of an miRNA (highlighted blue) are the most critical for targeting it to a specific mRNA. The CXCR4 miRNA shown at the bottom is a synthetic oligonucleotide introduced into cells by transfection. (b) miRNAs can also extensively pair to their target mRNAs, causing cleavage of the mRNA at the position indicated by the red arrow, triggering its rapid degradation. See P. D. Zamore and B. Haley, 2005, *Science* **309**:1519.

Most miRNAs are processed from RNA polymerase II transcripts of several hundred to thousands of nucleotides in length called *pri* (for *pri*mary transcript)-miRNAs (Figure 9-32). Pri-miRNAs can contain the sequence of one or more miRNAs. miRNAs are also processed out of some excised introns. Within these long transcripts are sequences that fold into hairpin structures of ~60 nucleotides in length with imperfect base pairing in the stem. A nuclear RNase specific for double-stranded RNA called **Drosha** acts in concert with a nuclear double-stranded RNA-binding protein called *DGCR8* in

humans and cleaves most of the hairpin region out of the long precursor RNA, generating a 60 nucleotide pre-miRNA. Pre-miRNAs are recognized and bound by a specific nuclear export factor, *Exportin5,* which allows the complex to diffuse through the inner channel of the nuclear pore complex. Once in the cytoplasm, a cytoplasmic double-stranded RNA-specific RNase called **Dicer** acts with a cytoplasmic double-stranded RNA-binding protein called *TRBP* in humans (for *T*ar *b*inding *p*rotein) to further process the pre-miRNA into a double-stranded miRNA. The double-stranded miRNA is approximately two turns of an A-form RNA helix in length, with strands 21–23 nucleotides long and with two unpaired 3′ nucleotides at each end.



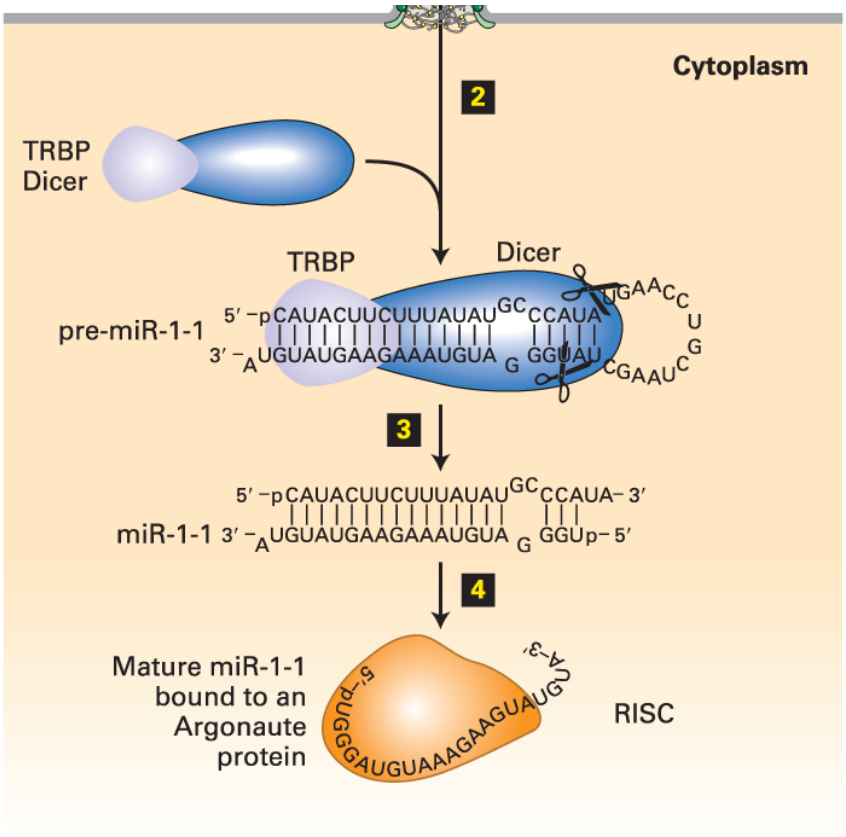


FIGURE 9-32 miRNA processing. This diagram shows transcription and processing of the miR-1-1 miRNA. The primary miRNA transcript (pri-miRNA) is transcribed by RNA polymerase II. The nuclear double-stranded RNA-specific endoribonuclease Drosha, with its partner double-stranded RNA-binding protein DGCR8, make the initial cleavages in the pri-miRNA, generating a ~70-nucleotide pre-miRNA (step 1) that is exported to the cytoplasm by nuclear transporter Exportin5 (step 2). The pre-miRNA is further processed in the cytoplasm to a double-stranded miRNA with two-base single-stranded 3′ ends by Dicer in conjunction with the double-stranded RNA-binding protein TRBP (step 3). Finally, one of the two strands is incorporated into an RISC complex, where it is bound by an Argonaute protein (step 4). See P. D. Zamore and B. Haley, 2005, *Science* **309**:1519.

Finally, one of the two strands is selected for assembly into a mature ***R*NA-*i*nduced *s*ilencing *c*omplex (RISC)** containing a single-stranded mature miRNA bound by a multidomain **Argonaute**

protein, a member of a protein family with a recognizable conserved sequence. Several Argonaute proteins are expressed in some organisms, especially plants, and are found in distinct RISC complexes with different functions. Humans express four Argonaute proteins that can each associate with a miRNA to produce a RISC with a miRNA with a seed sequence that can base-pair to a target mRNA, stimulating its deadenyation and degradation (Figure 9-30a). Ago2 is also able to catalyze mRNA cleavage. The four human Argonaute proteins have partially overlapping functions, demonstrated by experiments that show that knock-out of all four human Argonaute proteins is lethal to human embryonic stem cells, but any one of the four is sufficient for viability.

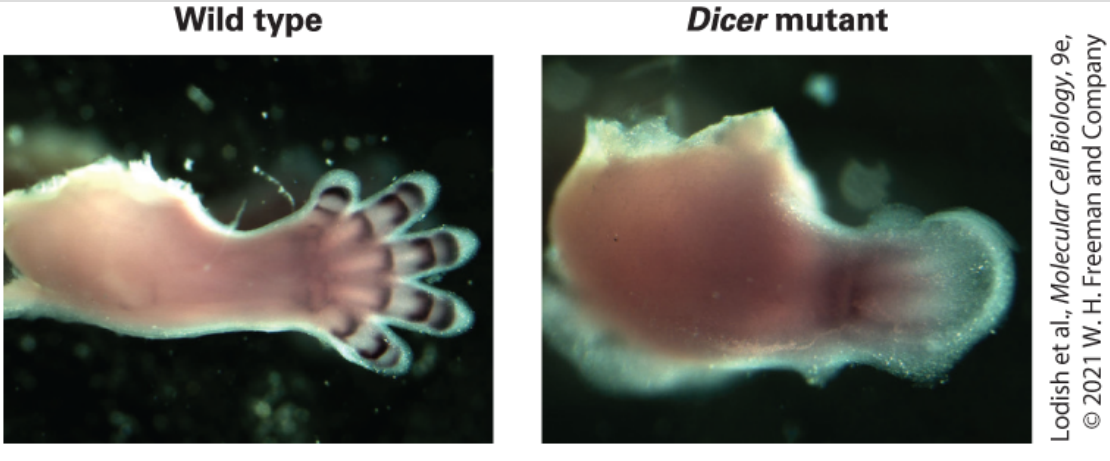
The miRNA-RISC complexes associate with target mRNPs by base pairing between the Argonaute-bound mature miRNA and complementary regions in the 3′ untranslated regions (3′ UTRs) of target mRNAs (see Figure 9-31). Generally, the more RISC complexes are bound to the 3′ UTR of an mRNA, the greater the repression. This allows combinatorial regulation of by separately regulating the transcription of two or more different pri-miRNAs, which are processed to miRNAs required in combination to repress a specific target mRNA.

Approximately 600 different human miRNAs have been observed, most of them expressed only in specific cell types. This has raised the question about the function of miRNAs in differentiation. In one

example, a specific miRNA called *miR-133* is induced when myoblasts differentiate into muscle cells. miR-133 suppresses the translation of PTB, a regulatory splicing factor that functions similarly to Sxl in *Drosophila* (see Figure 9-18). PTB binds to the 3′ splice-site region in the pre-mRNAs of many genes, leading to exon skipping or use of alternative 3′ splice sites. When miR-133 is expressed in differentiating myoblasts, the PTB concentration falls. As a result, alternative isoforms of multiple proteins important for muscle-cell function are expressed in the differentiated cells.

Other examples of miRNA regulation in differentiation in various organisms are being discovered at a rapid pace. Knocking out the *Dicer* gene eliminates the generation of miRNA in mammals. This causes embryonic death early in mouse development. However, when *Dicer* is knocked out only in limb primordia, the influence of miRNA on the development of limbs in the mouse can be observed because limbs are not essential for viability in a laboratory mouse (Figure 9-33). Although all major cell types differentiate and the fundamental aspects of limb patterning are maintained, development is abnormal — demonstrating the importance of miRNAs in regulating the proper level of translation of multiple mRNAs. In effect, miRNAs tune gene expression to the appropriate level for gene function in various cell types. Of the ∼600 human miRNAs, 53 appear to be unique to primates. It seems likely that new miRNAs arose readily during evolution by the duplication of a

pri-miRNA gene followed by mutation of bases encoding the mature miRNA.



EXPERIMENTAL FIGURE 9-33 miRNA function in limb development. Micrographs comparing normal (*left*) and *Dicer* knockout (*right*) limbs of embryonic development day-13 mouse embryos immunostained for the Gd5 protein, a marker of joint formation. *Dicer* is knocked out in developing mouse embryo limb buds by conditional expression of Cre to induce deletion of the *Dicer* gene only in these cells (see Figure 6-40).

[From B. D. Harfe et al., 2005, "The RNaseIII Enzyme Dicer Is Required for Morphogenesis but Not Patterning of the Vertebrate Limb," *Proc. Natl. Acad. Sci. USA* **102**(31):10898–10903.]

Alternative Polyadenylation and mRNA Regulation by miRNAs

The use of alternative polyadenylation sites in mRNA processing, discussed earlier, adds another avenue of regulation of gene expression by miRNAs. For mRNAs expressed from the same gene that use alternative polyadenylation sites, additional miRNA-binding

sites may be located in the mRNA with the longer 3′ exon. As a consequence, mRNAs with the same protein-coding sequence may be regulated differently in alternative cell types depending on the miRNAs expressed in the cells. Consequently, alternative polyadenylation can indirectly regulate mRNAs encoding the same protein as a consequence of miRNA control of translation and mRNA stability.

RNA Interference Induces Degradation of Precisely Complementary mRNAs

RNA interference (RNAi) was discovered unexpectedly during attempts to experimentally manipulate the expression of specific genes. Researchers tried to inhibit the expression of a gene in *C. elegans* by microinjecting a single-stranded, complementary RNA that would hybridize to the encoded mRNA and prevent its translation, a method called *antisense inhibition*. But in control experiments, perfectly base-paired double-stranded RNA a few hundred base pairs long was much more effective at inhibiting expression of the gene than the antisense strand alone (see Figure 6-41). Similar inhibition of gene expression by an introduced double-stranded RNA soon was observed in plants. In each case, the double-stranded RNA induced degradation of all cellular RNAs containing a sequence that was exactly the same as one strand of the double-stranded RNA. Because of the specificity of RNA interference in

targeting mRNAs for destruction, it has become a powerful experimental tool for studying gene function.

Subsequent biochemical studies with extracts of *Drosophila* embryos showed that a long double-stranded RNA that mediates interference is initially processed into a double-stranded short interfering RNA (siRNA). The strands in siRNA contain 21–23 nucleotides hybridized to each other so that the two bases at the 3′ end of each strand are single stranded. Further studies revealed that the cytoplasmic double-stranded RNA-specific ribonuclease that cleaves long double-stranded RNA into siRNAs is the same Dicer enzyme involved in processing pre-miRNAs after their nuclear export to the cytoplasm (see Figure 9-32). This discovery led to the realization that RNA interference and miRNA-mediated translational repression and target mRNA degradation are related processes. In both cases, the mature, short single-stranded RNA, either mature siRNA or mature miRNA, is assembled into a RISC complex in which the short RNAs are bound by an Argonaute protein. When bound to a short RNA that base-pairs extensively with its target RNA, the RISC complex induces target mRNA cleavage. By contrast, a RISC complex associated with a short RNA, which base pairs less extensively (see Figure 9-31a), results in inhibition of translation and a slower form of target mRNA degradation.

In humans, the Argonaute 2 protein is responsible for cleavage of target RNA in an RNAi complex; one domain of the Argonaute

protein is homologous to RNase H enzymes that degrade the RNA of an RNA-DNA hybrid. When the 5′ end of the short RNA of a RISC complex base-pairs precisely with a target mRNA, this domain cleaves the phosphodiester bond of the target RNA across from nucleotides 10 and 11 of the siRNA (see Figure 9-31b). The cleaved RNAs are released and subsequently degraded by cytoplasmic exosomes and the Xrn1 5′ exoribonuclease. If base pairing is less extensive, the Argonaute domain does not cleave or release the target mRNA. Instead, translation is inhibited and the mRNA it is degraded by a different and slower mechanism than the degradation pathway initiated by RISC cleavage of a perfectly complementary target RNA.

When double-stranded RNA is introduced into the cytoplasm of eukaryotic cells, it enters the pathway for assembly of siRNAs into a RISC complex because it is recognized by the cytoplasmic Dicer enzyme and TRBP double-stranded RNA-binding protein that process pre-miRNAs (see Figure 9-32). This process of RNA interference is believed to be an ancient cellular defense against viruses and mobile genetic elements in both plants and animals. Plants with mutations in the genes encoding Dicer and RISC proteins exhibit increased sensitivity to infection by RNA viruses and increased movement of *transposons* within their genomes. The double-stranded RNA intermediates generated during replication of RNA viruses are thought to be recognized by the Dicer ribonuclease, inducing an RNAi response that ultimately degrades viral mRNAs.

During transposition, transposons are inserted into cellular genes in a random orientation, and their transcription from different promoters produces complementary RNAs that can hybridize with each other, initiating the RNAi system that then interferes with the expression of transposon proteins required for additional transpositions.

In plants and *C. elegans* the RNAi response can be induced in all cells of the organism by introduction of double-stranded RNA into just a few cells. Such organism-wide induction requires production of a protein that is homologous to the RNA replicases of RNA viruses. It has been revealed that double-stranded siRNAs are replicated and then transferred to other cells in these organisms. In plants, transfer of siRNAs might occur through plasmodesmata, the cytoplasmic connections between plant cells that traverse the cell walls between them (see Figure 20-38). Organism-wide induction of RNA interference does not occur in *Drosophila* or mammals, presumably because their genomes do not encode RNA-replicase homologs.

In mammalian cells, the introduction of long RNA-RNA duplex molecules into the cytoplasm results in the generalized inhibition of protein synthesis via the PKR pathway, discussed further below. This greatly limits the use of long double-stranded RNAs to experimentally induce an RNAi response against a specific targeted mRNA in mammalian cells as opposed to *Drosophila* cells. Fortunately, researchers discovered that double-stranded siRNAs

21–23 nucleotides in length with two-base 3′-single-stranded regions lead to the generation of single-stranded RNAs that are incorporated into functional siRNA RISC complexes without inducing the generalized inhibition of protein synthesis. This has allowed researchers to use synthetic double-stranded siRNAs to "knock down" the expression of specific genes in human cells as well as in those of other mammals. Also, expression vectors for designed pri-miRNAs can be introduced into cells, leading to the processing of designed siRNAs that hybridize to specific, experimentally controlled target mRNAs, causing their degradation. These methods of **siRNA knockdown** are now widely used in studies of diverse processes, including the RNAi pathway itself!

Cytoplasmic Polyadenylation Promotes Translation of Some mRNAs

In addition to repression by miRNAs, other protein-mediated controls help regulate expression of some genes. Regulatory sequences, or elements, in mRNAs that interact with specific proteins to control translation are generally present in the untranslated region (UTR) at the 3′ or 5′ end of an mRNA. Here we discuss a type of protein-mediated translational control involving 3′ regulatory elements. A different mechanism involving RNA-binding proteins that interact with 5′ regulatory elements is discussed later.

Translation of many eukaryotic mRNAs is regulated by sequence-specific RNA-binding proteins that bind cooperatively to neighboring sites in 3′ UTRs. This allows them to function in a combinatorial manner, similar to the cooperative binding of transcription factors to regulatory sites in an enhancer or promoter region. In most cases studied, translation is repressed by protein binding to 3′ regulatory elements and regulation results from derepression at the appropriate time or place in a cell or developing embryo. The mechanism of such repression is best understood for mRNAs that must undergo *cytoplasmic polyadenylation* before they can be translated.

Cytoplasmic polyadenylation is a critical aspect of gene expression in the early embryo of animals. The egg cells (oocytes) of multicellular animals contain many mRNAs, encoding numerous different proteins that are not translated until after the egg is fertilized by a sperm cell. Some of these "stored" mRNAs have a short poly(A) tail, consisting of only ∼20–40 A residues, to which just a few molecules of cytoplasmic poly(A)-binding protein (PABPCI) can bind. Multiple PABPCI molecules bound to the long poly(A) tail of an mRNA interact with the eIF4G initiation factor, thereby stabilizing the interaction of the mRNA 5′ cap with eIF4E, which is required for translation initiation (see Figure 5-36). Because this stabilization cannot occur with mRNAs that have short poly(A) tails, such mRNAs stored in oocytes are not translated efficiently. At the appropriate time during oocyte maturation or after fertilization of

an egg cell, usually in response to an external signal, approximately 150 A residues are added to the short poly(A) tails on these mRNAs in the cytoplasm, stimulating their translation.

Studies with mRNAs stored in *Xenopus* oocytes have helped elucidate the mechanism of this type of translational control. Experiments in which short-tailed mRNAs are injected into oocytes have shown that two sequences in their 3′ UTR are required for their polyadenylation in the cytoplasm: the AAUAAA poly(A) signal that is also required for the nuclear polyadenylation of pre-mRNAs, and one or more copies of an upstream U-rich *cytoplasmic polyadenylation element (CPE)* (Figure 9-34). This regulatory element is bound by a highly conserved *CPE-binding protein (CPEB)* that contains an RRM domain and a zinc-finger domain.

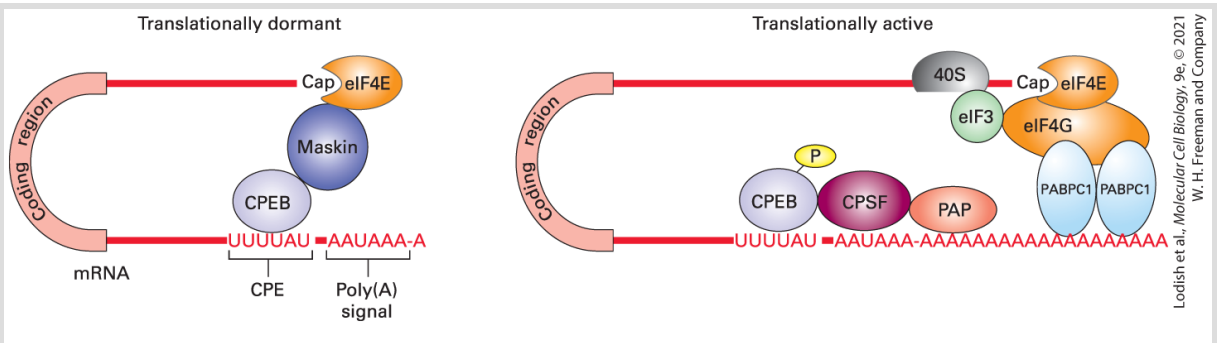


FIGURE 9-34 Model for control of cytoplasmic polyadenylation and translation initiation. *Left:* In immature oocytes, mRNAs containing the U-rich cytoplasmic polyadenylation element (CPE) have short poly(A) tails and are translationally dormant. CPE-binding protein (CPEB) mediates repression of translation through the interactions depicted, which prevent assembly of an initiation complex at the 5′ end of the mRNA. *Right:* Hormone stimulation of oocytes activates a protein kinase that phosphorylates CPEB, causing it to release Maskin. The cleavage and polyadenylation specificity factor (CPSF) then bind to the poly(A) site, interacting with both bound CPEB and the cytoplasmic form of poly(A) polymerase (PAP).

After the poly(A) tail is lengthened, multiple copies of the cytoplasmic poly(A)-binding protein I (PABPC1) can bind to it and interact with eIF4G, which functions with other initiation factors to bind the 40S ribosome subunit and initiate translation (translationally active). See R. Mendez and J. D. Richter, 2001, *Nat. Rev. Mol. Cell Biol.* **2**:521.

In the absence of a stimulatory signal, CPEB bound to the U-rich CPE interacts with the protein Maskin, which in turn binds to eIF4E associated with the mRNA 5′ cap (Figure 9-34, *left*). As a result, eIF4E cannot interact with other initiation factors and the 40S ribosomal subunit, so translation initiation is blocked. During oocyte maturation, a specific CPEB serine is phosphorylated, causing Maskin to dissociate from the complex. This allows cytoplasmic forms of the cleavage and polyadenylation specificity factor (CPSF) and poly(A) polymerase to bind to the mRNA cooperatively with CPEB. Once the poly(A) polymerase catalyzes the addition of A residues, PABPCI can bind to the lengthened poly(A) tail, leading to the stabilized interaction of all the factors needed to initiate translation (Figure 9-34, *right*; see also Figure 5-36). In the case of *Xenopus* oocyte maturation, the protein kinase that phosphorylates CPEB is activated in response to the hormone progesterone. Thus timing of the translation of stored mRNAs encoding proteins needed for oocyte maturation is regulated by this external signal.

Considerable evidence indicates that a similar mechanism of translational control plays a role in learning and memory. In the central nervous system, the axons from a thousand or so neurons can make connections (synapses) with the dendrites of a single

postsynaptic neuron (see Figure 23-3). When one of these axons is stimulated, the postsynaptic neuron "remembers" which one of these thousands of synapses was stimulated. The next time that synapse is stimulated, the strength of the response triggered in the postsynaptic cell differs from the first time. This change in response has been shown to result largely from the translational activation of mRNAs stored in the region of the synapse in the postsynaptic neuron, leading to the local synthesis of new proteins that increase the size and alter the neurophysiological characteristics of the synapse. The finding that CPEB is present in neuronal dendrites has led to the proposal that cytoplasmic polyadenylation stimulates translation of specific mRNAs in dendrites, much as it does in oocytes. In this case, presumably, synaptic activity (rather than a hormone) is the signal that induces local phosphorylation of CPEB and subsequent activation of translation in one specific synapse of the neuron.

Protein Synthesis Can Be Globally Regulated

Like proteins involved in other processes, translation initiation factors and ribosomal proteins can be regulated by post-translational modifications such as phosphorylation. Such mechanisms affect the translation rate of most mRNAs and hence the overall rate of cellular protein synthesis. The *TOR pathway*,

which represses translation when the supply of amino acids is low as sensed through communication between lysosomes and the cytosol, is discussed in detail in Chapter 21. Other mechanisms that control the global rate of cellular translation include mechanisms that control the activity of translation initiation factor *eIF2*. Here we will focus on the mechanisms that can regulate the initiation of translation.

Global regulation of protein synthesis by regulating translation initiation is mediated by eIF2 kinases. Figure 5-36 summarizes the steps in translation initiation. Translation initiation factor eIF2 brings the charged initiator tRNA to the small ribosome subunit P site. eIF2 is a *trimeric G protein* and exists in either a GTP-bound or a GDP-bound conformation. Only the GTP-bound form of eIF2 is able to bind the charged initiator tRNA and associate with the small ribosomal subunit. The small subunit with bound initiation factors and charged initiator tRNA then interacts with the eIF4 complex bound to the 5′ cap of an mRNA via its eIF4E subunit. The small ribosomal subunit then scans down the mRNA in the 3′ direction until it reaches an AUG initiation codon that can base-pair with the initiator tRNA in its P site. When this occurs, the GTP bound by eIF2 is hydrolyzed to GDP and the resulting eIF2·GDP complex is released. GTP hydrolysis results in an irreversible “proofreading” step that prepares the small ribosomal subunit to associate with the large subunit only when an initiator tRNA is properly bound in the P site and is properly base-paired with the AUG start codon. Before

eIF2 can participate in another round of initiation, its bound GDP must be replaced with a GTP. This process is catalyzed by the translation initiation factor eIF2B, a guanine nucleotide exchange factor (GEF) specific for eIF2.

A mechanism for inhibiting general protein synthesis in stressed cells involves phosphorylation of the eIF2α subunit at a specific serine. Phosphorylation at this site does not interfere with eIF2 function in protein synthesis directly. Rather, phosphorylated eIF2 has very high affinity for the eIF2 guanine nucleotide exchange factor, *eIF2B*, which cannot release the phosphorylated eIF2 and consequently is blocked from catalyzing GTP exchange by additional eIF2 factors. Since there is an excess of eIF2 over eIF2B, phosphorylation of a fraction of eIF2 results in inhibition of all the cellular eIF2B. The remaining eIF2 accumulates in its GDP-bound form, which cannot participate in protein synthesis, thereby inhibiting nearly all protein synthesis in the cell.

However, some mRNAs have 5′ regions that allow translation initiation at the low eIF2-GTP concentration that results from eIF2 phosphorylation. These mRNAs include those for chaperone proteins that function to refold cellular proteins denatured as the result of cellular stress, additional proteins that help the cell to cope with stress, and transcription factors that activate transcription of the genes encoding these stress-induced proteins.

Human cells contain four eIF2 kinases that phosphorylate the same inhibitory eIF2 α serine. Each of these is regulated by a different type of cellular stress, inhibiting protein synthesis and allowing cells to divert the large fraction of cellular resources usually devoted to protein synthesis in growing cells for use in responding to the stress:

- The *GCN2* (general control non-derepressible 2) eIF2-kinase is activated by binding uncharged tRNAs. The concentration of uncharged tRNAs increases when cells are starved for amino acids, activating GCN2 eIF2-kinase activity and greatly inhibiting protein synthesis.
- *PEK* (pancreatic eIF2 kinase) is activated when proteins translocated into the endoplasmic reticulum (ER) do not fold properly because of abnormalities in the ER lumen environment. Inducers include abnormal carbohydrate concentrations, because this inhibits the glycosylation of many ER proteins. Inactivating mutations in an ER chaperone required for proper folding of many ER proteins (Chapters 13 and 14) also results in PEK activation.
- *Heme-regulated inhibitor (HRI)* is another eIF2 kinase activated in developing red blood cells when the supply of the heme prosthetic group is too low to accommodate the rate of globin protein synthesis. This negative feedback loop lowers the rate of globin protein synthesis until it matches the rate of heme synthesis. HRI is also activated in other types of cells in response to oxidative stress or heat shock.

- *Protein kinase RNA activated (PKR)* is activated by double-stranded RNAs longer than ∼30 base pairs. Under normal circumstances in mammalian cells, such double-stranded RNAs are produced only during a viral infection. Long regions of double-stranded RNA are generated in replication intermediates of RNA viruses or from hybridization of complementary regions of RNA transcribed from both strands of DNA virus genomes. Inhibition of protein synthesis prevents the production of progeny virions, protecting neighboring cells from infection. Interestingly, adenoviruses evolved a defense against PKR: they express prodigious amounts of an ∼160-nucleotide virus-associated (VA) RNA with long double-stranded hairpin regions. VA RNA is transcribed by RNA polymerase III and exported from the nucleus by Exportin5, the exportin for pre-miRNAs (see Figure 9-32). VA RNA binds to PKR with high affinity, inhibiting its protein kinase activity and preventing the inhibition of protein synthesis observed in cells infected with a mutant adenovirus from which the VA gene was deleted.

Sequence-Specific RNA-Binding Proteins Control Specific mRNA Translation

In contrast to global mRNA regulation, mechanisms have also evolved for controlling the translation of certain specific mRNAs. This is usually done by sequence-specific RNA-binding proteins that

bind to a particular sequence or RNA structure in the mRNA. When binding is in the 5′ untranslated region (5′ UTR) of an mRNA, the 40S ribosomal subunit's ability to scan to the first initiation codon is blocked, inhibiting translation initiation. Binding in other regions can either promote or inhibit mRNA degradation.

Control of intracellular iron concentration is an elegant example of the protein-mediated regulation of the translation of one mRNA and the degradation of another. This is accomplished by ***Iron Response Element Binding Proteins* (IRE-BPs)**, including the bifunctional protein aconitase 1 (encoded by the ACO1 gene, alias *IRE-BP1*) and its paralog *IRE-BP2* (encoded by the IREB2 gene). Precise regulation of cellular iron ion concentration is critical to the cell. Multiple enzymes and proteins contain Fe^{2+} as a cofactor, such as enzymes of the TCA cycle (see Figure 12-16) and electron-carrying proteins involved in the generation of ATP by mitochondria and chloroplasts (Chapter 12). On the other hand, excess Fe^{2+} generates free radicals that react with and damage cellular macromolecules. When intracellular iron stores are low, a dual-control system operates to increase the level of cellular iron; when iron is in excess, the system operates to store the iron in an inert form that prevents accumulation of toxic levels of free Fe^{2+} ions.

IRE-BPs are 4Fe-4S cluster-containing proteins that act as cellular iron sensors. When the cytosolic Fe^{2+} concentration is too low, the iron in the Fe-S clusters dissociates from the protein and the IRE-BP

becomes an active RNA-binding protein that binds to a specific nucleotide sequence within stem-loop secondary structures in the untranslated regions of the ferritin and transferrin receptor mRNAs. These *I*ron *R*esponse *E*lements (IREs) in these mRNAs are located in the 5′ untranslated region of the ferritin mRNAs (Figure 9-35a) and in the 3′ untranslated sequence in the transferrin receptor (TfR) mRNA. Ferritin is composed of 24 subunits of a ferritin heavy chain and an equal number of ferritin light chains. They form a large intracellular protein complex that binds and stores excess cellular iron. When an IRE-BP is in its RNA-binding conformation at low cellular Fe^{2+} concentration, it recognizes and binds five specific bases in the IRE loop and the duplex nature of the stem in the mRNA for both heavy and light ferritin chains (Figure 9-35a). The bound IRE-BP blocks the small ribosomal subunit from scanning for the AUG start codon (see Figure 5-36), thereby inhibiting translation initiation. This causes a decrease in ferritin protein production and results in less iron being stored within the ferritin complex, making more iron available to iron-requiring enzymes. At high iron concentrations, the IRE-BP is in a conformation where it cannot bind the IREs in the 5′ untranslated region of the ferritin mRNAs. Consequently, ferritin translation is not inhibited, and newly synthesized ferritin can store free Fe^{2+}, thus prevent its accumulation to harmful levels within the cell.

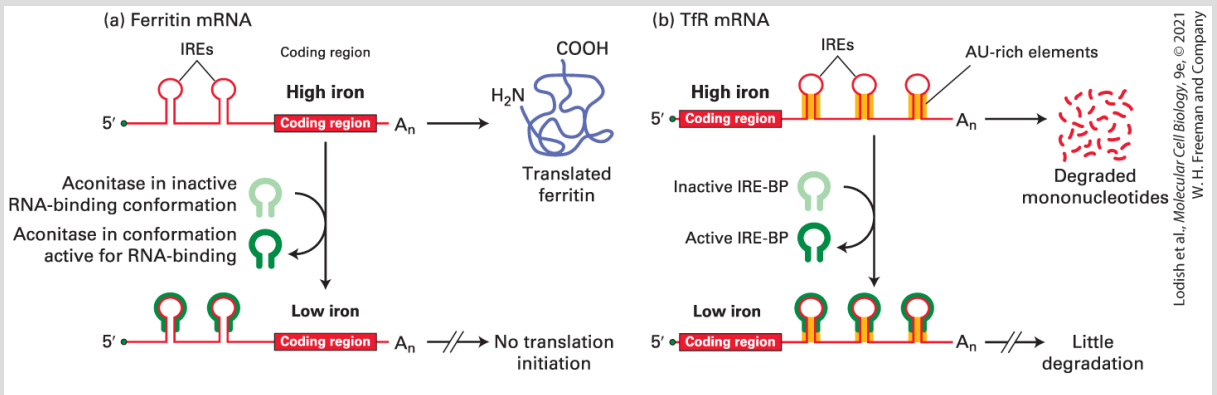


FIGURE 9-35 Iron-dependent regulation of mRNA translation and degradation. The iron response element–binding protein (IRE-BP) controls (a) translation of ferritin mRNA and (b) degradation of transferrin-receptor (TfR) mRNA. At low intracellular iron concentrations IRE-BP binds to iron-response elements (IREs) in the 5′ or 3′ untranslated region of these mRNAs. At high iron concentrations, IRE-BP undergoes a conformational change and cannot bind either mRNA. The dual control by IRE-BP precisely regulates the level of free iron ions within cells. See the text for discussion.

The second mRNA regulated by IRE-BPs is regulated differently: it encodes the transferrin receptor (TfR). In vertebrates, iron is carried through the circulatory system bound to a protein called *transferrin*. After binding to the transferrin receptor (TfR) in the plasma membrane of most cell types, the transferrin-iron complex is brought into cells by receptor-mediated endocytosis (see Chapter 14; Figure 14-31). The 3′-untranslated region of TfR mRNA contains multiple IREs with AU-rich element destabilizing sequences interspersed between the stem-loop structures (see above). At high Fe^{2+} concentrations, when the IRE-BP is unable to bind to the IRE, these AU-rich elements promote degradation of TfR mRNA by the same mechanism that leads to rapid degradation of other short-lived mRNAs with AU-rich elements, as described previously. The

resulting decrease in production of the transferrin receptor quickly reduces iron import, thus protecting the cell from excess iron. At low Fe^{2+}, concentrations, however, the IRE-BP binds to the 3′ IREs in TfR mRNA and blocks access of proteins to the destabilizing AU-rich elements. As a result, TfR mRNA levels increase, leading to increased production of the transferrin receptor protein and more iron transport into the cell. As a consequence of these two mechanisms of post-translational control, regulation of translation of the ferritin mRNAs and regulation of stability of the transferrin mRNA, the intracellular Fe^{2+} concentration is held at the concentration required for binding of Fe^{2+} to iron-requiring enzymes and is prevented from reaching higher concentrations that would be toxic.

Other regulated RNA-binding proteins may also function to control mRNA translation or degradation. For example, a heme-sensitive RNA-binding protein controls translation of the mRNA encoding aminolevulinate (ALA) synthase, a key enzyme in the synthesis of heme. Similarly, in vitro studies have shown that the mRNA encoding the milk protein casein is stabilized by the hormone prolactin and rapidly degraded in its absence.

Surveillance Mechanisms Prevent Translation of Improperly Processed mRNAs

Translation of an improperly processed mRNA could lead to production of a nonfunctional or abnormally functioning protein. This effect is equivalent to that resulting from dominant-negative mutations, discussed in Chapter 6 (see Figure 6-2). Several mechanisms collectively termed **mRNA surveillance** help cells avoid the translation of improperly processed mRNA molecules. We have previously mentioned two such surveillance mechanisms: the recognition of improperly processed pre-mRNAs in the nucleus and their degradation by the nuclear exosome, and the general restriction against nuclear export of incompletely spliced pre-mRNAs that remain associated with a spliceosome assembly intermediate.

Nonsense-Mediated Decay (NMD)

This mechanism of mRNA surveillance causes degradation of mRNAs in which one or more exons have been incorrectly spliced. Such incorrect splicing will often alter the open reading frame of the mRNA 3′ to the improper exon junction, resulting in introduction of an out-of-frame missense mutation and an incorrect stop codon. For nearly all properly spliced mRNAs, the stop codon is in the last exon. The process of nonsense-mediated decay results in the rapid degradation of mRNAs with stop codons that occur before the last splice junction in the mRNA, since in most cases such mRNAs arise from errors in RNA splicing. However, NMD can also result from a mutation creating a stop codon within a gene or a frame-shifting

deletion or insertion. NMD was initially discovered during the study of patients with β^0-thalasemia, who produce a low level of β-globin protein associated with a low level of β-globin mRNA (Figure 9-36).

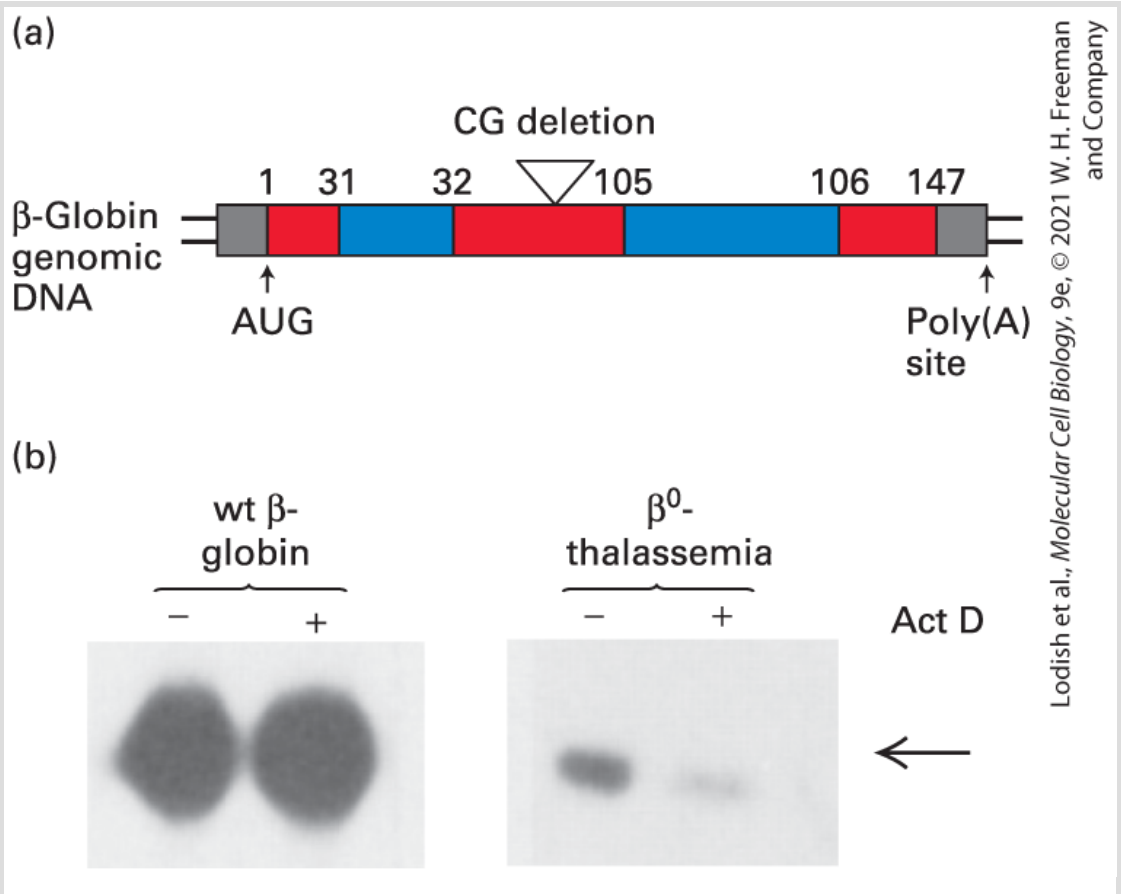


FIGURE 9-36 Discovery of nonsense-mediated mRNA decay (NMD). (a) Patients with β^0-thalassemia express very low levels of β-globin mRNA. A common cause of this syndrome is a single-base-pair deletion in exon 1 or exon 2 of the β-globin gene. Ribosomes translating the mutant mRNA read out of frame following the deletion and encounter a stop codon in the wrong reading frame before they translate across the last exon junction in the mRNA. Consequently, they do not displace an exon-junction complex (EJC) from the mRNA. Cytoplasmic proteins associate with the EJC and induce degradation of the mRNA. (b) Bone marrow was obtained from a patient with a wild-type β-globin gene and from a patient with β^0-thalassemia. RNA was isolated from the bone marrow cells shortly after collection, or 30 minutes after incubation in media with Actinomycin D, a drug that inhibits transcription. The

amount of β^0-globin RNA was measured using the S1-nuclease protection method (arrow). The patient with β^0-thalassemia had much less β-globin mRNA than the patient with a wild-type β-globin gene (−Act D). The mutant β-globin mRNA decayed rapidly when transcription was inhibited (+Act D), whereas the wild-type β-globin mRNA remained stable.

[Republished with permission from Elsevier, from L. E. Maquat et al., 1981, "Unstable β-Globin mRNA in mRNA-Deficient β^0 Thalassemia," *Cell* **27**(3):543–553; permission conveyed through Copyright Clearance Center, Inc.]

A search for possible molecular signals that might indicate the positions of splice junctions in a processed mRNA led to the discovery of exon-junction complexes. As noted already, these complexes of several proteins (including a core complex of eIF4AIII, MLN51, MAGOH, and Y14 proteins and peripheral factors including REF), bind ∼20 nucleotides 5′ to an exon-exon junction following RNA splicing, and stimulate export of mRNPs from the nucleus by interacting with the mRNA exporter, NXT1/NXF1 (see Figure 9-26). Analysis of yeast mutants indicated that one of the peripheral proteins associated with exon-junction complexes in the nucleus, UPF3, functions in nonsense-mediated decay (see Figure 9-37, step **1**). UPF3 remains associated with the EJC as it is exported through a nuclear pore complex, and associates with the cytoplasmic NMD factor UPF2 (see Figure 9-37, step **2**). During translation termination, eukaryotic release factors 1 and 3 (eRF1 and eRF3) associate with ribosomes with a stop codon in the A site (see Figure 5-38). NMD factors UPF1 and SMG1, a protein kinase, associate with eRF1 and eRF3 forming a complex called the *SURF* complex (step **3**) associated with the terminating ribosome. The SURF complex

associates, via an interaction between UPF1 and UPF2, with any UPF2 that remains associated with EJC complexes on the mRNA that were not displaced from the mRNA by a translating ribosome (step **4**). This leads to the phosphorylation of UPF1 by SMG1, which promotes the release of eRF1 and eRF3 and the binding of NMG factor SMG7. Finally, SMG7 causes the mRNP to associate with P bodies, inhibiting translation. UPF2 then binds a P body–associated deadenylase complex that rapidly removes the poly(A) tail from an associated mRNA, leading to its decapping and degradation by the P body–associated $5' \rightarrow 3'$ exonuclease Xrn1 (see Figure 9-30a).

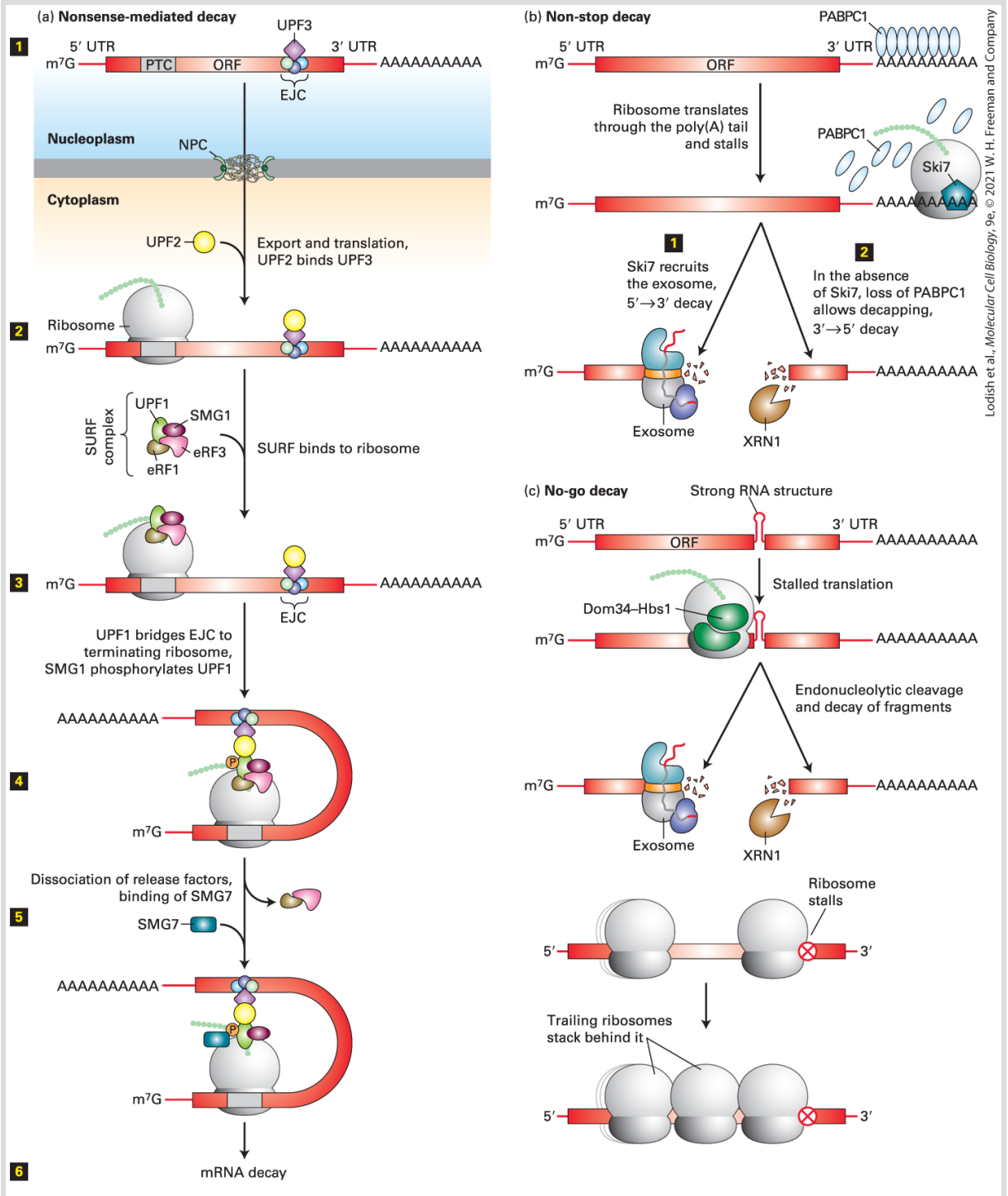


FIGURE 9-37 (a) Nonsence Mediated Decay (NMD). Step **1**: An abnormal mRNA with a premature termination codon (PTC) is produced in the nucleus as the result of abnormal RNA splicing. An exon junction complex (EJC) is left at a splice junction following RNA-splicing and associates with NMD factor UPF3. Step **2**: After export to the cytoplasm, NMD factor UPF2 associates with UPF3 bound to the EJC. A ribosome also initiates translation

and reaches the PTC. Step 3: Eukaryotic release factors 1 (eRF1) and 3 (eRF3) (see Figure 5-38) associate with the ribosome stopped at the termination codon and associate with UPF1 and protein kinase SMG1, forming a "SURF complex" associated with the terminating ribosome. Step 4: Any EJCs associated with mRNA that were not displaced by a translating ribosome, interact with the terminating ribosome through an interaction between UPF2 associated with the EJC and UPF1 associated with the SURF complex, leading to SMG1 phorphorylation of UPF1. Step 5: NMD factor SMG7 binds phosphorylated UPF1 as eRF1 and eRF3 are released. Step 6: Bound SMG7 causes the mRNP complex to associate with P-bodies where translation is inhibited and the abnormal mRNA is degraded. (b) Non-stop decay. (c) No-go decay. Bottom, diagram of ribosomes "stacking up" 5′ to a stalled ribosome. See NL Garneau, J Wilusz, and CJ Wilusz Nature Reviews *Mol Cell Biol* **8**:113 (2007); Part (c) Simms CL et al. (2017) *Mol. Cell* **68**, 361–373.

In the case of a properly spliced mRNA, the current hypothesis is that during translation of the mRNA by the first "pioneer" ribosome to translate the mRNA, EJCs are displaced from the mRNP. However, for mRNAs with a stop codon before the final exon junction, one or more EJCs remain associated with the mRNA, resulting in nonsense-mediated decay (Figure 9-37a).

Non-Stop Decay

This alternative mechanism of mRNA surveillance leads to the inhibition of translation and degradation of mRNAs that have been prematurely polyadenylated so that a translating ribosome does not encounter an in-frame stop codon and translation continues to the end of the poly(A) where the ribosome arrests, still tightly bound to the mRNA (Figure 9-37b). Such ribosomes associated with the 3′ end of an mRNA without terminating translation by the normal

mechanism involving peptide release factors (Figure 5-38) are recognized by the protein Ski7. The C-terminus of Ski7 is similar in structure to the GTPase domains of elongation factor-1A (EF1A) (see Figure 5-37) and peptide release factor eRF3 (see Figure 5-38) that bind to the empty A site of the ribosome. Ski7 is proposed to bind to the A site similarly, causing release of the ribosome from the mRNA. Ski7 then recruits the exosome to deadenylate and then rapidly decay the transcript in the $3' \rightarrow 5'$ direction.

No-Go Decay

This is the final mechanism of mRNA quality control in the cytoplasm we will discuss in this chapter. This releases ribosomes that are stalled during translation due to damage to the mRNA template or the presence of an unusually large and stable secondary structure in the mRNA $5'$ of the stop codon (see Figure 9-37c). A recent study indicates that ribosome collision is a critical trigger for mRNA cleavages that occur in the initial step of no-go decay. Ribosome collision occurs when ribosomes stack-up along the mRNA $5'$ to the block to ribosome translocation (see Figure 9-37c, *bottom*). This was detected experimentally because mRNAs with engineered long, stable stem loops that block ribosome translocation ("no-go mRNAs") exhibit regions of protection in 30 base intervals $5'$ to the block when RNase is added to mildly cross-linked, permeabilized cells treated as in the procedure for chromatin immunoprecipitation (Figure 8-12a, Step **1**). This

protection of the no-go mRNA in 30-base intervals presumably results from protection of the mRNA in the channel of the 30S ribosomal subunits (Figure 5-34) stacked up along the mRNA.

In vivo, this stacking up of ribosomes 5′ to the block to ribosome translocation causes endonucleolytic cleavage of the no-go mRNA. Although the endonuclease that does this has not been identified, it seems likely to be associated with the ribosome since deep sequencing revealed that cleavage sites in no-go mRNAs occur in regions of the mRNA within the mRNA channels of 30S subunits (see Figure 5-34) phased behind the block to translocation. In humans, the E3 ubiquitin ligase ZNF598 polyubiquitinylates several ribosomal proteins leading to their proteosomal degradation as well as degradation of nascent polypeptide and release of the ribosomal subunits from the cleaved ends of the no-go mRNA. This is followed by degradation of the 5′ fragment of the cleaved no-go mRNA by cytoplasmic exosomes and degradation of the 3′ fragment by cytoplasmic XRN1 (Figure 9-37c).

Localization of mRNAs Permits Production of Proteins at Specific Regions Within the Cytoplasm

Many cellular processes depend on localization of particular proteins to specific structures or regions of the cell. In later

chapters, we examine how some proteins are transported *after* their synthesis to their proper cellular location. Alternatively, protein localization can be achieved by localization of mRNAs to specific regions of the cytoplasm in which their encoded proteins function. In most cases examined thus far, such mRNA localization is specified by sequences in the 3′ untranslated region of the mRNA. A recent genomic-level study of mRNA localization in *Drosophila* embryos revealed that ∼70 percent of 3000 mRNAs analyzed were localized to specific subcellular regions, raising the possibility that this is a much more general phenomenon in all metazoa than previously appreciated.

Localization of mRNAs to the Bud in *S. Cerevisiae*

The most thoroughly understood example of mRNA localization occurs in the budding yeast *S. cerevisiae*. As discussed in Chapter 8, whether a haploid yeast cell exhibits the **a** or α mating type is determined by whether **a** or α genes are present at the expressed *MAT* locus on chromosome III (see Figure 8-31). The process that transfers **a** or α genes from a silent mating-type locus near the telomeres of chromosome III to the expressed *MAT* locus is initiated by a sequence-specific endonuclease called *HO*. Transcription of the *HO* gene is dependent on the SWI/SNF chromatin-remodeling complex (see Section 8.4). Daughter yeast cells that arise by budding from mother cells contain a transcriptional repressor called *Ash1* (for *A*symmetric *s*ynthesis of *H*O) that prevents recruitment of the

SWI/SNF complex to the *HO* gene, thereby preventing its transcription. The absence of Ash1 from mother cells allows them to transcribe the *HO* gene. As a consequence, mother cells switch their mating type, while daughter cells generated by budding do not (Figure 9-38).

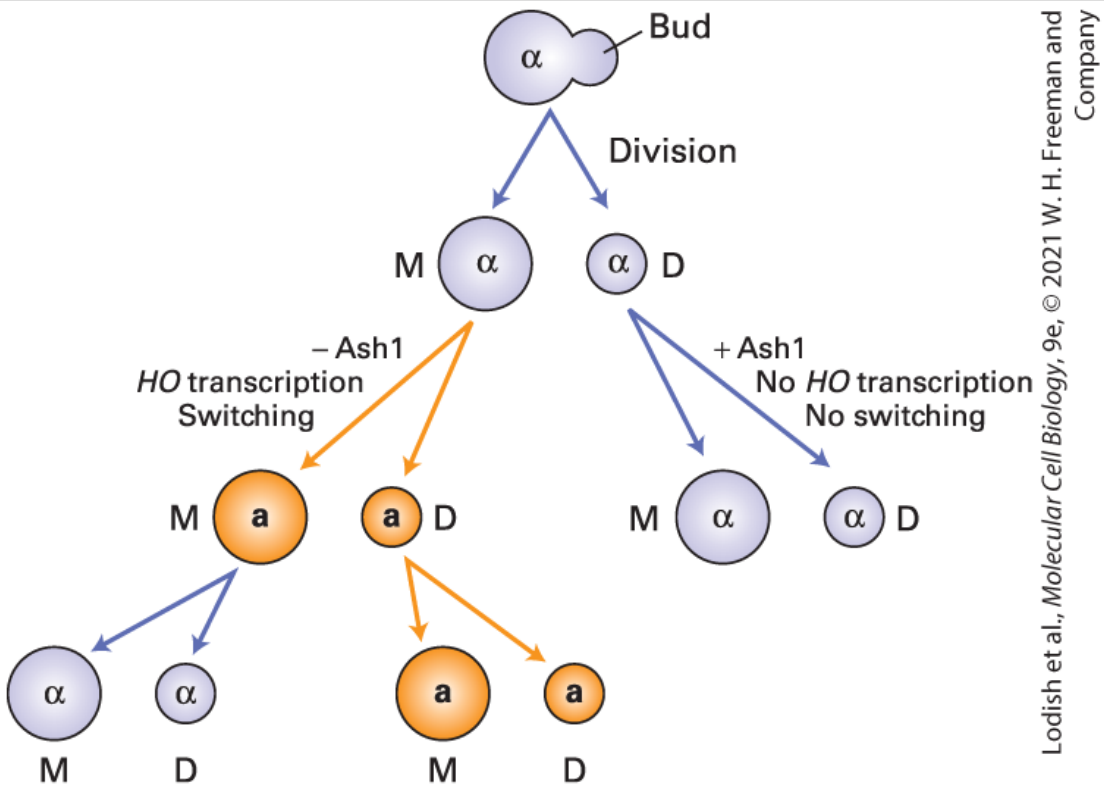


FIGURE 9-38 Mating-type switching in haploid budding yeast cells. Following nuclear DNA synthesis in *S. cerevisiae*, a haploid yeast cell undergoes mitosis and generates a small daughter cell from a daughter cell bud (Figure 1-24) and a larger mother cell [Cell division 1]. The resulting daughter and mother cells have the same mating type as the parent cell (α in this example, represented by purple color). Before the next cell division, the smaller daughter cell grows in size to the size of a mother cell and then divides [Cell division 2] into a smaller daughter cell and a larger mother cell that have the same mating type as the parental cell (α in this example). In contrast, the larger mother cell generated in Cell division 1 switches its mating type before DNA synthesis and divides into a mother and daughter cell with the

opposite mating type (**a** in this example, represented by orange). This process continues in further divisions so that daughter cells have progeny of the same mating type, while the mother cells have progeny that have switched their mating type compared to their mother.

Ash1 protein accumulates only in daughter cells because the mRNA encoding it is localized to daughter cells. The localization process requires three proteins: She2 (for *S*WI-dependent *HO e*xpression), an RNA-binding protein that binds specifically to a localization signal with a specific RNA structure in the ASH1 mRNA; Myo4, a myosin motor protein that moves cargos on actin filaments (see Chapter 17); and She3, which links She2 and therefore ASH1 mRNA to Myo4 (Figure 9-39). ASH1 mRNA is transcribed in the nucleus of the mother cell before mitosis. Movement of Myo4 with its bound ASH1 mRNA along actin filaments that extend from the mother cell into the bud carries the ASH1 mRNA into the growing bud before cell division.

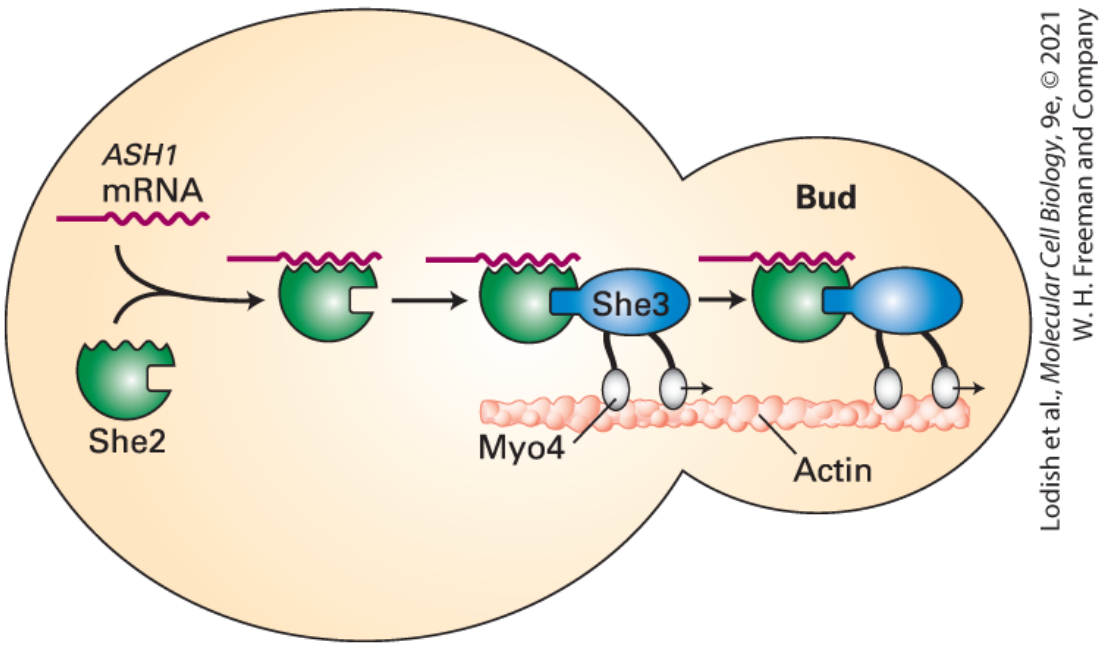
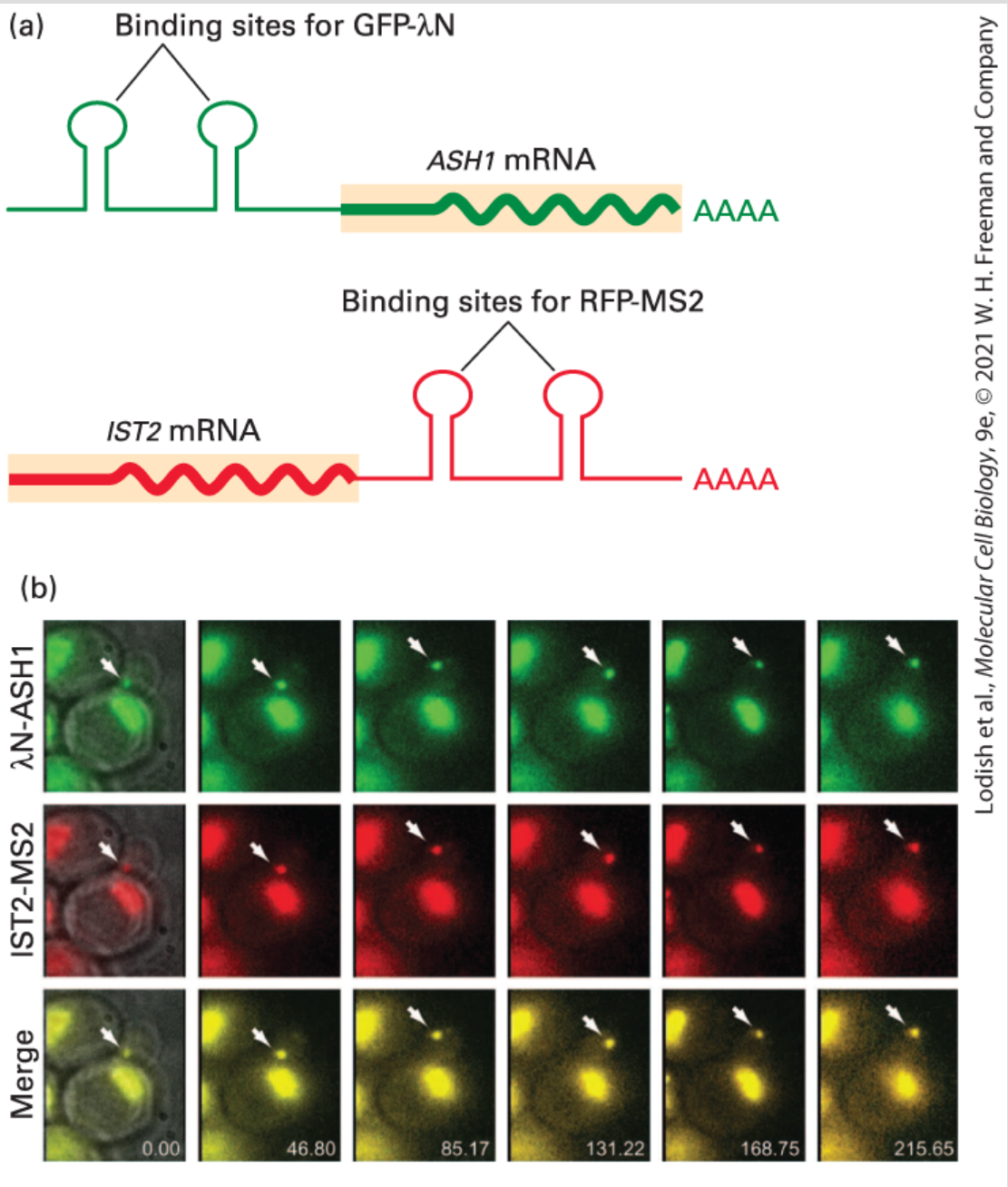


FIGURE 9-39 Model for restriction of mating-type switching to mother cells in *S. cerevisiae*. Ash1 protein prevents a cell from transcribing the *HO* gene whose encoded protein initiates the DNA rearrangement that results in mating-type switching from α to **a** or **a** to α. Switching occurs only in the mother cell, after it separates from a newly budded daughter cell, because the Ash1 protein is present only in the daughter cell. The molecular basis for this differential localization of Ash1 is the one-way transport of ASH1 mRNA into the bud. A linking protein, She2, binds to specific 3′ untranslated sequences in the ASH1 mRNA and also binds to She3 protein. This protein in turn binds to a myosin motor, Myo4, which moves along actin filaments into the bud. See S. Koon and B. J. Schnapp, 2001, *Curr. Biol.* **11**:R166.

At least 23 other mRNAs were found to be transported by the She2, She3, Myo4 system. All have an RNA localization signal to which She2 binds, usually in their 3′ UTR. The process can be visualized in live cells by the experiment shown in Figure 9-40. RNAs can be fluorescently labeled by including in their sequence high-affinity binding sites for RNA-binding proteins, such as bacteriophage

proteins MS2 coat protein and bacteriophage λN protein, that bind to different stem loops of specific sequence (Figure 9-40a). When such engineered mRNAs are expressed in budding yeast cells along with the bacteriophage proteins fused to proteins that fluoresce different colors, the fusion proteins bind to these specific RNA sequences, thereby labeling the RNAs that contain them with different colors. In the experiment shown in Figure 9-40b, ASH1 mRNA was labeled by the binding of green fluorescent protein fused to λN. Another mRNA localized to the bud by this system, the IST2 mRNA, encoding a component of the growing bud membrane, was labeled by the binding of red fluorescent protein fused to MS2 coat protein. Video of a budding cell showed that the differently labeled ASH1 and IST2 mRNAs accumulated in the same large cytoplasmic RNP particle containing multiple mRNAs in the mother cell cytoplasm, as can be seen from the merge of the green and red fluorescent signals. The RNP particle was then transported into the bud within about one minute.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 9-40 Transport of mRNP particles from a yeast mother cell into the bud. (a) Yeast cells were engineered to express an Ash1 mRNA with binding sites for the bacteriophage λN protein in its 5′ untranslated region and an IST2 mRNA with binding sites for bacteriophage MS2 coat protein in its 3′ untranslated region. A fusion of green fluorescent protein to λN protein (GFP-λN) and a fusion of red fluorescent protein to MS2 coat protein (RFP-MS2) also were expressed in the same cells. In other experiments, these fluorescently tagged sequence-specific RNA-binding proteins were shown to bind to their own specific binding sites engineered

into the ASH1 and IST2 mRNAs, and not to each others' binding sites. Both fluorescently tagged proteins also contained a nuclear localization signal so that the fluorescent proteins that were not bound to their high-affinity binding sites in these mRNAs were transported into nuclei through nuclear pore complexes (see Chapter 13). This was necessary to prevent high fluorescence from excess GFP-λN and RFP-MS2 in the cytoplasm. At the right, GFP-λN and RFP-MS2 were independently visualized by using millisecond alternating laser excitation of GFP and RFP. (b) Frames from a video of fluorescing cells are shown. The nucleus next to the large vacuole in the mother cell near the center of the micrographs, as well as nuclei in neighboring cells, was observed by green and red fluorescence as shown in the top and middle rows. A merge of the two images is shown in the bottom row, which also indicates the time elapsed between images. An RNP particle containing both the ASH1 mRNA with λN-binding sites and the IST2 mRNA with MS2-binding sites was observed in the mother cell cytoplasm in the left column of images (arrow). The particle increased in intensity between 0.00 and 46.80 seconds, indicating that more of these mRNAs joined the RNP particle. The RNP particle was transported into the bud between 46.80 and 85.17 seconds and then became localized to the bud tip.

[Republished with permission of John Wiley & Sons, Inc., from Lange, S. et al., "Simultaneous transport of different localized mRNA species revealed by live-cell imaging," 2008, *Traffic*, **9**:(8)1256–67; permission conveyed through Copyright Clearance Center, Inc.]

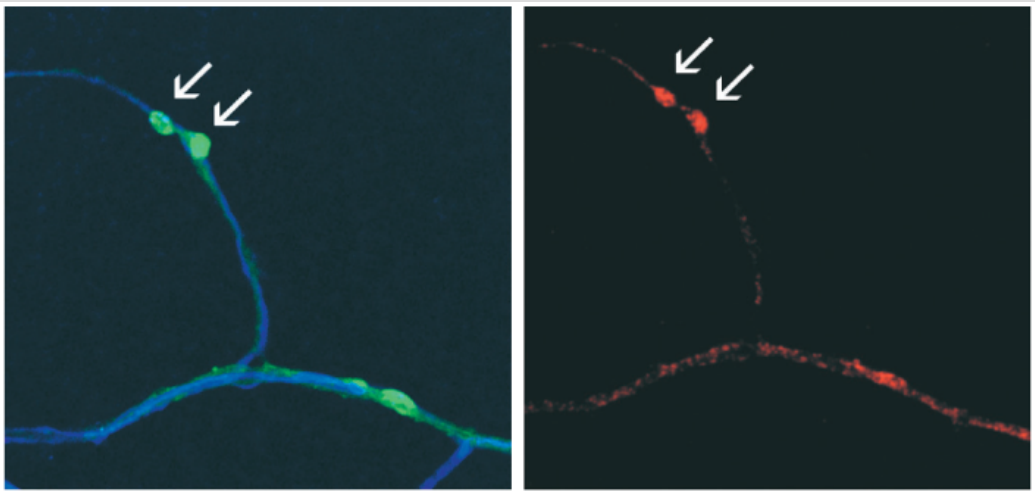
As discussed above in the section on P bodies, cytoplasmic sites of mRNA decay, formation of the large cytoplasmic granules of RNPs observed here (Figure 9-40b) and in other examples of transported RNA granules in cells of multicellular eukaryotes requires sequences in the RNA-binding protein composed of a low complexity amino acid sequence such as repeats of [G/S]Y[G/S]. Peptides containing these low complexity sequences spontaneously associate in vitro by a process referred to as "liquid-liquid phase separation" discussed in Chapter 8 in regard to formation of

transcriptional condensates in the nucleus (Figure 8-38). The condensates formed in vitro from these peptides with simple sequence repeats can be dissociated by phosphorylation of serines within them. Such liquid-liquid phase separated droplets are probably involved in the formation of the large RNP complexes transported on actin cables in yeast and on microtubules (see Chapters 17 and 18) in large asymmetric cells in higher eukaryotes, such as in the axons and dendrites of nerve cells (see Chapter 23) described in the next section. Regulated phosphorylation of these low complexity sequences in RNA-binding proteins associated with RNP granules may well account for the regulated formation and dissociation of RNP granules such as those observed in Figure 9-40b.

Localization of mRNAs to Synapses in the Mammalian Nervous System

As mentioned earlier, in neurons, localization of specific mRNAs at synapses far from the nucleus in the cell body plays an essential function in learning and memory (Figure 9-41). Like the localized mRNAs in yeast, these mRNAs contain RNA localization signals in their 3′ untranslated region. Some of these mRNAs are initially synthesized with short poly(A) tails that do not allow translation initiation. Once again, large RNP particles containing multiple mRNAs bearing localization signals form in the cytoplasm near the cell nucleus. In this case, the RNP particles are transported down the axon to synapses by kinesin motor proteins that travel down

microtubules extending the length of the axon (see Chapter 18). Membrane depolarization at a given synapse by the regulated opening of ion channels in the plasma membrane (see Chapter 23) may then stimulate the mRNAs' polyadenylation in the region of the synapse, activating the translation of encoded proteins that increase the size and alter the neurophysiological properties of the one synapse while leaving unaffected the hundreds to thousands of other synapses made by the same neuron.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 9-41 A specific neuronal mRNA localizes to synapses. Sensory neurons from the sea slug *Aplysia californica* were cultured with target motor neurons so that axons from the sensory neurons formed synapses with dendrites from the motor neurons in culture. The micrograph at the left shows motor neuron axons visualized with a blue fluorescent dye. GFP-VAMP (green) was expressed in sensory neurons and marks the location of synapses formed between sensory neuron axons and motor neuron dendrites (arrows). The micrograph on the right shows red fluorescence from in situ hybridization of an anti-sensorin mRNA probe. Sensorin is a neurotransmitter expressed only by the sensory neuron; sensory neuron axons are not otherwise visualized in this preparation, but they lie adjacent to the motor neuron dendrites. The in situ hybridization results indicate that sensorin mRNA in the sensory neuron is localized to synapses.

[Republished with permission of Elsevier, from Lyles, V., et al., "Synapse formation and mRNA localization in cultured *Aplysia* neurons," *Neuron*, 2006, **49**(3):349–356; permission conveyed through Copyright Clearance Center, Inc.]

KEY CONCEPTS OF SECTION 9.4

Cytoplasmic Mechanisms of Post-Transcriptional Control

- Most mRNAs are degraded in the cytoplasm as the result of the gradual shortening of the poly(A) tail (deadenylation) followed by exosome-mediated $3' \rightarrow 5'$ digestion, or removal of the $5'$ cap and digestion by a $5' \rightarrow 3'$ exoribonuclease, XRN1 (see Figure 9-30).
- Eukaryotic mRNAs encoding proteins that are expressed in short bursts generally have repeated copies of an AU-rich sequence (AU-rich element) in their $3'$ UTR. Specific proteins that bind to these elements also interact with a deadenylating enzyme complex and cytoplasmic exosomes, promoting rapid mRNA degradation.
- Translation can be repressed by microRNAs (miRNAs), which form imperfect hybrids with sequences in the $3'$ untranslated region (UTR) of specific target mRNAs. mRNAs bound by several miRNAs are concentrated in P bodies in the cytoplasm, where they are degraded by decapping followed by digestion by the cytoplasmic exosome.
- The related phenomenon of RNA interference (RNAi), which probably evolved as an early defense system against viruses and transposons, leads to rapid degradation of mRNAs that form perfect hybrids with short interfering RNAs (siRNAs).
- Both miRNAs and siRNAs are 21–23 nucleotides in length, are generated from longer precursor molecules and are bound by an Argonaute protein and assembled into a multiprotein RNA-induced silencing complex (RISC). RISC complexes either repress translation of target mRNAs and induce their localization to P bodies, where they are degraded (miRNAs), or cleave them (siRNAs), generating unprotected ends that are rapidly degraded by cytoplasmic exosomes and the $5' \rightarrow 3'$ exonuclease XRN1 (see Figures 9-31 and 9-32).
- Cytoplasmic polyadenylation is required for the translation of mRNAs with a short poly(A) tail. This step is regulated in cells: binding of a specific protein to regulatory elements in the $3'$ UTRs represses translation of these mRNAs, and phosphorylation of this RNA-binding protein leads to lengthening of the $3'$ poly(A) tail and thus translation (see Figure 9-34).

- The global rate of cellular protein synthesis is regulated in part by controlling the activity of eukaryotic initiation factor 2 (eIF2) required at an early step in selecting mRNAs for translation (see Figure 5-24). eIF2 kinases respond to various types of cellular stress by phosphorylating a regulatory site in eIF2, inhibiting eIF2 exchange of GTP for GDP and consequently inhibiting an early step in protein synthesis.
- In addition to regulation of the global rate of protein synthesis by most mRNAs, the translation, stability, and cytoplasmic localization of specific mRNAs can be controlled. This is generally accomplished by sequence-specific RNA-binding proteins that bind to regulatory elements in the $3'$ or less commonly the $5'$ UTRs of some mRNAs. For example, translation of ferritin mRNA and degradation of transferrin receptor (TfR) mRNA are both regulated by the same iron-sensitive RNA-binding proteins, IREBP1 and 2. At low iron concentrations, these proteins have an active conformation that binds to specific sequences that form stem-loops in the mRNAs, inhibiting ferritin mRNA translation and degradation of TfR mRNA (see Figure 9-35). This dual control precisely regulates the intracellular Fe^{2+} concentration.
- Nonsense-mediated decay and other mRNA surveillance mechanisms prevent the translation of improperly processed mRNAs encoding abnormal proteins that might interfere with the functioning of the corresponding normal proteins.
- Many mRNAs are transported to specific subcellular locations by sequence-specific RNA-binding proteins that bind localization sequences usually found in the $3'$ UTR. These RNA-binding proteins then associate, directly or via intermediary proteins, with motor proteins that carry large RNP particles, containing many mRNAs bearing the localization signals, on actin or microtubule fibers to specific locations in the cytoplasm.

9.5 Processing of rRNA and tRNA

Approximately 80 percent of the total RNA in rapidly growing mammalian cells (e.g., cultured HeLa cells) is rRNA, and 15 percent is tRNA; protein-coding mRNA thus constitutes only a small portion of the total RNA. The primary transcripts produced from most rRNA genes and from tRNA genes, like pre-mRNAs, are extensively processed to yield the mature, functional forms of these RNAs.

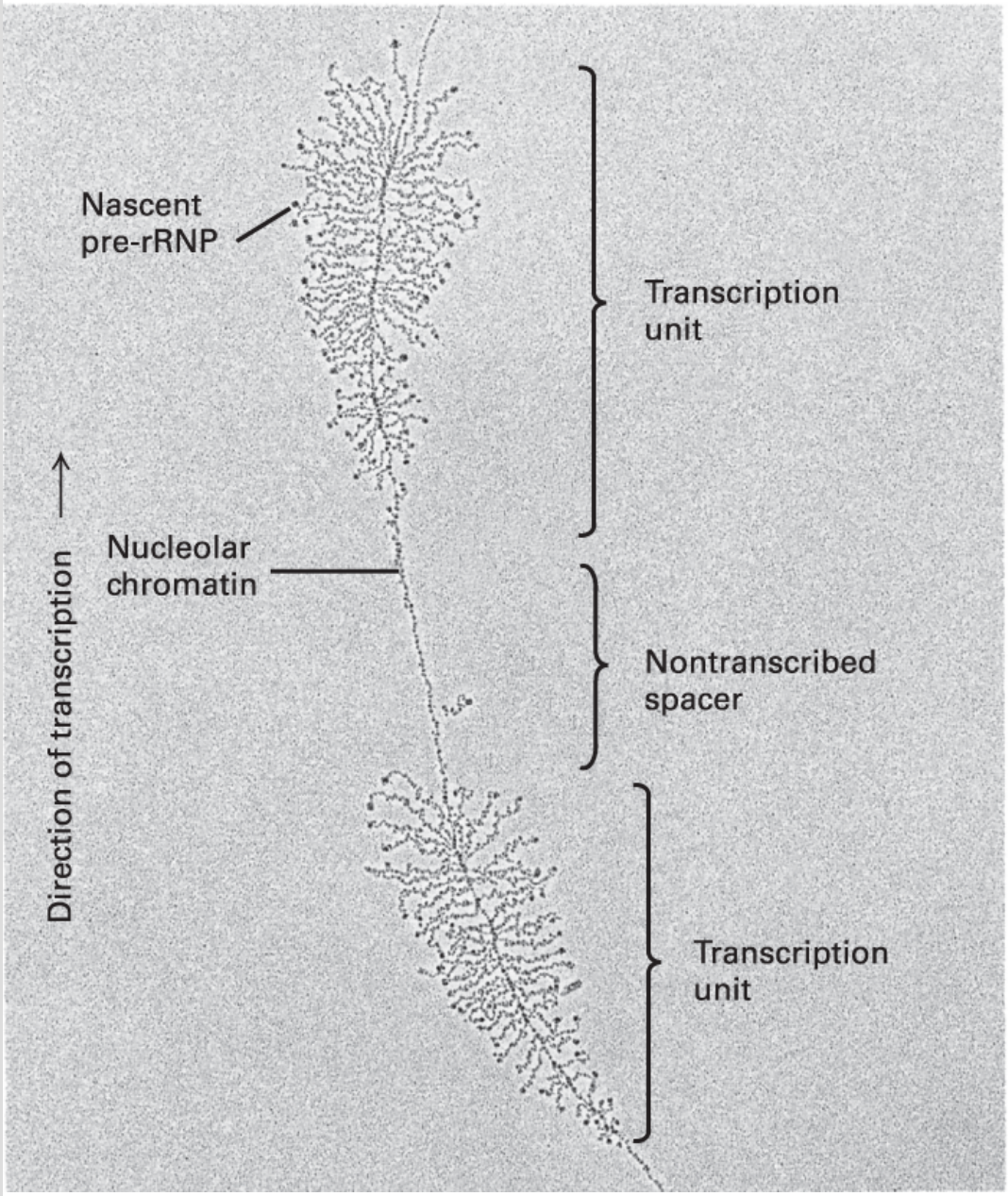
The ribosome is a highly evolved, complex structure (see Figure 5-34), optimized for its function in protein synthesis. Ribosome synthesis requires the function and coordination of all three nuclear RNA polymerases, RNA pol I, II, and III. The 28S and 5.8S rRNAs associated with the large ribosomal subunit and the single 18S rRNA of the small subunit are transcribed by RNA polymerase I. The 5S rRNA of the large subunit is transcribed by RNA polymerase III, and the mRNAs encoding the ribosomal proteins are transcribed by RNA polymerase II. In addition to the four rRNAs and ∼70 ribosomal proteins, at least 150 other RNAs and proteins interact transiently with the two ribosomal subunits during their assembly through a series of coordinated steps. Furthermore, multiple specific bases and riboses of the mature rRNAs are modified to optimize their function in protein synthesis. Although most of the steps in ribosomal subunit synthesis and assembly occur in the nucleolus, some occur in the nucleoplasm during passage from the nucleolus to nuclear pore complexes. A quality-control step occurs before

nuclear export so that only fully functional subunits are exported to the cytoplasm, where the final steps of ribosome subunit maturation occur. tRNAs also are processed from precursor primary transcripts in the nucleus and modified extensively before they are exported to the cytoplasm and used in protein synthesis. First, we'll discuss the processing and modification of rRNA and the assembly and nuclear export of ribosomes. Then we'll consider the processing and modification of tRNAs.

Pre-rRNA Genes Are Similar in All Eukaryotes and Function as Nucleolar Organizers

The 28S and 5.8S rRNAs associated with the large (60S) ribosomal subunit and the 18S rRNA associated with the small (40S) ribosomal subunit in multicellular eukaryotes (and the functionally equivalent rRNAs in all other eukaryotes) are encoded by a single type of pre-rRNA transcription unit. In human cells, transcription by RNA polymerase I yields a 45S ($\sim$13.7 kb) primary transcript (pre-rRNA), which is processed into the mature 28S, 18S, and 5.8S rRNAs found in cytoplasmic ribosomes. The fourth rRNA, 5S, is encoded separately and transcribed outside the nucleolus by RNA polymerase III. Sequencing of the DNA encoding the 45S pre-rRNA from many species showed that this DNA shares several properties in all eukaryotes. First, the pre-rRNA genes are arranged in long

tandem arrays separated by nontranscribed spacer regions ranging in length from $\sim$2 kb in frogs to $\sim$30 kb in humans (Figure 9-42). Second, the genomic regions corresponding to the three mature rRNAs are always arranged in the same $5' \rightarrow 3'$ order: 18S, 5.8S, and 28S. Third, in all eukaryotic cells (and even in bacteria), the pre-rRNA gene codes for regions that are removed during processing and rapidly degraded. These regions probably contribute to proper folding of the rRNAs but are not required once the folding has occurred. The general structure of pre-rRNA genes in several organisms is diagrammed in Figure 9-43.



EXPERIMENTAL FIGURE 9-42 Electron micrograph of pre-rRNA transcription units from the nucleolus of a frog oocyte. Each "feather" represents multiple pre-rRNA molecules associated with protein in a pre-ribonucleoprotein complex (pre-rRNP) emerging from a transcription unit. Note the dense "knob" at the $5'$ end of each nascent pre-RNP thought to be a processome. Pre-rRNA transcription units are arranged in tandem, separated by nontranscribed spacer regions of nucleolar chromatin.

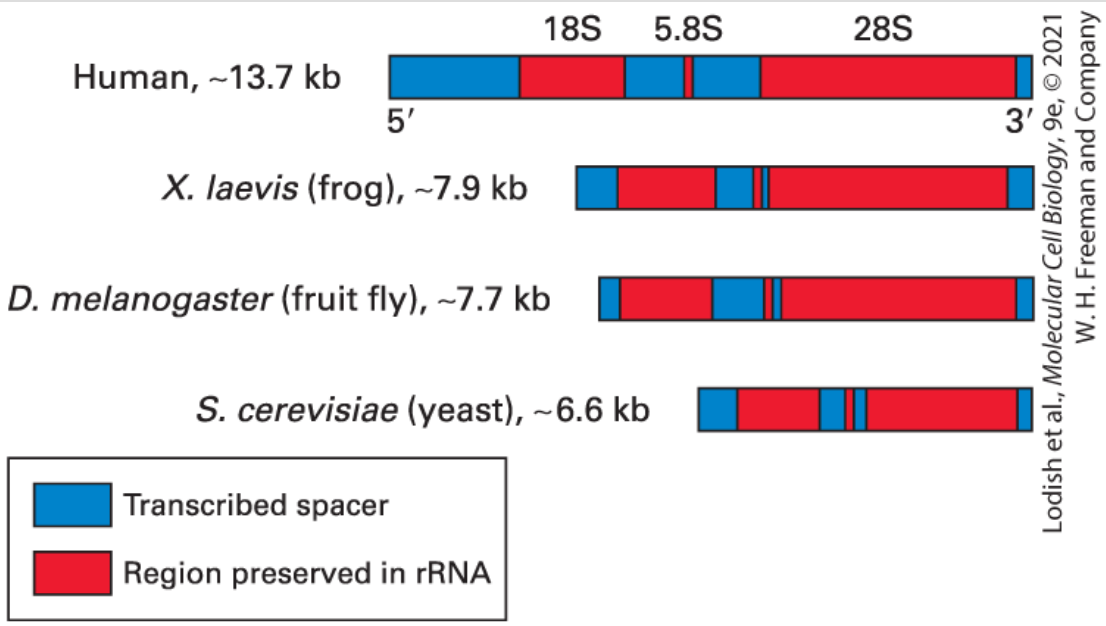


FIGURE 9-43 General structure of eukaryotic pre-rRNA transcription units. The three coding regions (red) encode the 18S, 5.8S, and 28S rRNAs found in ribosomes of multicellular eukaryotes or their equivalents in other species. The order of these coding regions in the genome is always $5' \rightarrow 3'$. Variations in the lengths of the transcribed spacer regions (blue) account for the major difference in the lengths of pre-rRNA transcription units in different organisms.

The synthesis and most of the processing of pre-rRNA occurs in the nucleolus. When pre-rRNA genes initially were identified in the nucleolus by in situ hybridization, it was not known whether any other DNA was required to form the nucleolus. Subsequent experiments with transgenic *Drosophila* strains demonstrated that a single complete pre-rRNA transcription unit induces formation of a small nucleolus. Thus a single pre-rRNA gene is sufficient to be a "nucleolar organizer," and all the other components of the ribosome diffuse to the newly formed pre-rRNA. The structure of the nucleolus observed by light and electron microscopy results

primarily from the processing of pre-RNA and the assembly of ribosomal subunits.

Small Nucleolar RNAs Assist in Processing Pre-rRNAs

Ribosomal subunit assembly, maturation, and export to the cytoplasm are best understood in the yeast *S. cerevisiae*. However, nearly all the proteins and RNAs involved are highly conserved in multicellular eukaryotes, where the fundamental aspects of ribosome biosynthesis are likely to be the same. As we have seen with pre-mRNAs, nascent pre-rRNA transcripts are immediately bound by proteins, forming pre-ribosomal ribonucleoprotein particles (pre-rRNPs). Cleavage of the pre-rRNA does not begin until transcription of the pre-rRNA is nearly complete. In yeast, it takes approximately 6 minutes for a pre-rRNA to be transcribed. Once transcription is complete, the rRNA is cleaved, and bases and riboses are modified in about 10 seconds. In a rapidly growing yeast cell, ~40 pairs of ribosomal subunits are synthesized, processed, and transported to the cytoplasm every second. This extremely high rate of ribosome synthesis despite the seemingly long period required to transcribe a pre-rRNA is possible because pre-rRNA genes are packed with RNA polymerase I molecules transcribing the same gene simultaneously (see Figure 9-42) and because there are

100–200 such genes on chromosome XII, the yeast nucleolar organizer.

The primary transcript of ~6.6 kb is cut in a series of cleavage and exonucleolytic steps that ultimately yield the mature rRNAs found in ribosomes (Figure 9-44). During processing, pre-rRNA is also extensively modified, mostly by methylation of the 2′-hydroxyl group of specific riboses and conversion of specific uridine residues to pseudouridine. Virtually all of these modifications occur in the most conserved core structure of the ribosome, which is directly involved in protein synthesis. The positions of the specific sites of 2′-*O*-methylation and pseudouridine formation are determined by approximately 150 different small nucleolus- restricted RNA species, called **small nucleolar RNAs (snoRNAs)**, which hybridize transiently to pre-rRNA molecules. Like the snRNAs that function in pre-mRNA processing, snoRNAs associate with proteins, forming ribonucleoprotein particles called *snoRNPs*. One class of more than 40 snoRNPs (box C/D containing snoRNAs, see Figure 9-45a) positions a methyltransferase enzyme near methylation sites in the pre-mRNA. The multiple different box C/D snoRNAs direct methylation at multiple sites through a similar mechanism. They share common sequence and structural features, including the C and D box sequences (Figure 9-45a) and are bound by a common set of proteins. One or two regions of each of these snoRNAs are precisely complementary to sites on the pre-rRNA and direct the methyltransferase to specific riboses in the hybrid region (see Figure

9-45a). A second major class of snoRNPs (box H/ACA containing snoRNAs, Figure 9-45b) positions the enzyme that converts uridine to pseudouridine by rotation of the pyrimidine ring (Figure 9-45c). Bases on either side of the modified uridine in the pre-rRNA base-pair with bases in the bulge of a stem in the H/ACA snoRNAs, leaving the uridine to be modified bulged out of the helical double-stranded region, like the branch-point A bulges out in pre-mRNA spliceosomal splicing (see Figure 9-9d). Other modifications of pre-rRNA nucleotides, such as adenine dimethylation, are carried out by specific proteins without the assistance of guiding snoRNAs.

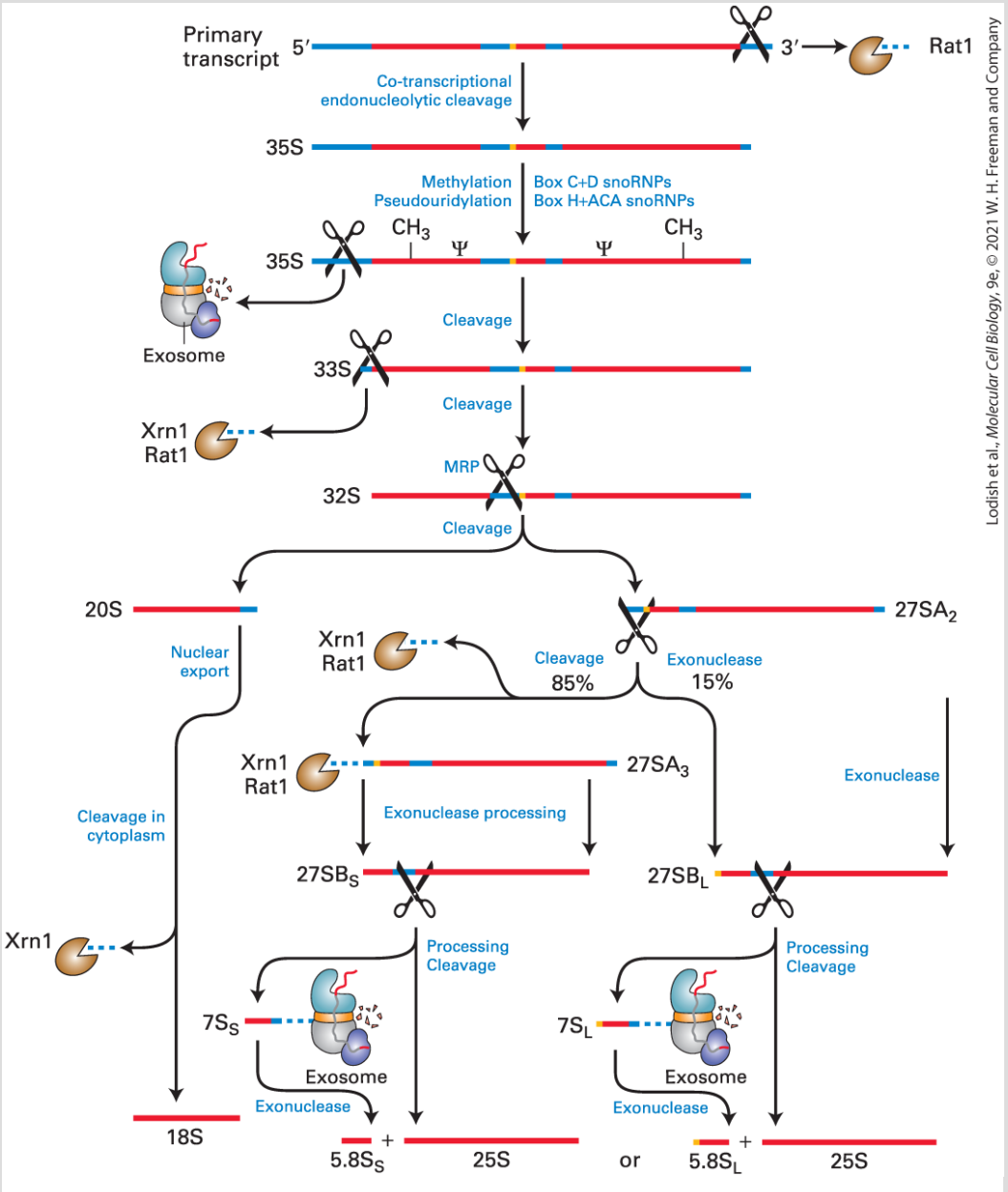


FIGURE 9-44 rRNA processing. Endoribonucleases that make internal cleavages are represented as scissors. Exoribonucleases that digest from one end, either $5'$ or $3'$, are shown as Pac-Men. Most $2'$-O-ribose methylation (CH_3) and generation of pseudouridines in the rRNAs occurs following the initial cleavage at the $3'$ end, before the initial cleavage at the $5'$ end. Proteins and snoRNPs known to participate in these steps are indicated. See J. Venema and D. Tollervey, 1999, *Ann. Rev. Genetics* **33**:261.

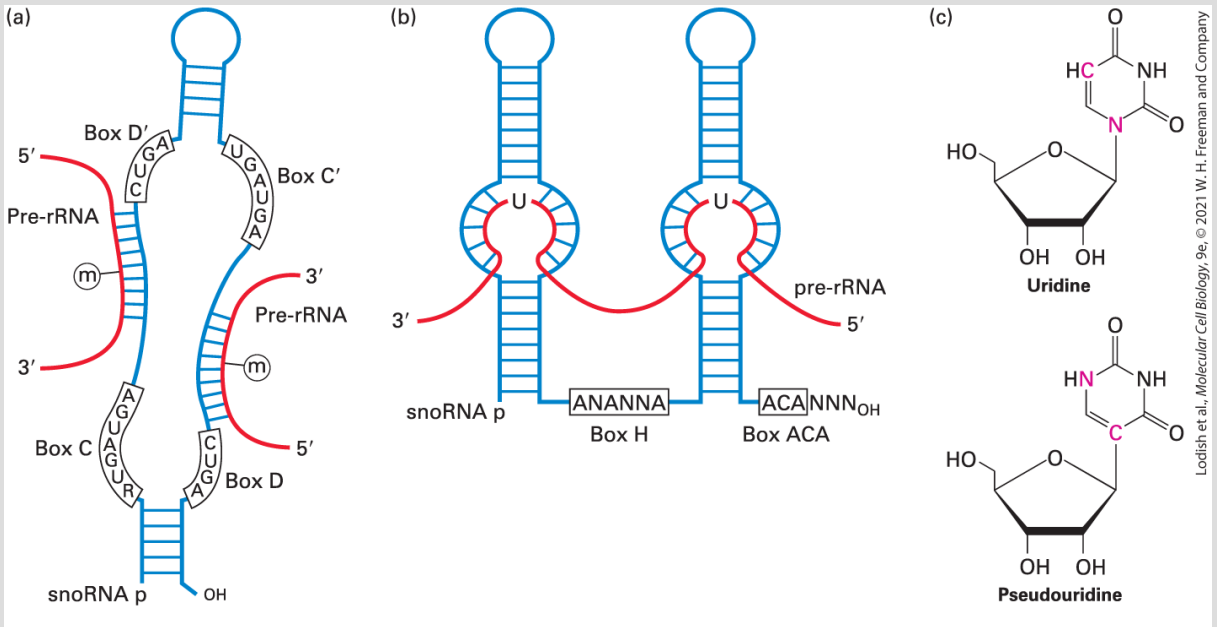


FIGURE 9-45 snoRNP-directed modification of pre-rRNA. (a) SnoRNAs called *box C/D snoRNAs* are involved in ribose 2′-O-methylation. Sequences in these snoRNAs hybridize to multiple different regions in the pre-rRNA, directing methylation at the indicated sites. (b) Box H/ACA snoRNAs fold into two stem loops with internal single-stranded bulges in the stems. Pre-rRNA hybridizes to the single-stranded bulges, demarcating a site of pseudouridylation. (c) Conversion from uridine to pseudouridine directed by the box H/ACA snoRNAs of part (b). See T. Kiss, 2001, *EMBO J.* **20**:3617. Part (b) from U. T. Meier, 2005, *Chromosoma* **114**:1.

The U3 snoRNA is assembled into a large snoRNP containing ~72 proteins called the *small subunit (SSU) processome*, which specifies cleavage at site A_0, the initial cut near the 5′ end of the pre-rRNA (see Figure 9-44). The processome helps to fold the 6.6-kb pre-rRNA into a structure required for the multiple cleavage events that follow transcription of the primary pre-rRNA transcript. U3 snoRNA base-pairs with an upstream region of the pre-rRNA to specify the location of the cleavage. The processome is thought to form the "5′ knob" visible in electron micrographs of pre-rRNPs (see Figure 9-

42). Base pairing of other snoRNPs specify additional cleavage reactions that remove transcribed spacer regions. The first cleavage to initiate processing of the 5.8S and 25S rRNAs of the large subunit is performed by RNase MRP, a complex of nine proteins with an RNA. Once cleaved from pre-rRNAs, these sequences are degraded by the same exosome-associated $3' \rightarrow 5'$ nuclear exonuclease that degrades introns spliced from pre-mRNAs. Nuclear $5' \rightarrow 3'$ exoribonucleases (Rat1; Xrn1) also remove some regions of $5'$ spacer.

Some snoRNAs are expressed from their own promoters by RNA polymerase II or III. Remarkably, however, the large majority of snoRNAs are processed from spliced-out introns of genes encoding functional mRNAs for proteins involved in ribosome synthesis or translation. Some snoRNAs are processed from introns spliced from apparently nonfunctional mRNAs. The genes encoding these mRNAs seem to exist only to express snoRNAs from excised introns.

Unlike 18S, 5.8S, and 28S genes, 5S rRNA genes are transcribed by RNA polymerase III in the nucleoplasm outside the nucleolus. With only minor additional processing to remove nucleotides at the $3'$ end, 5S rRNA diffuses to the nucleolus, where it assembles with the pre-rRNA precursor and remains associated with the region that is cleaved into the precursor of the large ribosomal subunit.

Most of the ribosomal proteins of the small 40S ribosomal subunit associate with the nascent pre-rRNA during transcription (Figure 9-46). Cleavage of the full-length pre-rRNA in the 90S RNP precursor releases a pre-40S particle that requires only a few more remodeling steps, that is gain or loss of associated proteins, base modifications, and changes in conformation, before it is transported to the cytoplasm, as diagrammed in Figure 9-46. Once the pre-40S particle leaves the nucleolus (purple in Figure 9-46), it traverses the nucleoplasm (blue) quickly and is exported through nuclear pore complexes (NPCs) to the cytoplasm (tan), as discussed below. Final maturation of the small ribosomal subunit occurs in the cytoplasm: exonucleolytic processing of the 20S rRNA into mature small subunit 18S rRNA by the cytoplasmic $5' \rightarrow 3'$ exoribonuclease Xrn1 and the dimethylation of two adjacent adenines near the $3'$ end of 18S rRNA by the cytoplasmic enzyme Dim1.

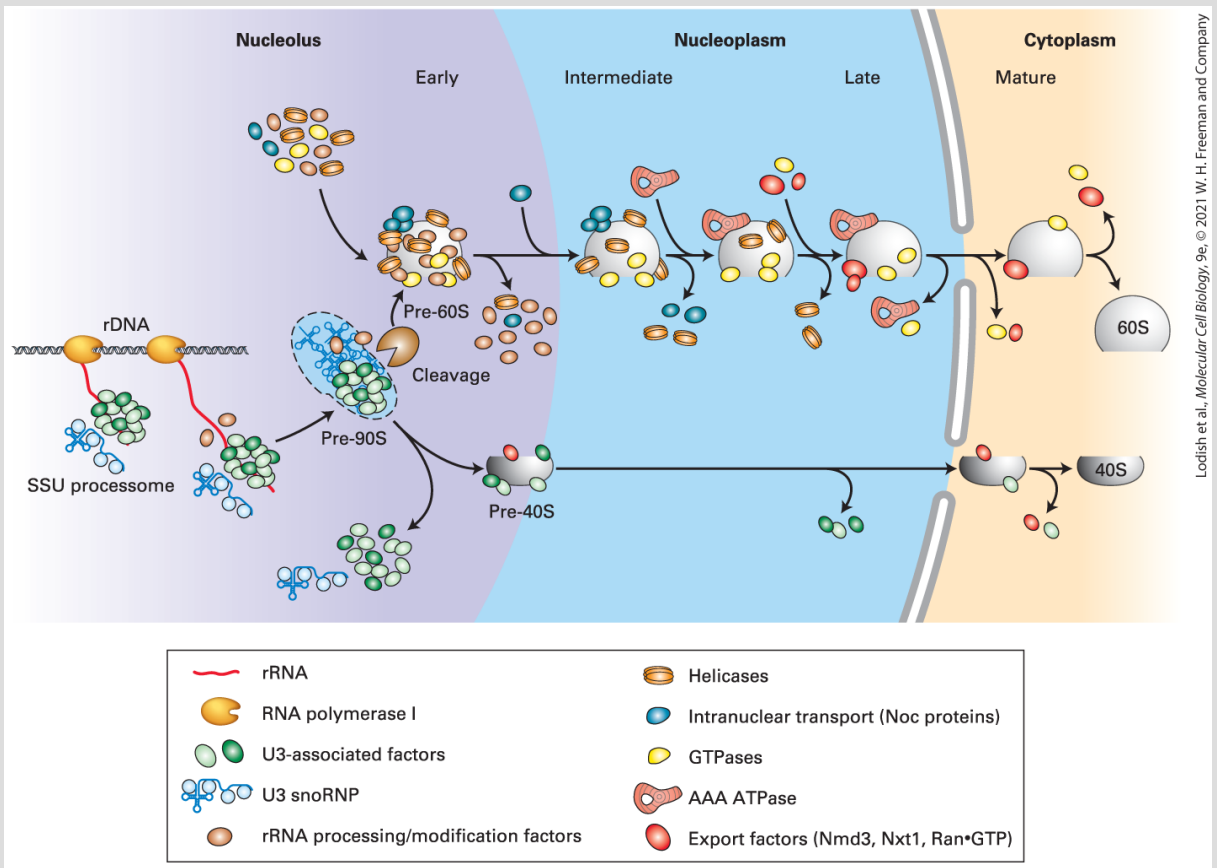


FIGURE 9-46 Ribosomal subunit assembly. Ribosomal proteins and RNAs in the maturing small and large ribosomal subunits are depicted in blue, with a shape similar to the icons for the mature subunits in the cytoplasm. Other factors that associate transiently with the maturing subunits are depicted in different colors, as shown in the key. See H. Tschochner and E. Hurt, 2003, *Trends Cell Biol.* **13**:255.

In contrast to the pre-40S particle, the precursor of the large subunit requires considerably more remodeling through many more transient interactions with nonribosomal proteins before it is sufficiently mature for export to the cytoplasm. Consequently, it takes a considerably longer period for the maturing 60S subunit to exit the nucleus (30 minutes compared to 5 minutes for export of the 40S subunit in cultured human cells). Multiple RNA helicases and

small G proteins are associated with the maturing pre-60S subunits. Some RNA helicases are necessary to dislodge the snoRNPs that base-pair perfectly with pre-rRNA over up to 30 base pairs. Other RNA helicases may function in the disruption of protein-RNA interactions. The requirement for so many GTPases suggests that there are many quality-control checkpoints in the assembly and remodeling of the large subunit RNP, where one step must be completed before a GTPase is activated to allow the next step to proceed. Members of the AAA ATPase family also bind transiently. This class of proteins is often involved in large molecular movements and may be required to fold the complex, large rRNA into the proper conformation. Some steps in 60S subunit maturation occur in the nucleoplasm (Figure 9-46, blue), during passage from the nucleolus to nuclear pore complexes. Much remains to be learned about the complex, fascinating, and essential remodeling processes that occur during formation of the ribosomal subunits.

The large ribosomal subunit is one of the largest structures to pass through the nuclear pore complexes. Maturation of the large subunit in the nucleoplasm leads to the generation of binding sites for a nuclear export adapter called *Nmd3*. Nmd3 is bound by the nuclear transporter Exportin 1 (also called *Crm1*). This is another quality-control step, because only correctly assembled subunits can bind Nmd3 and be exported. The small subunit of the mRNP exporter (NXT1) also becomes associated with the nearly mature large ribosomal subunit. These nuclear transporters permit diffusion of

the large subunit through the central channel of the NPC, which is filled with a cloud of unstructured protein domains that extend from the structured parts of proteins that line the wall of the channel (see Chapter 13). Several additional subunits that form the walls of the NPC central channel are also required for ribosomal subunit export and may have additional functions specific for this task. The dimensions of ribosomal subunits (~25–30 nm in diameter) and the central channel of the NPC are comparable, so passage may not require distortion of either the ribosomal subunit or the channel. Final maturation of the large subunit in the cytoplasm includes removal of these export factors. As for the export of most macromolecules from the nucleus, including tRNAs and pre-miRNAs (but not most mRNPs), ribosome subunit export requires the function of a small G protein called *Ran*, as discussed in Chapter 13.

Self-Splicing Group I Introns Were the First Examples of Catalytic RNA

During the 1970s, all the pre-rRNA genes of the protozoan *Tetrahymena thermophila* were discovered to contain an intron, indicating that splicing is required to produce mature rRNA in these organisms. In 1982, in vitro studies showing that the pre-rRNA was spliced at the correct sites in the absence of any protein provided

the first indication that RNA can function as a catalyst, like protein enzymes.

A whole raft of self-splicing sequences were subsequently found in pre-rRNAs from other single-celled organisms, in mitochondrial and chloroplast pre-rRNAs, in several pre-mRNAs from certain *Escherichia coli* bacteriophages, and in some bacterial tRNA primary transcripts. The self-splicing sequences in all these precursors, referred to as *group I introns*, use guanosine as a cofactor and can fold by internal base pairing to juxtapose closely the two exons that must be joined. As discussed earlier, certain mitochondrial and chloroplast pre-mRNAs and tRNAs contain a second type of self-splicing intron, designated *group II*.

The splicing mechanisms used by group I introns, group II introns, and spliceosomes are generally similar, involving two transesterification reactions, which require no input of energy (Figure 9-47). Structural studies of the group I intron from *Tetrahymena* pre-rRNA combined with mutational and biochemical experiments revealed that the RNA folds into a precise three-dimensional structure that, like protein enzymes, contains deep grooves for binding substrates and solvent-inaccessible regions that function in catalysis. The group I intron functions like a metalloenzyme to precisely orient the atoms that participate in the two transesterification reactions adjacent to catalytic Mg^{2+} ions. Considerable evidence now indicates that splicing by group II

introns and by snRNAs in the spliceosome also involves bound catalytic Mg^{2+} ions. In both the group I and II self-splicing introns, RNA functions as a **ribozyme**, an RNA sequence with catalytic ability. As discussed above, spliceosomal splicing is also catalyzed by Mg^{2+} ions held in the proper positions by U6 snRNA and its interactions with U1 and U5 snRNAs (see Figure 9-14c), but only with considerable help from multiple proteins that also shape the catalytic site and help bind and release substrates.

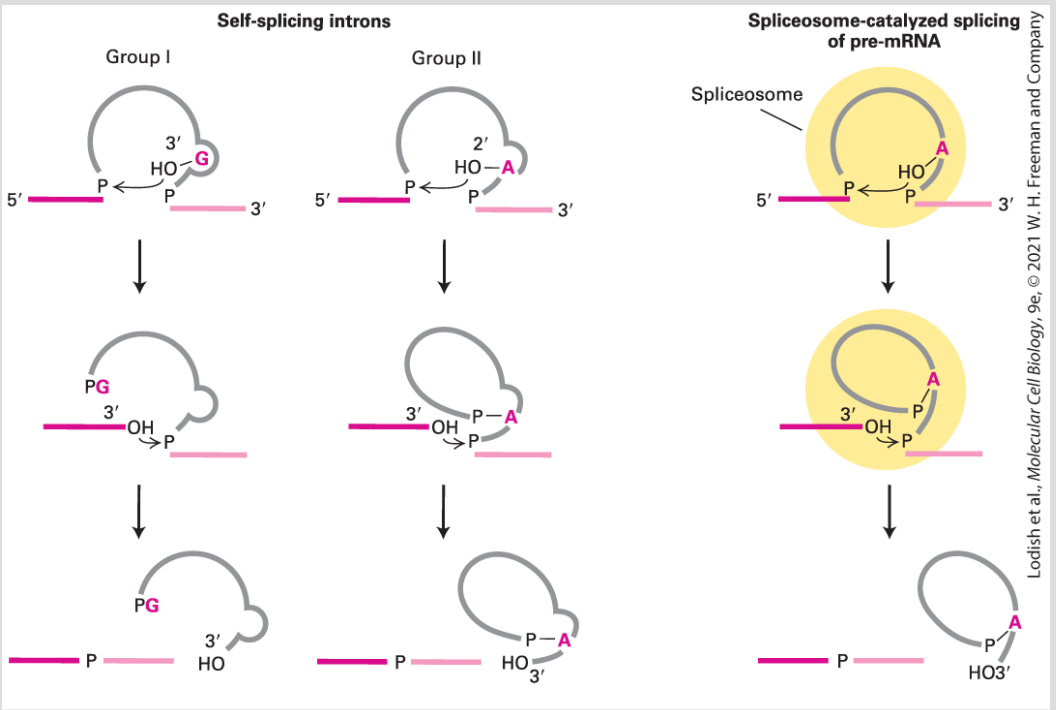


FIGURE 9-47 Splicing mechanisms in group I and group II self-splicing introns and spliceosome-catalyzed splicing of pre-mRNA. The intron is shown in gray, the exons to be joined in red. In group I introns, a guanosine cofactor (G) that is not part of the RNA chain associates with the active site. The 3′-hydroxyl group of this guanosine participates in a transesterification reaction with the phosphate at the 5′ end of the intron; this reaction is analogous to that involving the 2′-hydroxyl groups of branch-site A's in group II introns and pre-mRNA introns spliced in spliceosomes (see Figure 8-8). The subsequent

transesterification that links the 5′ and 3′ exons is similar in all three splicing mechanisms. Note that spliced-out group I introns are linear structures, unlike the branched intron products in the other two cases. See P. A. Sharp, 1987, *Science* **235**:769.

Pre-tRNAs Undergo Extensive Modification in the Nucleus

Mature cytosolic tRNAs, which average 75–80 nucleotides in length, are produced from larger precursors (pre-tRNAs) synthesized by RNA polymerase III in the nucleoplasm. Mature tRNAs also contain numerous modified bases that are not present in tRNA primary transcripts. Cleavage and base modification occur during processing of all pre-tRNAs; some pre-tRNAs also are spliced during processing. All of these processing and modification events occur in the nucleus.

A 5′ sequence of variable length that is absent from mature tRNAs is present in all pre-tRNAs (Figure 9-48). This occurs because the 5′ end of mature tRNAs is generated by an endonucleolytic cleavage specified by the tRNA three-dimensional structure rather than the start site of transcription. These extra 5′ nucleotides are removed by ribonuclease P (RNase P), a ribonucleoprotein endonuclease. Studies with *E. coli* RNase P indicate that at high Mg^{2+} concentrations the RNA component alone can recognize and cleave *E. coli* pre-tRNAs. The RNase P polypeptide increases the rate of

cleavage by the RNA, allowing it to proceed at physiological Mg^{2+} concentrations. A comparable RNase P functions in eukaryotes.

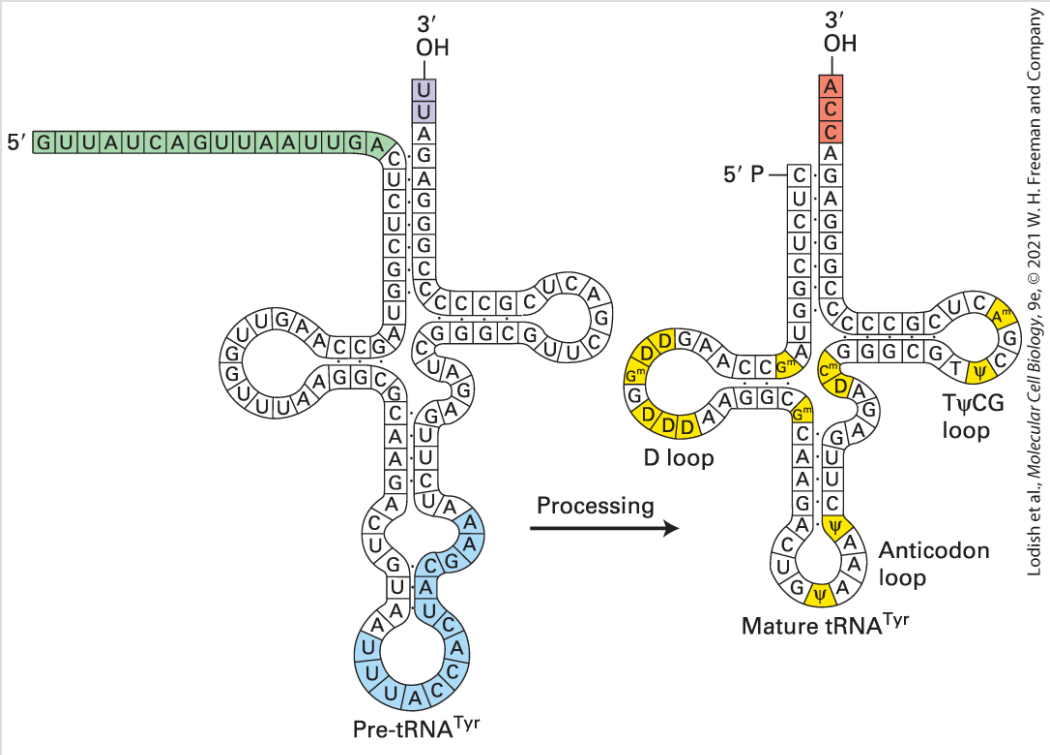


FIGURE 9-48 Changes that occur during the processing of tyrosine pre-tRNA. A 14-nucleotide intron (blue) in the anticodon loop is removed by splicing. A 16-nucleotide sequence (green) at the 5′ end is cleaved by RNase P. U residues at the 3′ end are replaced by the CCA sequence (red) found in all mature tRNAs. Numerous bases in the stem-loops are converted to characteristic modified bases (yellow). Not all pre-tRNAs contain introns that are spliced out during processing, but they all undergo the other types of changes shown here. D = dihydrouridine; Ψ = pseudouridine.

About 10 percent of the bases in pre-tRNAs are modified enzymatically during processing. Three classes of base modifications have been described (Figure 9-48): (1) U residues at the 3′ end of pre-tRNA are replaced with a CCA sequence. The CCA

sequence is found at the 3′ end of all tRNAs and is required for their charging by amino-acyl-tRNA synthetases during protein synthesis. This step in tRNA synthesis likely functions as a quality-control point, since only properly folded tRNAs are recognized by the CCA addition enzyme. (2) Methyl and isopentenyl groups are added to the heterocyclic ring of purine bases, and the 2′-OH groups in the ribose of specific residues are methylated. (3) Specific uridines are converted to dihydrouridine, pseudouridine, or ribothymidine residues. The functions of these base and ribose modifications are not well understood, but since they are highly conserved, they probably function in protein synthesis.

As shown in Figure 9-48, the pre-tRNA expressed from the yeast tyrosine tRNA (tRNATyr) gene contains a 14-base intron that is not present in mature tRNATyr. Some other eukaryotic tRNA genes and some archaeal tRNA genes also contain introns. The introns in nuclear pre-tRNAs are shorter than those in pre-mRNAs and lack the consensus splice-site sequences found in pre-mRNAs (see Figure 9-7c). Pre-tRNA introns are also clearly distinct from the much longer self-splicing group I and group II introns found in chloroplast and mitochondrial pre-rRNAs. The mechanism of pre-tRNA splicing differs greatly from spliceosomal and self-splicing group I and group II intron splicing (see Figure 9-47). First, splicing of pre-tRNAs is catalyzed by proteins, not by RNAs. Second, a pre-tRNA intron is excised in one step that entails simultaneous cleavage at both ends of the intron. Finally, hydrolysis of GTP and ATP is required to join

the two tRNA halves generated by cleavage on either side of the intron.

After pre-tRNAs are processed in the nucleoplasm, the mature tRNAs are transported to the cytoplasm through nuclear pore complexes by Exportin-t, a transporter (see Chapter 13) dedicated to the nuclear export of tRNAs. In the cytoplasm, tRNAs are passed between aminoacyl-tRNA synthetases, elongation factors, and ribosomes during protein synthesis (Chapter 5). Thus tRNAs generally are associated with proteins and spend little time free in the cell, as is also the case for mRNAs and rRNAs.

KEY CONCEPTS OF SECTION 9.5

Processing of rRNA and tRNA

- A large precursor pre-rRNA (13.7 kb in humans) transcribed by RNA polymerase I undergoes cleavage, exonucleolytic digestion, and base modifications to yield mature 28S, 18S, and 5.8S rRNAs, which associate with ribosomal proteins into ribosomal subunits.
- Transcription and processing of pre-rRNA occur in the nucleolus. The 5S rRNA component of the large ribosomal subunit is synthesized in the nucleoplasm by RNA polymerase III.
- Approximately 150 snoRNAs, associated with proteins in snoRNPs, base-pair with specific sites in pre-rRNA, where they direct ribose methylation, modification of uridine to pseudouridine, and cleavage at specific sites during rRNA processing in the nucleolus.
- Group I and group II self-splicing introns, and probably snRNAs in spliceosomes, all function as ribozymes, or catalytically active RNA sequences, that carry out splicing by analogous transesterification reactions requiring bound Mg^{2+} ions (see Figure 9-47).
- Pre-tRNAs synthesized by RNA polymerase III in the nucleoplasm are processed by removal of the $5'$-end sequence, addition of CCA to the $3'$ end, and modification of

multiple internal bases (see Figure 9-48).

- Some pre-tRNAs contain a short intron that is removed by a protein-catalyzed mechanism distinct from the splicing mechanisms used by pre-mRNAs and self-splicing introns.
- All species of RNA molecules are associated with proteins in various types of ribonucleoprotein particles, both in the nucleus and after export to the cytoplasm.

9.6 Nuclear Bodies Are Functionally Specialized Nuclear Domains

High-resolution visualization of plant- and animal-cell nuclei by electron microscopy and by staining with fluorescently labeled antibodies has revealed specialized nuclear domains, called **nuclear bodies**. These domains are not surrounded by membranes. Nonetheless, they are regions of high concentrations of specific proteins and RNAs that form distinct, often roughly spherical structures within the nucleus (Figure 9-49). The most prominent nuclear bodies are nucleoli, the sites of ribosomal subunit synthesis and assembly discussed earlier. Several other types of nuclear bodies also have been described in structural studies.

Experiments with fluorescently labeled nuclear proteins have shown that the nucleus is a highly dynamic environment, with rapid diffusion of proteins through the nucleoplasm. Proteins associated with nuclear bodies are often also observed at lower concentrations in the nucleoplasm outside the nuclear bodies, and fluorescence studies indicate that they diffuse in and out of the nuclear bodies. Based on these measurements of molecular mobility in living cells, nuclear bodies can be mathematically modeled as the expected steady state for diffusing proteins that interact with sufficient affinity to form self-organized regions of high concentrations of specific proteins, but with low enough affinity for each other to be able to diffuse in and out of the structure. This is similar to the

situation for transcriptional condensates discussed in Chapter 8 (Figure 8-38) and cytoplasmic ribonuclear protein particles involved in localization of mRNAs associated with motor proteins that "walk" on actin filaments and microtubules (see Figures 9-40b and 9-41). In electron micrographs, these structures appear to be a heterogeneous, spongelike network of interacting components. We discuss a few examples of nuclear bodies here.

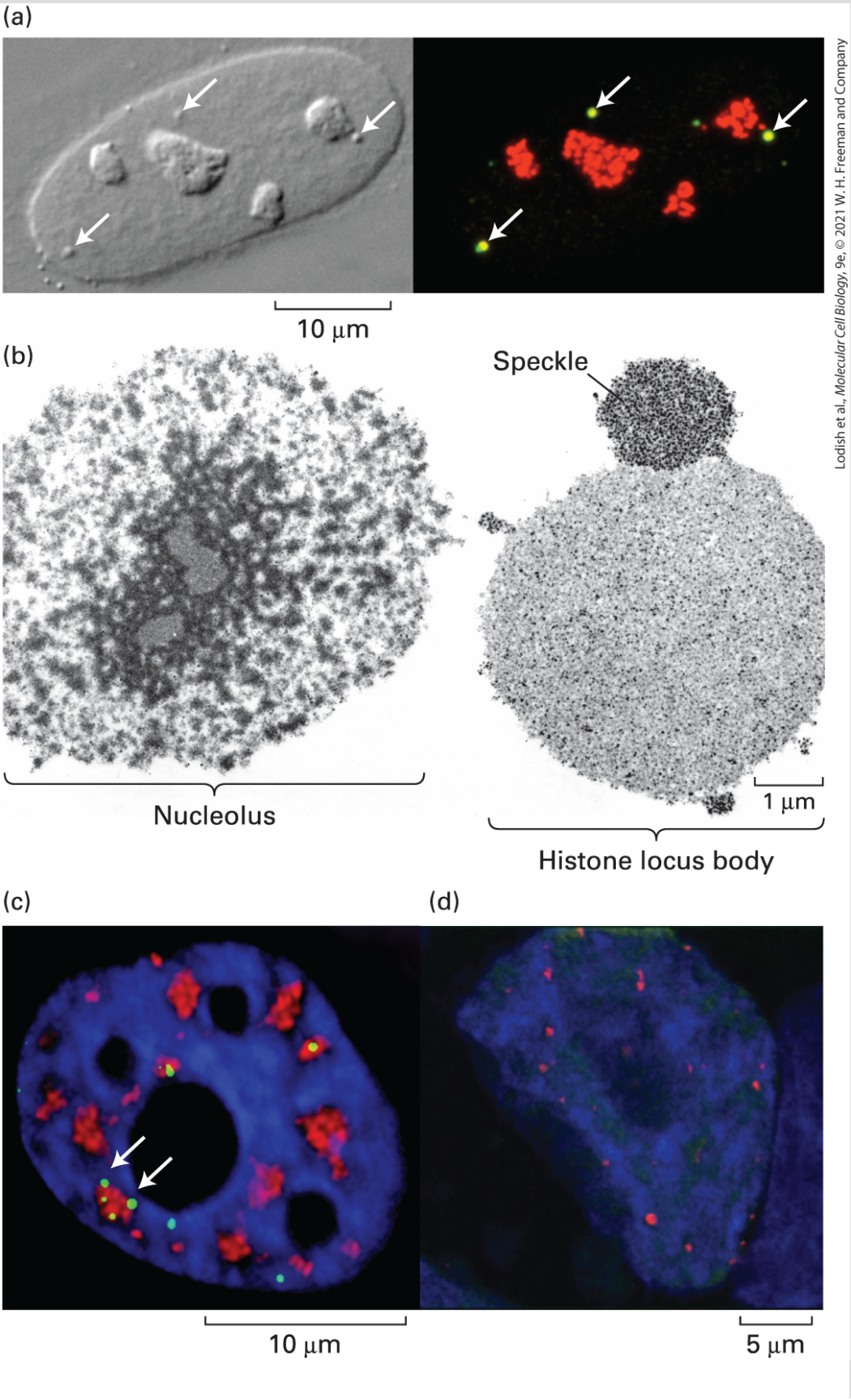
Cajal Bodies

Cajal bodies are ~0.2–1 μm spherical structures that have been observed in large nuclei for more than a century (Figure 9-49a). Current research indicates that like nucleoli, Cajal bodies are centers of RNP-complex assembly, in this case for spliceosomal snRNPs and other nuclear RNPs. Like rRNAs, snRNAs undergo specific modifications, such as the conversion of specific uridine residues to pseudouridine and addition of methyl groups to the 2′-hydroxyl groups of specific riboses. These post-transcriptional modifications are important for the proper assembly and function of snRNPs in pre-mRNA splicing. These modifications occur in Cajal bodies, where they are directed by a class of snoRNA-like guide RNA molecules called *scaRNAs* (small Cajal body–associated RNAs). There is also evidence that the Cajal body is the site of reassembly of the U4/U6/U5 tri-snRNP complexes required for pre-mRNA splicing from the free U4, U5, and U6 snRNPs released during the removal of each intron (see Figure 9-13). Since Cajal bodies also contain a high

concentration of the U7 snRNP involved in the specialized 3′-end processing of the major histone mRNAs, it is likely that this process also occurs in Cajal bodies, as may the assembly of the telomerase RNP.

Nuclear Speckles

Nuclear speckles were observed using fluorescently labeled antibodies to snRNP proteins and other proteins involved in pre-mRNA splicing, as approximately 25–50 irregular, amorphous structures 0.5–2 μm in diameter that are distributed through the nucleoplasm of vertebrate cells (Figure 9-49c, d). Since speckles are not located at sites of co-transcriptional pre-mRNA splicing, which are associated closely with chromatin, they are thought to be storage regions for snRNPs and proteins involved in pre-mRNA splicing that are released into the nucleoplasm when required.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 9-49 Examples of nuclear bodies. (a, b) Cajal bodies and nucleoli in a HeLa cell nucleus. (a, *left*) DIC image showing four nucleoli and three Cajal bodies (arrowheads). Scale bar = 10 μm. (a, *right*) The same nucleus immunostained with antibodies against coilin (green) and fibrillarin (red). The three Cajal bodies appear yellow because they stain with both antibodies. The nucleoli stain only for fibrillarin. Fibrillarin is the methyltransferase for $2'$-O-methylation of rRNA in the nucleoli and snRNAs in the Cajal Bodies. (b) Transmission electron microscopy of nuclear bodies in a Xeopus oocyte nucleus. These extraordinarily large nuclei are larger than in most vertebrate cells. A nucleolus and a histone body with an associated speckle were relatively close to each other in this oocyte nucleus, and consequently were observed in the same electron micrograph. Scale bar = 1 μm. (c) HeLa cell stained with DAPI (blue), antibody to SC35 (red), a splicing factor stored in nuclear bodies called "speckles," and antibody to PSPC1, a protein found in nuclear bodies called "paraspeckles" (white arrows) because they are most often observed close to speckles. Scale bar = 10 μm. (d) PML-nuclear bodies in an H1299 cell (a lung canrcinoma cell line) nucleus. DNA was stained with DAPI (blue) and PML-nuclear bodies were immunostained with antibody to the major protein in these bodies, PML. Scale bar = 1 μm.

[Part (a) Reprinted with permission of Nature Publishing Group, from Gall, J.G., "The centennial of the Cajal body," *Nat Rev Mol Cell Biol*., 2003, **4**(12):975–980; permission conveyed through Copyright Clearance Center, Inc. Part (b) Republished with permission of Elsevier, from Handwerger, K. E. and Gall, J. G., "Subnuclear organelles: new insights into form and function," *Trends Cell Biol*. 2006, **16**(1):19–26; permission conveyed through Copyright Clearance Center, Inc. Part (c) Republished with permission from Cold Spring Harbor Laboratory Press, from A. H. Fox and A. I. Lamond, 2010, "Paraspeckles," *Cold Spring Harb. Perspect. Biol*. **2**(7):a000687. Part (d) Republished with permission from American Society for Microbiology, from M. A. Pennella et al., 2010, "Adenovirus E1B 55-Kilodalton Protein Is a p53-SUMO1 E3 Ligase That Represses p53 and Stimulates Its Nuclear Export Through Interactions with Promyelocytic Leukemia Nuclear Bodies," *J. Virol*. **84**(23):12210–12225; permission conveyed through Copyright Clearance Center, Inc.]

Nuclear Paraspeckles

Paraspeckles are composed of RNPs formed by the interaction between a long nonprotein-coding RNA species (lncRNA), NEAT1, and members of the DBHS (*Drosophila* Behavior Human Splicing) family of proteins, P54NRB/NONO, PSPC1, and PSF/SFPQ. Paraspeckles are critical to the control of gene expression through the nuclear retention of RNA containing double-stranded RNA regions that have been subjected to adenosine-to-inosine editing. In this way, they may function in the incompletely understood mRNA quality-control mechanisms that operate in the nucleus.

Promyelocytic Leukemia (PML) Nuclear Bodies

The PML gene was originally discovered when chromosomal translocations within the gene were observed in the leukemic cells of patients with the rare disease promyelocytic leukemia (PML). When antibodies specific for the PML protein were used in immunofluorescence microscopy studies, the protein was found to localize to ∼10–30 roughly spherical regions 0.3–1 μm in diameter in the nuclei of mammalian cells. Multiple functions have been proposed for these PML nuclear bodies, but a consensus is emerging that they function as sites for the assembly and modification of protein complexes involved in DNA repair and the induction of

apoptosis. For example, the important p53 tumor suppressor protein appears to be post-translationally modified by phosphorylation and acetylation in PML nuclear bodies in response to DNA damage, increasing its ability to activate the expression of DNA-damage response genes. PML nuclear bodies are also required for cellular defenses against DNA viruses that are induced by interferons, proteins secreted by virus-infected cells and T-lymphocytes involved in the immune response (see Chapter 24).

PML nuclear bodies are also sites of protein post-translational modification through the addition of a small, ubiquitin-like protein called *SUMO1* (small ubiquitin-like moiety-1), which can control the activity and subcellular localization of the modified protein. Many transcriptional activators are inhibited when they are sumoylated, and mutation of their site of sumoylation increases their activity in stimulating transcription. These observations indicate that PML nuclear bodies are involved in a mechanism of transcriptional repression that remains to be thoroughly understood.

Nucleolar Functions in Addition to Ribosomal Subunit Synthesis

The first nuclear bodies to be observed, the nucleoli, have specialized regions of substructure (Figure 9-49b) that are dedicated to functions other than ribosome biogenesis. There is evidence that

immature SRP ribonucleoprotein complexes involved in protein secretion and ER membrane insertion (Chapter 13) are assembled in nucleoli and then exported to the cytoplasm, where their final maturation takes place. The Cdc14 protein phosphatase that regulates processes in the final stages of mitosis is sequestered in nucleoli until chromosomes have been properly segregated into daughter cells (Chapter 19). Also, a tumor-suppressor protein called *ARF*, which is involved in the regulation of the protein encoded by the most frequently mutated gene in human cancers, *p53*, is sequestered in nucleoli and released in response to DNA damage (Chapter 25). In addition, heterochromatin often forms on the surface of nucleoli (see Figure 7-29a), suggesting that proteins associated with nucleoli also participate in the formation of this repressing chromatin structure.

KEY CONCEPT OF SECTION 9.6

Nuclear Bodies Are Functionally Specialized Nuclear Domains

- Nuclear bodies are functionally specialized regions in the nucleus where interacting proteins form self-organized structures. Many of these bodies, including the nucleolus, are regions of assembly of RNP complexes.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

5′ capping
3′ cleavage/polyadenylation
alternative splice sites
alternative splicing (alternative RNA splicing)
Argonaute protein
cross-exon recognition complex
Dicer
Drosha
exons
exon shuffling
exosome
group I introns

group II introns
introns
Iron Response Element Binding Proteins (IRE-BPs)
microexons
microRNAs (miRNAs)
mRNA surveillance
mRNP exporter
nuclear bodies
nuclear envelope
nuclear pore complexes (NPCs)
P bodies
pervasive transcription
poly(A) tail
pre-mRNA
pre-rRNA
ribonucleoprotein (RNP) complexes
ribozyme
RNA editing
RNA-induced silencing complex (RISC)
RNA interference (RNAi)
RNA splicing
short interfering RNAs (siRNAs)
siRNA knockdown
small nuclear RNAs (snRNAs)
small nucleolar RNAs (snoRNAs)
snRNPs (small ribonuclear particles) U1, U2, U4, U5, and U6
spliceosome
splicing enhancers

splicing silencers
SR proteins

Review the Concepts

1. Describe three types of post-transcriptional regulation of protein-coding genes.
2. True or false? The CTD is responsible for mRNA-processing steps that are specific for mRNA and not for other forms of RNA. Explain why you chose true or false.
3. There are a number of conserved sequences found in an mRNA that dictate where splicing occurs. Where are these sequences found relative to the exon-intron junctions? What is the significance of these sequences in the splicing process? One of these important regions is the branch-point A found in the intron. What is the role of the branch-point A in the splicing process, and can this be accomplished with the OH group on either the 2′ or the 3′ carbon?
4. What are the differences between hnRNAs, snRNAs, miRNAs, siRNAs, and snoRNAs?
5. What are the mechanistic similarities between group II intron self-splicing and spliceosomal splicing? What is the evidence that there may be an evolutionary relationship between the two?
6. You obtain the sequence of a gene containing 10 exons, 9 introns, and a 3′ UTR containing a polyadenylation

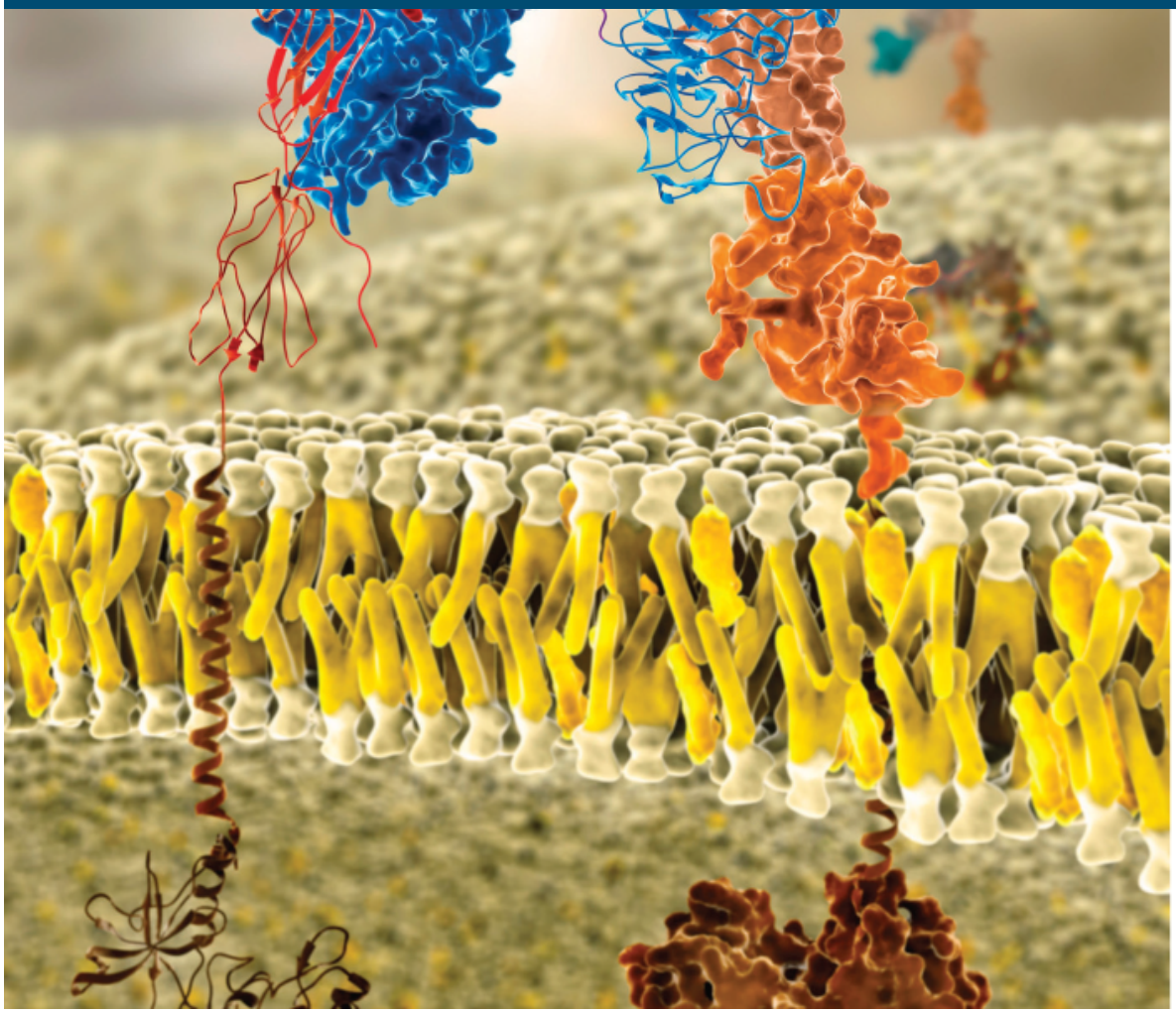
consensus sequence. The fifth intron also contains a polyadenylation site. To test whether both polyadenylation sites are used, you isolate mRNA and find a longer transcript from muscle tissue and a shorter transcript from all other tissues. Speculate about the mechanism involved in the production of these different transcripts.

7. RNA editing is a common process in the mitochondria of trypanosomes and plants as well as in chloroplasts, and in rare cases it occurs in higher eukaryotes. What is RNA editing, and what benefit does it demonstrate in the documented example of ApoB in humans?
8. Because DNA is found in the nucleus, transcription is a nuclear-localized process. Ribosomes responsible for protein synthesis are found in the cytoplasm. Why is hnRNP trafficking to the cytoplasm restricted to the nuclear pore complexes?
9. A protein complex in the nucleus is responsible for transporting mRNA molecules into the cytoplasm. Describe the proteins that form this exporter. What two protein groups are probably behind the mechanism involved in the directional movement of the mRNP and exporter into the cytosol?
10. RNA knockdown has become a powerful tool in the arsenal of methods used to repress gene expression. Briefly describe how gene expression can be knocked down.
11. Speculate about why plants deficient in Dicer activity show increased sensitivity to infection by RNA viruses.

12. mRNA stability is a key regulator of protein levels in a cell. Briefly describe the three mRNA degradation pathways. Suppose that a yeast cell has a mutation in the DCP1 gene, resulting in decreased uncapping activity. Would you expect to see a change in the P bodies found in this mutant cell?
13. mRNA localization now appears to be a common phenomenon. What benefit does mRNA localization have for a cell? What is the evidence that some mRNAs are directed to accumulate in specific subcellular locations?

CHAPTER 10

Biomembrane Structure



Ramón Andrade, 3Dciencia/Science Source.

Molecular model of a lipid bilayer with embedded membrane proteins. Integral membrane proteins have distinct exoplasmic, cytosolic, and membrane-spanning domains. Shown here are portions of the insulin receptor, which regulates cell metabolism.

OUTLINE

10.1 The Lipid Bilayer: Composition and Structural Organization
10.2 Membrane Proteins: Structure and Basic Functions
10.3 Phospholipids, Sphingolipids, and Cholesterol: Synthesis and Intracellular Movement

Membranes participate in many aspects of cell structure and function. The **plasma membrane** defines the cell and separates the inside from the outside. In eukaryotes, membranes also define intracellular organelles such as the nucleus, mitochondrion, and lysosome. These *biomembranes* all have the same basic architecture: a phospholipid bilayer in which proteins are embedded (Figure 10-1). By preventing the unassisted movement of most water-soluble substances across the membrane, the phospholipid bilayer serves as a permeability barrier, helping to maintain the characteristic differences between the inside and the outside of the cell or organelle. In turn, the embedded proteins endow the membrane with specific functions, such as regulated transport of substances from one side to the other. Each type of cellular membrane has its own set of proteins that allow it to carry out many different functions.

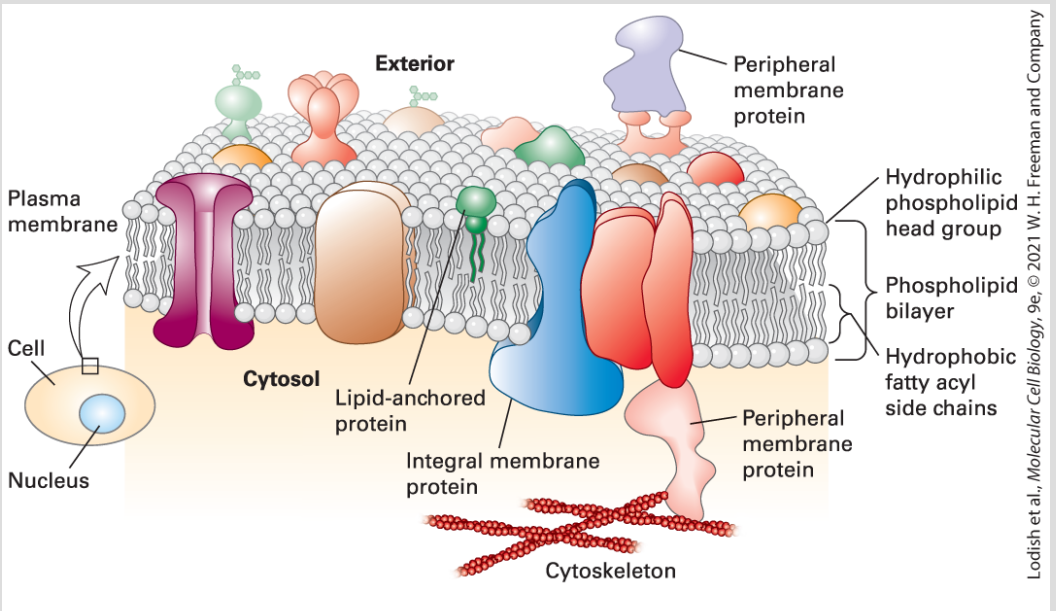


FIGURE 10-1 Fluid mosaic model of biomembranes. A bilayer of phospholipids about 3–4 nm thick provides the basic architecture of all cellular membranes; membrane proteins give each cellular membrane its unique set of functions. Individual phospholipids can move laterally and spin within the plane of the membrane, giving the membrane a fluidlike consistency similar to that of olive oil. Noncovalent interactions between phospholipids, and between phospholipids and proteins, lend strength and resilience to the membrane, while the hydrophobic core of the bilayer prevents the unassisted movement of water-soluble substances from one side to the other. Integral membrane proteins (transmembrane proteins) span the bilayer and often form dimers and higher order oligomers. Lipid-anchored proteins are tethered to one leaflet by a covalently attached hydrocarbon chain. Peripheral membrane proteins associate with the membrane primarily by specific noncovalent interactions with integral membrane proteins or membrane lipids. See D. Engelman, 2005, *Nature* **438**:578–580.]

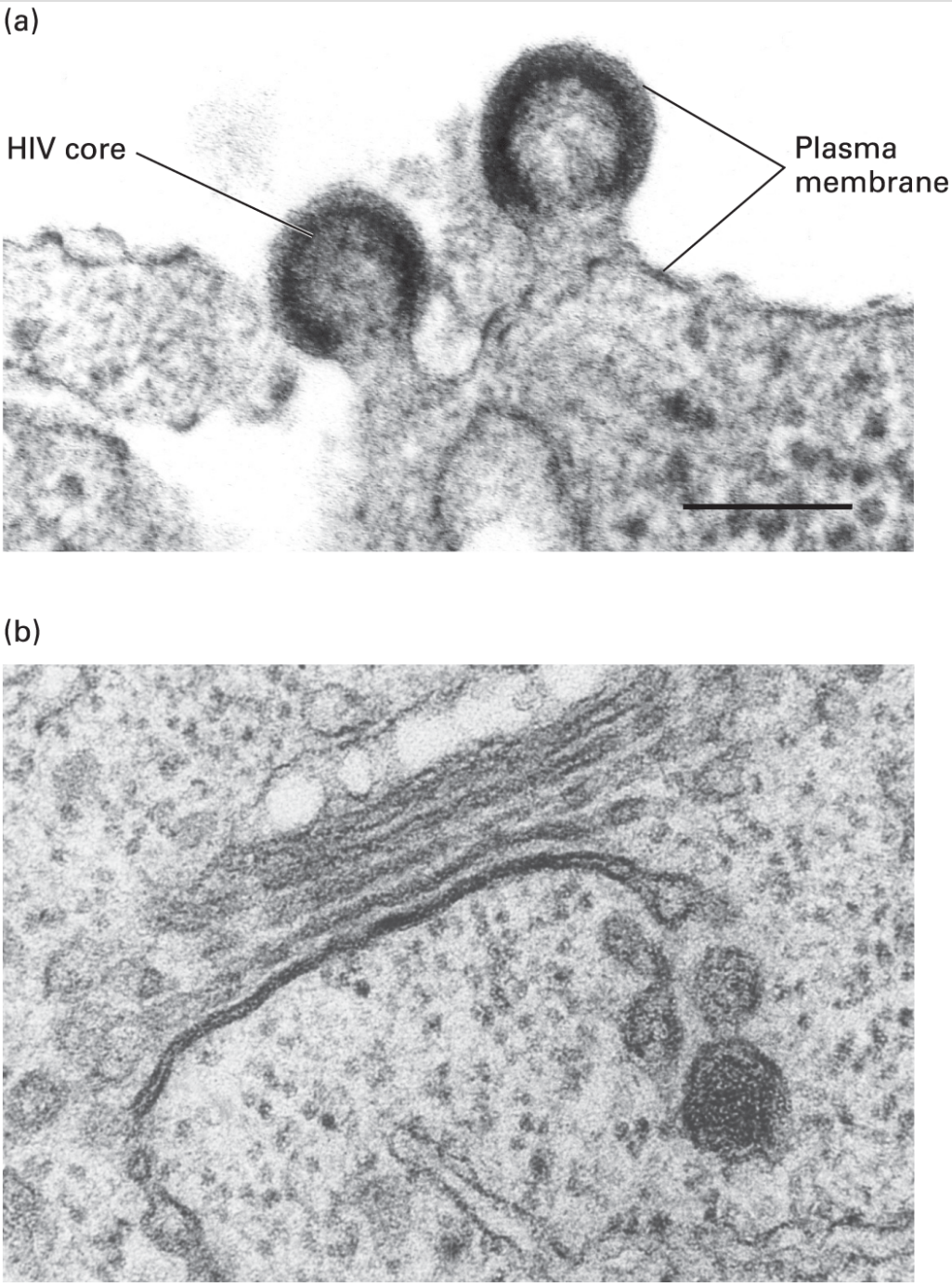
Prokaryotes, the simplest and smallest cells, are about 1–2 μm in length and are usually surrounded by a single plasma membrane; in most cases, they contain no internal membrane-limited subcompartments (see Figure 1-12). However, this single plasma

membrane contains hundreds of different types of proteins that are integral to the function of the cell. Some of these proteins catalyze ATP synthesis and initiation of DNA replication, for instance. Others include the many types of **membrane transport proteins** that enable specific ions, sugars, amino acids, and vitamins to cross the otherwise impermeable phospholipid bilayer to enter the cell and that allow specific metabolic products to exit. **Receptor proteins** in the plasma membrane are proteins that allow the cell to recognize chemical signals present in its environment and adjust its metabolism or pattern of gene expression in response. *Peripheral membrane proteins*, discussed in Section 10.2, associate with the membrane primarily by specific noncovalent interactions with integral membrane proteins or membrane lipids.

Eukaryotes also have a plasma membrane studded with many different proteins that perform a variety of functions, including membrane transport, cell signaling, and connecting cells into tissues. In addition, eukaryotic cells — which are generally much larger than prokaryotes — have a variety of internal membrane-bounded organelles (see Figure 1-13). Each type of organelle membrane has a unique complement of proteins that enable it to carry out its characteristic cellular functions, such as ATP generation (in mitochondria) and DNA synthesis (in the nucleus). Many plasma membrane proteins also bind components of the **cytoskeleton**, a dense network of protein filaments that crisscrosses the cytosol to provide mechanical support for cellular membranes.

These interactions are essential for the cell to assume and maintain its specific shape and for many types of cell movements.

Despite playing a structural role in cells, membranes are not rigid structures. They can bend and flex in three dimensions while still maintaining their integrity, in part because of the abundant noncovalent interactions that hold the lipids and proteins together. Moreover, there is considerable mobility of individual lipids and proteins within the plane of the membrane. According to the *fluid mosaic model* of biomembranes, first proposed by researchers in the 1970s, the phospholipid bilayer behaves in some respects like a two-dimensional fluid, with individual lipid molecules able to move past one another as well as spin in place. Such fluidity and flexibility not only allow organelles to assume their typical shapes but also confer on the membrane the dynamic property that enables membrane budding and fusion/fission, as occurs when viruses enter a cell or are released from an infected cell (Figure 10-2a). Internal cellular membranes of the Golgi complex bud into vesicles in the cytosol (Figure 10-2b) and then fuse with other membranes to transport their contents from one organelle to another (see Chapter 14).



Part (a) Courtesy of Wes Sundquist and Uta von Schwedler, University of Utah. Part (b) Biology Pics/Science Source.

FIGURE 10-2 Eukaryotic cellular membranes are dynamic structures. (a) An electron micrograph of the plasma membrane of an HIV-infected cell, showing HIV particles budding into the culture medium. As the virus core buds from the cell, it becomes enveloped by a membrane, derived from the cell's plasma membrane, that contains specific viral proteins. (b) Stacked membranes of the

Golgi complex with budding vesicles. Note the irregular shape and curvature of these membranes.

We begin our examination of biomembranes by considering their lipid components. These molecules not only affect membrane shape and function but also help anchor proteins to the membrane, modify membrane protein activities, and transduce signals to the cytoplasm. We then consider the structure of membrane proteins. Many of these proteins have large segments that are embedded in the hydrocarbon core of the phospholipid bilayer, and we focus on the principal classes of such transmembrane proteins. Finally, we consider how lipids such as phospholipids and cholesterol are synthesized in cells and distributed to their many membranes and organelles.

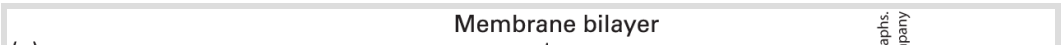
10.1 The Lipid Bilayer: Composition and Structural Organization

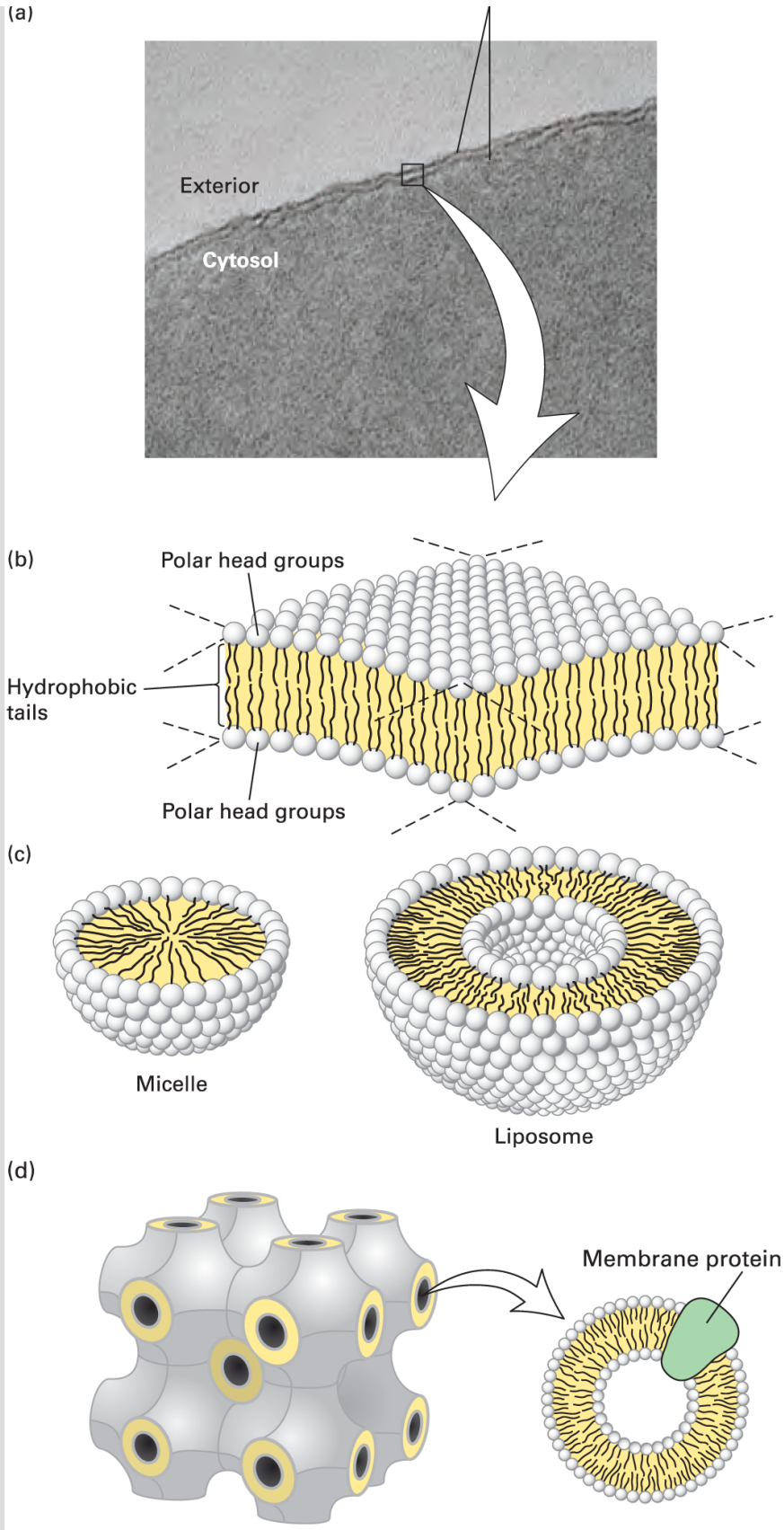
In Chapter 2, we learned that phospholipids are the principal building blocks of biomembranes. There are multiple types of phospholipids, but the most common phospholipids in membranes are the phosphoglycerides (see Figure 2-20). All phospholipids are **amphipathic** molecules, meaning that they consist of two segments with very different chemical properties: a fatty acid–based (fatty acyl) hydrocarbon "tail" that is **hydrophobic** ("water fearing") and partitions away from water, and a polar "head group" that is strongly **hydrophilic** ("water loving") and tends to interact with water molecules. The interactions of phospholipids with one another and with water largely determine the structure of biomembranes.

Besides phospholipids, biomembranes contain smaller amounts of other amphipathic lipids, such as glycolipids and cholesterol, which contribute to membrane function in important ways. We first consider the structure and properties of pure phospholipid bilayers and then discuss the composition and behavior of natural cellular membranes. Then we consider how the precise lipid composition of a given membrane influences its physical properties.

Phospholipids Spontaneously Form Bilayers

The amphipathic nature of phospholipids, which governs their interactions, is critical to the structure of biomembranes. When a suspension of phospholipids is mechanically dispersed (e.g., by ultrasound, or "sonication") in an aqueous solution, the phospholipids aggregate into one of three structures: spherical **micelles** or **liposomes**, or sheet-like **phospholipid bilayers** that are two molecules thick (Figure 10-3). The type of structure formed by pure phospholipids or a mixture of phospholipids in the absence of added proteins depends on several factors, including the length of the fatty acyl chains in the hydrophobic tails, their degree of saturation (i.e., the number of C—C and C=C bonds), and temperature. In all three structures, the hydrophobic effect (see Chapter 2) causes the fatty acyl chains to aggregate and exclude water molecules from the core of the structure. Micelles are rarely formed from natural phospholipids, whose fatty acyl chains are generally too bulky to fit into the interior of a micelle. Micelles are formed, however, if one of the two fatty acyl chains that make up the tail of a phospholipid is removed by hydrolysis to form a lysophospholipid, as occurs upon treatment with the enzyme phospholipase. In aqueous solutions, common detergents and soaps form micelles that behave like the balls in tiny ball bearings, thus giving soap solutions their slippery feel and lubricating properties.





Part (a) Warren Rosenberg/Fundamental Photogr
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Com

FIGURE 10-3 The bilayer structure of biomembranes. (a) Electron micrograph of a thin section through an erythrocyte membrane stained with osmium tetroxide. The characteristic "railroad track" appearance of the membrane indicates the presence of two polar layers, consistent with the bilayer structure of phospholipid membranes. (b) Schematic interpretation of the phospholipid bilayer. Polar groups face outward to shield the hydrophobic fatty acyl tails from water. The hydrophobic effect and van der Waals interactions between the fatty acyl tails drive the assembly of the bilayer (see Chapter 2). (c) Cross-sectional views of two other structures formed by dispersal of phospholipids in water. A spherical micelle has a hydrophobic interior composed entirely of fatty acyl chains; a spherical liposome consists of a phospholipid bilayer surrounding an aqueous center. (d) Under certain circumstances, lipids can assume yet other forms of organization. Shown here is the cubic phase of lipids, a highly regular recurring structure that has helped the formation of crystals of membrane proteins that were otherwise difficult to crystallize.

Phospholipid mixtures of the composition present in cells spontaneously form a symmetric phospholipid bilayer. Each phospholipid layer in this lamellar structure is called a *leaflet*. The hydrophobic fatty acyl chains in each leaflet minimize their contact with water by aligning themselves tightly together in the center of the bilayer, forming a hydrophobic core that is about 3–4 nm thick (see Figure 10-3b). The close packing of these nonpolar tails is stabilized by van der Waals interactions between the hydrocarbon chains. Ionic and hydrogen bonds stabilize the interactions of the polar head groups with one another and with water. Electron microscopy of thin sections of cells stained with osmium tetroxide, which binds strongly to the polar head groups of phospholipids, shows the bilayer structure (see Figure 10-3a). A cross section of a single membrane stained with osmium tetroxide looks like a

railroad track: two thin dark lines (the stained head group complexes) with a uniform light space of about 2 nm between them (the hydrophobic tails).

A phospholipid bilayer can be of almost unlimited size — from micrometers (μm) to millimeters (mm) in length or width — and can contain tens of millions of phospholipid molecules. The phospholipid bilayer is the basic structural unit of nearly all biological membranes. Its hydrophobic core prevents most water-soluble substances from crossing from one side of the membrane to the other. Although biomembranes contain other molecules (e.g., cholesterol, glycolipids, proteins), it is the phospholipid bilayer that separates two aqueous solutions and acts as a permeability barrier. The lipid bilayer thus defines cellular compartments and separates the cell's interior from the outside world.

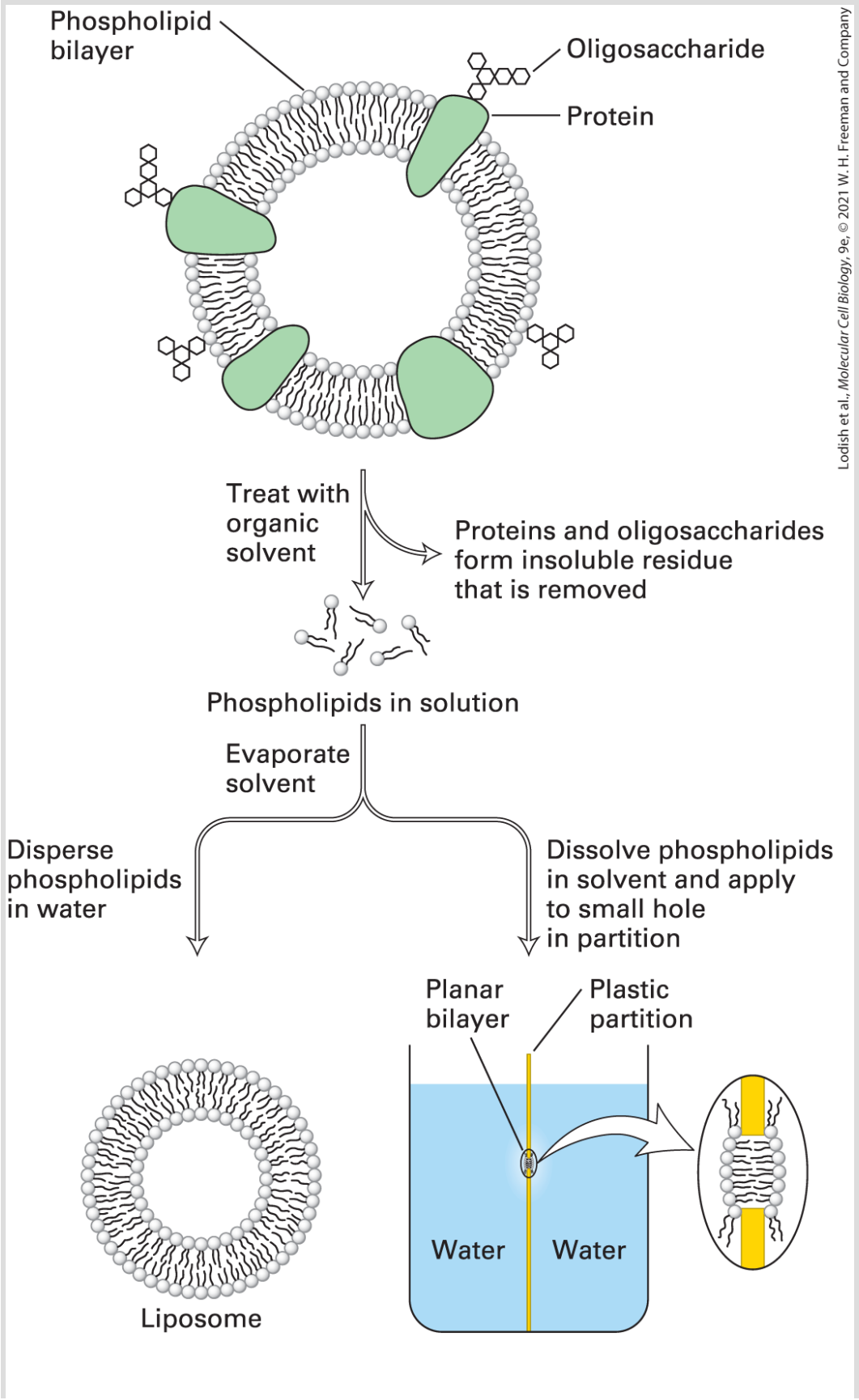
The three structures mentioned above are not the only forms that lipids can assume in an aqueous environment. Unusual configurations of lipids have been instrumental in enforcing order on otherwise difficult-to-crystallize membrane proteins, including G protein–coupled receptors, enabling crystallographic analysis of membrane proteins in a true lipid environment (Figure 10-3d).

Phospholipid Bilayers Form a Sealed Compartment Surrounding an Internal

Aqueous Space

Phospholipid bilayers can be generated in the laboratory by simple means, using either chemically pure phospholipids or lipid mixtures of the composition found in cellular membranes (Figure 10-4). Such synthetic bilayers possess three important properties. First, they are virtually impermeable to water-soluble (hydrophilic) solutes, which do not readily diffuse across the bilayer. These solutes include salts, sugars, and most other small hydrophilic molecules — including water itself. The second property of a bilayer is its stability. The hydrophobic and van der Waals interactions between the fatty acyl chains maintain the integrity of the interior of the bilayer structure. Even though the exterior aqueous environment may vary widely in ionic strength and pH, the bilayer has the strength to retain its characteristic architecture. Third, all synthetic phospholipid bilayers can spontaneously form sealed compartments in which the aqueous space on the inside is separated from the exterior environment. An edge of a phospholipid bilayer, as depicted in Figure 10-3b, with the hydrocarbon core of the bilayer exposed to an aqueous solution, would be unstable; the exposed fatty acyl chains would be in an energetically much more stable state if they were not adjacent to water molecules but rather surrounded by other fatty acyl chains. Thus, in an aqueous solution, sheets of phospholipid bilayers spontaneously seal their edges, forming a spherical bilayer that encloses an aqueous central compartment. The liposome depicted in Figure 10-3c is an example of such a structure viewed in cross section. Phospholipid bilayers can even be formed and

immobilized on a solid support such as a microscope slide, which allows the incorporation of membrane proteins whose behavior in the plane of the membrane can then be studied by various microscopical techniques.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 10-4 Formation and study of pure phospholipid bilayers. *Top:* A preparation of biological membranes is treated with an organic solvent, such as a mixture of chloroform and methanol (3:1), which selectively solubilizes the phospholipids and cholesterol. Proteins and carbohydrates remain in an insoluble residue. The solvent is removed by evaporation. *Bottom left:* If the extracted materials are mechanically dispersed in water, they spontaneously form a liposome, shown in cross section, with an internal aqueous compartment. *Bottom right:* A planar bilayer, also shown in cross section, can form over a small hole in a partition separating two aqueous phases; such a system can be used to study the physical properties of bilayers, such as their permeability to solutes.

This physical chemical property of a phospholipid bilayer has important implications for cellular membranes: no membrane in a cell can have an edge with exposed hydrocarbon fatty acyl chains. All membranes form closed compartments, similar in basic architecture to liposomes. Because all cellular membranes enclose an entire cell or an internal compartment, they have an *internal face* (the surface oriented toward the interior of the compartment) and an *external face* (the surface presented to the environment). More commonly, we designate the two surfaces of a cellular membrane as the **cytosolic face** and the **exoplasmic face**. This nomenclature is useful in highlighting the topological equivalence of the faces in different membranes, as diagrammed in Figures 10-5 and 10-6. For example, the exoplasmic face of the plasma membrane is directed away from the cytosol, toward the extracellular space or external environment, and defines the outer limit of the cell. The cytosolic face of the plasma membrane faces the cytosol.

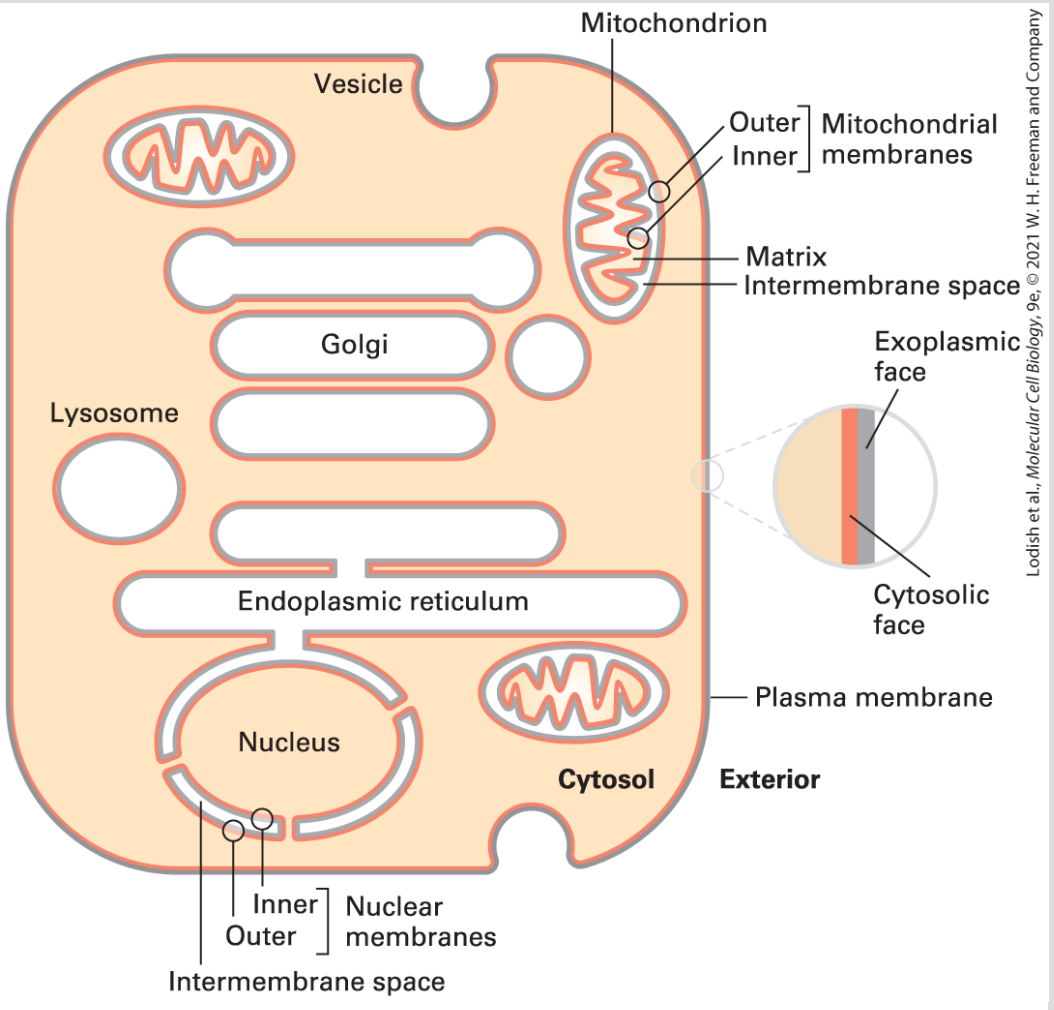


FIGURE 10-5 The faces of cellular membranes. The plasma membrane, a single bilayer, encloses the cell. In this highly schematic representation, internal cytosol (tan) and external environment (white) define the cytosolic (red) and exoplasmic (gray) faces of the bilayer. Vesicles and some organelles have a single membrane, and their internal aqueous space (white) is topologically equivalent to the outside of the cell. Three organelles — the nucleus, mitochondrion, and chloroplast (not shown) — are enclosed by two membranes separated by a small intermembrane space. The exoplasmic faces of the inner and outer membranes around these organelles border the intermembrane space between them. For simplicity, the hydrophobic membrane interior is not indicated in this diagram.

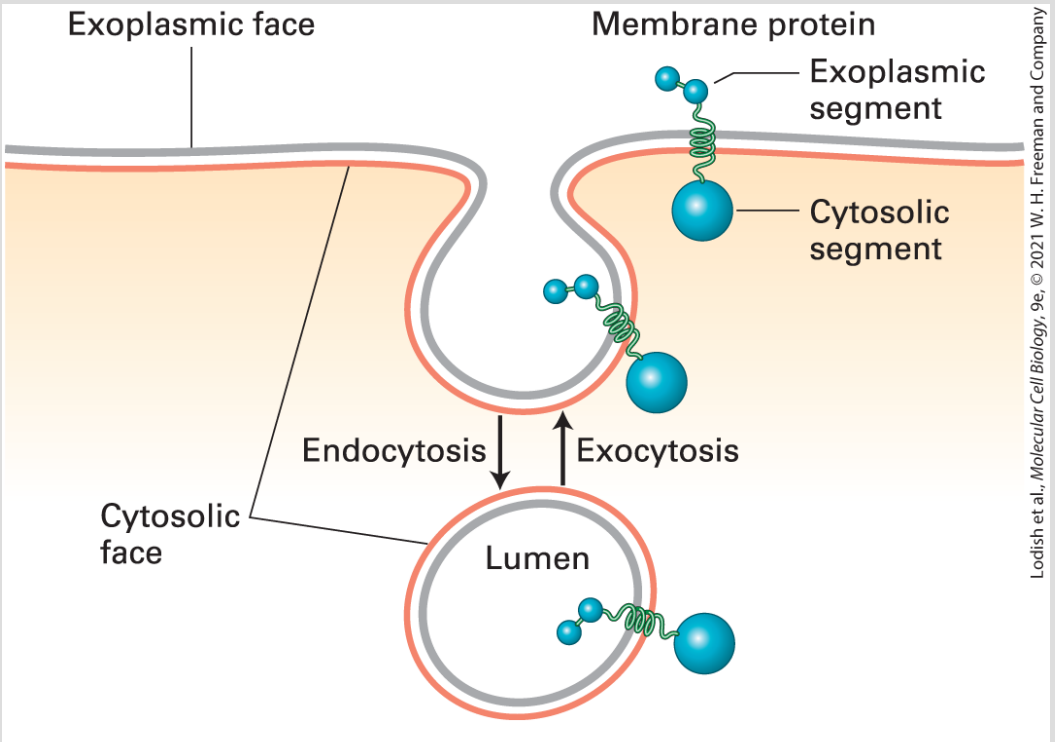


FIGURE 10-6 The faces of cellular membranes are conserved during membrane budding and fusion. Red membrane surfaces are cytosolic faces; gray membrane surfaces are exoplasmic faces. During endocytosis, a segment of the plasma membrane buds inward toward the cytosol and eventually pinches off a separate vesicle. During this process, the cytosolic face of the plasma membrane remains facing the cytosol, and the exoplasmic face of the new vesicle membrane faces the vesicle lumen. During exocytosis, an intracellular vesicle fuses with the plasma membrane, and the lumen of the vesicle (exoplasmic face) connects with the extracellular medium. Proteins that span the membrane retain their asymmetric orientation during vesicle budding and fusion; in particular, the same segment always faces the cytosol.

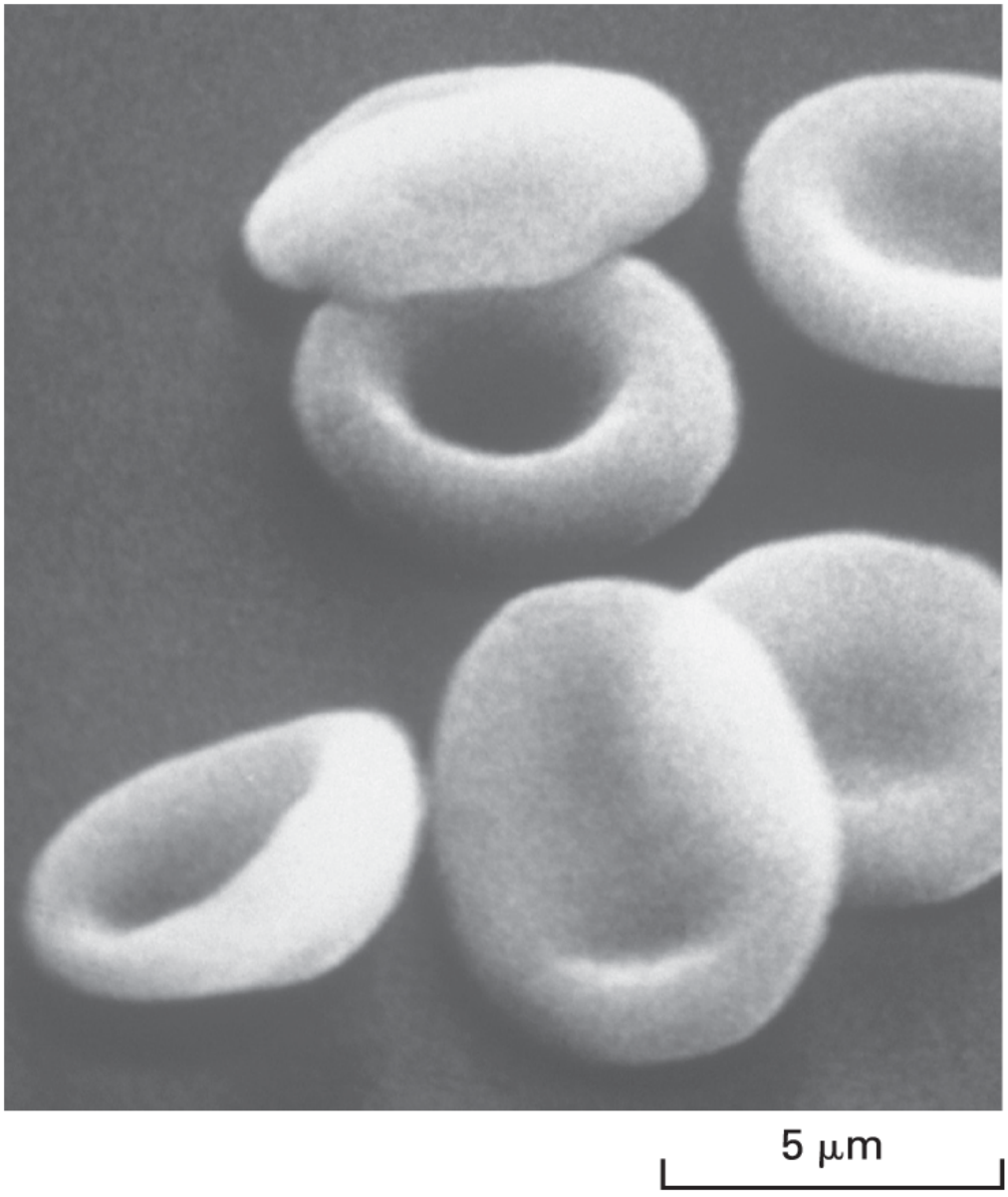
Similarly for organelles and vesicles surrounded by a single membrane, the cytosolic face is directed toward the cytosol. The exoplasmic face is always directed away from the cytosol, and in this case it is on the inside of the organelle, in contact with the internal aqueous space, or *lumen*. The lumen of a vesicle is topologically

equivalent to the extracellular space, a concept most easily understood for vesicles that arise by invagination (endocytosis) of the plasma membrane. The external face of the plasma membrane becomes the internal face of the vesicle membrane, while in the vesicle the cytosolic face of the plasma membrane still faces the cytosol (see Figure 10-6).

Three organelles — the nucleus, mitochondrion, and chloroplast — are surrounded not by a single membrane, but by two. The exoplasmic surface of each membrane faces the space between the two membranes (the intermembrane space). This relationship can perhaps be best understood by reference to the *endosymbiont hypothesis*, discussed in Chapter 12, which posits that mitochondria and chloroplasts arose early in the evolution of eukaryotic cells by the engulfment of bacteria capable of oxidative phosphorylation or photosynthesis, respectively (see Figure 12-7). Gram-negative bacteria such as *Escherichia coli* are likewise surrounded by two membranes (see Figure 1-12).

Natural membranes from different cell types exhibit a variety of shapes, which complement a cell's function. The smooth, flexible surface of the erythrocyte plasma membrane allows the cell to squeeze through narrow blood capillaries (Figure 10-7). Some cells have a long, slender extension of the plasma membrane, called a **cilium** or **flagellum**, which beats in a whiplike manner (see Figure 1-15). This motion causes fluid to flow across the surface of a sheet of

cells or a sperm cell to swim toward an egg. The differing shapes and properties of biomembranes raise a key question in cell biology, namely, how the composition of biological membranes is regulated to establish and maintain the identity of the different membrane structures and membrane-limited compartments. We return to this question in Section 10.3 and in Chapter 14.



Omikron/Science Source.

FIGURE 10-7 Variation in biomembranes in different cell types. A smooth, flexible membrane covers the surface of the discoid erythrocyte cell, as seen in this scanning electron micrograph.

Biomembranes Contain Three Principal Classes of Lipids

The term *phospholipid* refers to any amphipathic lipid with a phosphate-based head group and a two-chain hydrophobic tail. A typical biomembrane is not composed of phospholipids alone, but actually contains three classes of amphipathic lipids: phosphoglycerides, sphingolipids, and sterols, which differ in their chemical structures, abundance, and functions in the membrane (Figure 10-8). While all phosphoglycerides are phospholipids, only certain sphingolipids are, and no sterols are.

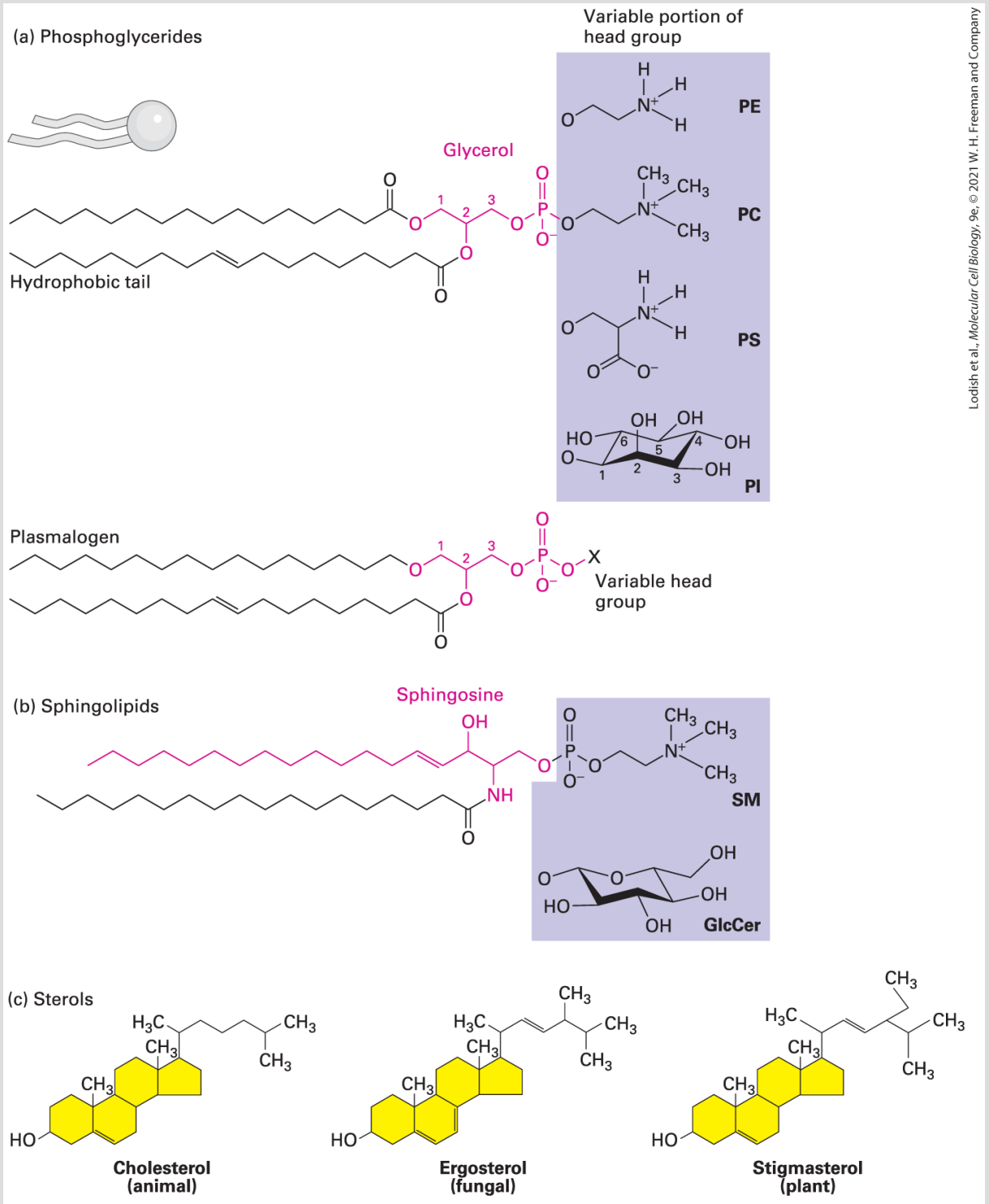


FIGURE 10-8 Three classes of membrane lipids. (a) Most phosphoglycerides are derivatives of glycerol 3-phosphate (red), which contains two esterified fatty acyl chains that constitute the hydrophobic "tail" and a polar "head group" esterified to the phosphate. The fatty acids can vary in length and be saturated (no double bonds) or unsaturated (one, two, or three

double bonds). In phosphatidylcholine (PC), the head group is choline. Also shown are the molecules attached to the phosphate group in three other common phosphoglycerides: phosphatidylethanolamine (PE), phosphatidylserine (PS), and phosphatidylinositol (PI). Plasmalogens contain one fatty acyl chain attached to glycerol by an ester linkage and one attached by an ether linkage; they contain the same head groups as other phosphoglycerides. (b) Sphingolipids are derivatives of sphingosine (red), an amino alcohol with a long hydrocarbon chain. Various fatty acyl chains are connected to sphingosine by an amide bond. The sphingomyelins (SM), which contain a phosphocholine head group, are phospholipids. Other sphingolipids are glycolipids in which a single sugar residue or branched oligosaccharide is attached to the sphingosine backbone. For instance, the simple glycolipid glucosylcerebroside (GlcCer) has a glucose head group. (c) The major sterols in animals, fungi, and plants differ slightly in structure, but all serve as key components of cellular membranes. The basic structure of sterols is a four-ring hydrocarbon (yellow). Like other membrane lipids, sterols are amphipathic. The single hydroxyl group is equivalent to the polar head group in other lipids; the conjugated ring and short hydrocarbon chain form the hydrophobic tail. See H. Sprong, P. van der Sluijs, and G. van Meer, 2001, *Nature Rev. Mol. Cell Biol.* **2**:504–513.

Phosphoglycerides, the most abundant class of phospholipids in most membranes are derivatives of glycerol 3-phosphate (see Figure 10-8a). A typical phosphoglyceride molecule consists of a hydrophobic tail composed of two fatty acyl chains esterified to the two hydroxyl groups in glycerol phosphate and a polar head group attached to the phosphate group. The structure comprising the glycerol moiety and the two fatty acyl chains is referred to as a *diacylglycerol*. The two fatty acyl chains may differ in the number of carbons that they contain (commonly 16 or 18) and their degree of saturation (0, 1, or 2 double bonds). A phosphoglyceride is classified according to the nature of its head group. In phosphatidylcholines, the most abundant phospholipids in the plasma membrane, the

head group consists of choline, a positively charged alcohol, esterified to the negatively charged phosphate. In other phosphoglycerides, an OH-containing molecule such as ethanolamine, serine, or the sugar derivative inositol is linked to the phosphate group. In the case of the inositol head group, the hydroxyl groups may be further modified with phosphates, yielding a class of phospholipids called *phosphoinositides*. The phosphoinositides fulfill an important function in signal transduction, as discussed in Chapters 15 and 16. The negatively charged phosphate group and the positively charged groups or hydroxyl groups on the head group interact strongly with water. At neutral pH, some phosphoglycerides (e.g., phosphatidylcholine and phosphatidylethanolamine) carry no net electric charge, whereas others (e.g., phosphatidylinositol and phosphatidylserine) carry a single net negative charge. Nonetheless, the polar head groups in all these phospholipids can pack together into the characteristic bilayer structure. When phospholipases act on phosphoglycerides, they produce lysophospholipids, which lack one of the two acyl chains. Lysophospholipids are not only important signaling molecules, released from cells and recognized by specific receptors; their presence can also affect the physical properties of the membranes in which they reside.

The *plasmalogens* are a group of phosphoglycerides that contain one fatty acyl chain attached to carbon 2 of glycerol by an ester linkage and one long hydrocarbon chain attached to carbon 1 of glycerol by

an ether (C—O—C) rather than an ester linkage. Plasmalogens are particularly abundant in human brain and heart tissue. The greater chemical stability of the ether linkage in plasmalogens compared with the ester linkage, and the subtle differences in three-dimensional structure between plasmalogens and other phosphoglycerides, may have as yet unrecognized physiological significance.

A second class of membrane lipids is the **sphingolipids**. All these compounds are derived from sphingosine, an amino alcohol with a long hydrocarbon chain. Sphingolipids contain a long-chain fatty acid attached in amide linkage to the sphingosine amino group (see Figure 10-8b). Like phosphoglycerides, some sphingolipids have a phosphate-based polar head group. In sphingomyelin, the most abundant sphingolipid, phosphocholine is attached to the terminal hydroxyl group of sphingosine (see Figure 10-8b; SM). Thus sphingomyelin is a phospholipid, and its overall structure is quite similar to that of phosphatidylcholine. Sphingomyelins are similar in shape to phosphoglycerides and can form mixed bilayers with them. Other sphingolipids are amphipathic **glycolipids** whose polar head groups are sugars that are not linked to the tails via a phosphate group (so they are not phospholipids). Glucosylcerebroside, the simplest glycosphingolipid, contains a single glucose unit attached to sphingosine. In the complex glycosphingolipids called *gangliosides*, one or two branched sugar chains (oligosaccharides) containing sialic acid groups are attached

to sphingosine. Glycolipids constitute 2–10 percent of the total lipid content of plasma membranes; they are most abundant in nervous tissue.

Cholesterol and its analogs constitute the third important class of membrane lipids, the *sterols*. The basic structure of sterols is a four-ring isoprenoid-based hydrocarbon. The structures of the principal yeast sterol (ergosterol) and plant phytosterols (e.g., stigmasterol) differ slightly from that of cholesterol, the major animal sterol (see Figure 10-8c). The small differences in the biosynthetic pathways and structures of fungal and animal sterols are the basis of most antifungal drugs currently in use. Cholesterol, like the two other sterols, has a hydroxyl substituent on one ring. Although cholesterol is almost entirely hydrocarbon in composition, it is amphipathic because its hydroxyl group can interact with water. Because it lacks a phosphate-based head group, it is not a phospholipid. Cholesterol is especially abundant in the plasma membranes of mammalian cells but is absent from most prokaryotic and all plant cells. As much as 30–50 percent of the lipids in plant plasma membranes consist of certain steroids unique to plants. Between 50 and 90 percent of the cholesterol in most mammalian cells is present in the plasma membrane and associated vesicles. Cholesterol and other sterols are too hydrophobic to form a bilayer structure on their own. Instead, at the concentrations found in natural membranes, these sterols must intercalate between phospholipid molecules to be incorporated into biomembranes. When so intercalated, sterols

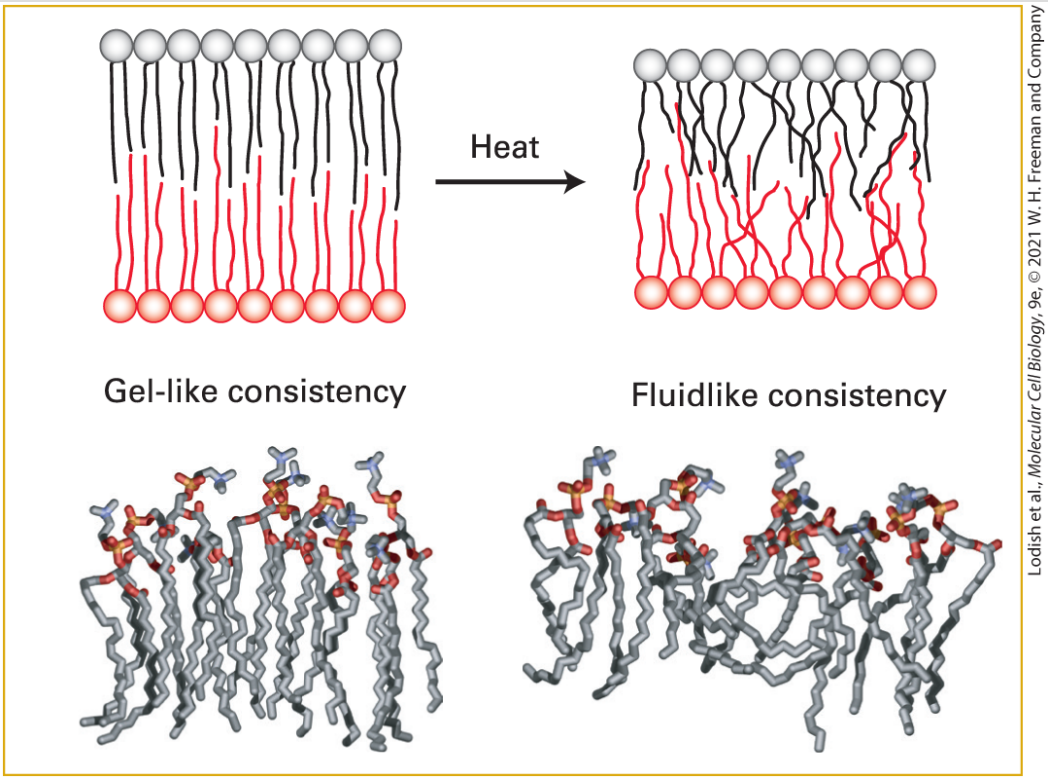
provide structural support to membranes, preventing too close a packing of the phospholipids' acyl chains to maintain a significant measure of membrane fluidity and at the same time conferring the necessary rigidity required for mechanical support. Some of these effects can be highly local, as in the case of lipid rafts, discussed below.

In addition to its structural role in membranes, cholesterol is the precursor for several important bioactive molecules. They include *bile acids*, which are made in the liver and help emulsify dietary fats for digestion and absorption in the intestines; steroid hormones produced by endocrine cells (e.g., adrenal gland, ovary, testes); and vitamin D produced in the skin and kidneys. Another critical function of cholesterol is its covalent addition to Hedgehog protein, a key signaling molecule in embryonic development (see Chapter 16).

Most Lipids and Many Proteins Are Laterally Mobile in Biomembranes

In the two-dimensional plane of a bilayer, thermal motion permits lipid molecules to rotate freely around their long axes and to diffuse laterally within each leaflet. Because such movements are lateral or rotational, the fatty acyl chains remain in the hydrophobic interior of the bilayer. In both natural and artificial membranes, a typical

lipid molecule exchanges places with its neighbors in a leaflet about 10^7 times per second and diffuses several micrometers per second at 37 °C. These diffusion rates indicate that the bilayer is 100 times more viscous than water — about the same as the viscosity of olive oil. Even though lipids diffuse more slowly in the bilayer than in an aqueous solvent, a membrane lipid could diffuse the length of a typical bacterial cell (1 μm) in only 1 second and the length of an animal cell in about 20 seconds. When artificial pure phospholipid membranes are cooled below 37 °C, the lipids can undergo a *phase transition* from a liquid-like (fluid) state to a gel-like (semisolid) state, analogous to the liquid-solid transition when molten butter solidifies upon cooling or liquid water freezes (Figure 10-9). Below the phase-transition temperature, the rate of diffusion of the lipids drops precipitously. At usual physiological temperatures, the hydrophobic interior of natural membranes generally has a low viscosity and a fluidlike consistency, in contrast to the gel-like consistency observed at lower temperatures.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

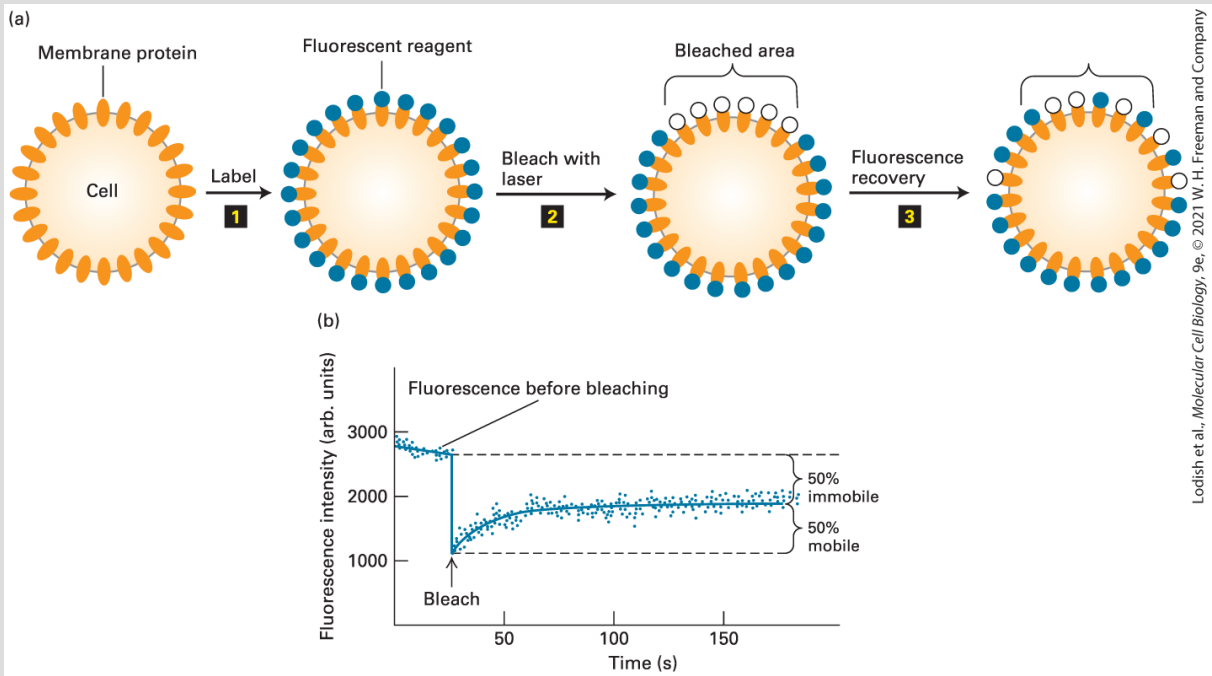
FIGURE 10-9 Gel-like and fluidlike forms of the phospholipid bilayer. *Top:* Depiction of gel-to-fluid transition. Phospholipids with long saturated fatty acyl chains tend to assemble into a highly ordered, gel-like bilayer in which there is little overlap of the nonpolar tails in the two leaflets. Heat disorders the nonpolar tails and induces a transition from a gel to a fluid within a temperature range of only a few degrees. As the chains become disordered, the bilayer also decreases in thickness. *Bottom:* Molecular models of phospholipid monolayers in gel-like and fluidlike states, as determined by molecular dynamics calculations.

[Data from H. Heller, M. Schaefer, and K. Schulten, 1993, *J. Phys. Chem.* **97**:8343.]

In pure lipid bilayers (i.e., in the absence of protein), phospholipids and sphingolipids rotate and move laterally, but they do not usually spontaneously migrate, or flip-flop, from one leaflet to the other. The energetic barrier is too high; migration would require moving

the polar head group from its aqueous environment through the hydrocarbon core of the bilayer to the aqueous solution on the other side. Special membrane proteins discussed in Chapter 11 are required to flip membrane lipids and other polar molecules from one leaflet to the other.

The lateral movements of specific plasma-membrane proteins and lipids can be quantified by a technique called *fluorescence recovery after photobleaching (FRAP)*. Phospholipids containing a fluorescent substituent are used to monitor lipid movement. For proteins, a fragment of a monoclonal antibody (see Chapter 23) that is specific for the exoplasmic domain of the desired protein and that has only a single antigen-binding site is tagged with a fluorescent dye (alternatively, a genetic fusion of the desired protein with fluorescent green protein can accomplish the same goal). With this method, described in Figure 10-10, the rate at which membrane molecules move — the diffusion coefficient — can be determined, as well as the proportion of the molecules that are laterally mobile.



EXPERIMENTAL FIGURE 10-10 Fluorescence recovery after photobleaching (FRAP) experiments can quantify the lateral movement of proteins and lipids within the plasma membrane. (a) Experimental protocol. Step **1**: Cells are first labeled with a fluorescent reagent that binds uniformly to a specific membrane lipid or protein. Genetic fusion of the desired protein with green fluorescent protein (GFP) can accomplish the same goal. Step **2**: A laser light is then focused on a small area of the cell surface, irreversibly bleaching the bound reagent (or GFP) and thus reducing the fluorescence in the illuminated area. Step **3**: In time, the fluorescence of the bleached patch increases as unbleached fluorescent surface molecules diffuse into it and bleached ones diffuse outward. The extent of recovery of fluorescence in the bleached patch is proportional to the fraction of labeled molecules that are mobile in the membrane. (b) Results of a FRAP experiment with human hepatoma cells treated with a fluorescent antibody specific for the asialoglycoprotein receptor protein. The finding that 50 percent of the fluorescence returned to the bleached area indicates that 50 percent of the receptor molecules in the illuminated membrane patch were mobile and 50 percent were immobile. Because the rate of fluorescence recovery is proportional to the rate at which labeled molecules move into the bleached region, the diffusion coefficient of a protein or lipid in the membrane can be calculated from such data.

[Data from Y. I. Henis et al., 1990, *J. Cell Biol.* **111**:1409.]

The results of FRAP studies with fluorescence-labeled phospholipids have shown that in fibroblast plasma membranes, all phospholipids are freely mobile over distances of about 0.5 μm, but most cannot diffuse over much longer distances. These findings suggest that protein-rich regions of the plasma membrane about 1 μm in diameter separate lipid-rich regions containing the bulk of the membrane phospholipids. Phospholipids are free to diffuse within such regions, but not from one lipid-rich region to an adjacent one. Furthermore, the rate of lateral diffusion of lipids in the plasma membrane is nearly an order of magnitude slower than in pure phospholipid bilayers: diffusion constants of 10^{-8} cm^2 per second and 10^{-7} cm^2 per second are characteristic of the plasma membrane and a pure phospholipid bilayer, respectively. This difference suggests that lipids may be tightly but not irreversibly bound to certain integral proteins in some membranes, as indeed has recently been demonstrated (see the discussion of annular phospholipids below).

Lipid Composition Influences the Physical Properties of Membranes

A typical cell contains many different types of membranes, each with unique properties derived from its particular mix of lipids and proteins. The data in Table 10-1 illustrate the variation in lipid composition in different biomembranes. Several phenomena

contribute to these differences. For instance, the relative abundances of phosphoglycerides and sphingolipids differ between membranes in the endoplasmic reticulum (ER), where phospholipids are synthesized, and the Golgi complex, where sphingolipids are synthesized. The proportion of sphingomyelin as a percentage of total membrane lipid phosphorus is about six times as high in Golgi membranes as it is in ER membranes. In other cases, the movement of membranes from one cellular compartment to another can selectively enrich certain membranes in lipids such as cholesterol. In responding to differing environments throughout an organism, different types of cells generate membranes with differing lipid compositions. In the cells that line the intestinal tract, for example, the membranes that face the harsh environment in which dietary nutrients are digested have a sphingolipid-to-phosphoglyceride-to-cholesterol ratio of 1:1:1, rather than the 0.5:1.5:1 ratio found in cells subject to less stress. The relatively high concentration of sphingolipids in these intestinal membranes may increase their stability because of extensive hydrogen bonding by the free −OH group in the sphingosine moiety (see Figure 10-8).

TABLE 10-1 • Major Lipid Components of Selected Biomembranes

	Composition (mol %)			
Source/Location	**PC**	**PE + PS**	**SM**	**Cholesterol**
Plasma membrane (human erythrocytes)	21	29	21	26
Myelin membrane (human	16	37	13	34

neurons)				
Plasma membrane (mung bean)	47	43	0	0
Inner mitochondrial membrane (cauliflower)	42	38	0	0
Outer mitochondrial membrane (cauliflower)	47	27	0	0
Plasma membrane *(E. coli)*	0	85	0	0
Endoplasmic reticulum membrane (rat)	60	25	3	7
Golgi membrane (rat)	51	26	8	13
Inner mitochondrial membrane (rat)	40	37	2	7
Outer mitochondrial membrane (rat)	54	31	2	11
Primary leaflet location	Exoplasmic	Cytosolic	Exoplasmic	Both

PC = phosphatidylcholine; PE = phosphatidylethanolamine; PS = phosphatidylserine, SM = sphingomyelin.
SOURCE: Data from S. E. Horvath and G. Daum, 2013, "Lipids of Mitochondria," *Prog. Lipid Res.* **52**:590–614.

The degree of bilayer fluidity depends on lipid composition, the structure of the phospholipid hydrophobic tails, and temperature. As already noted, van der Waals interactions and the hydrophobic effect cause the nonpolar tails of phospholipids to aggregate. Long,

saturated fatty acyl chains have the greatest tendency to aggregate, packing tightly together into a gel-like state. Phospholipids with short fatty acyl chains, which have less surface area and therefore fewer van der Waals interactions, form more fluid bilayers. Likewise, the kinks in cis-unsaturated fatty acyl chains (see Chapter 2) result in their forming less stable van der Waals interactions with other lipids, and hence more fluid bilayers, than do straight saturated chains, which can pack more tightly together.

Cholesterol is important in maintaining the appropriate fluidity of natural membranes, a property that appears to be essential for normal cell growth and reproduction. Cholesterol restricts the random movement of phospholipid head groups at the outer surfaces of the leaflets, but its effect on the movement of long phospholipid tails depends on its concentration. At the cholesterol concentrations normally present in the plasma membrane, the interaction of the steroid ring with the long hydrophobic tails of phospholipids tends to immobilize those lipids and thus decreases biomembrane fluidity. It is this property that can help organize the plasma membrane into discrete subdomains of unique lipid and protein composition. At lower cholesterol concentrations, however, the steroid ring separates and disperses phospholipid tails, causing the inner regions of the membrane to become slightly more fluid.

The lipid composition of a bilayer also influences its thickness, which in turn may influence the distribution of other membrane components, such as proteins, in a particular membrane. It has

been argued that relatively short transmembrane segments of certain Golgi-resident enzymes (glycosyltransferases) are an adaptation to the lipid composition of the Golgi membrane and contribute to the retention of these enzymes in the Golgi apparatus. The results of biophysical studies on artificial membranes demonstrate that sphingomyelin associates into a more gel-like and thicker bilayer than phosphoglycerides do (Figure 10-11a). Cholesterol and other molecules that decrease membrane fluidity also increase membrane bilayer thickness. Because sphingomyelin tails are already optimally stabilized, the addition of cholesterol has no effect on the thickness of a sphingomyelin bilayer.

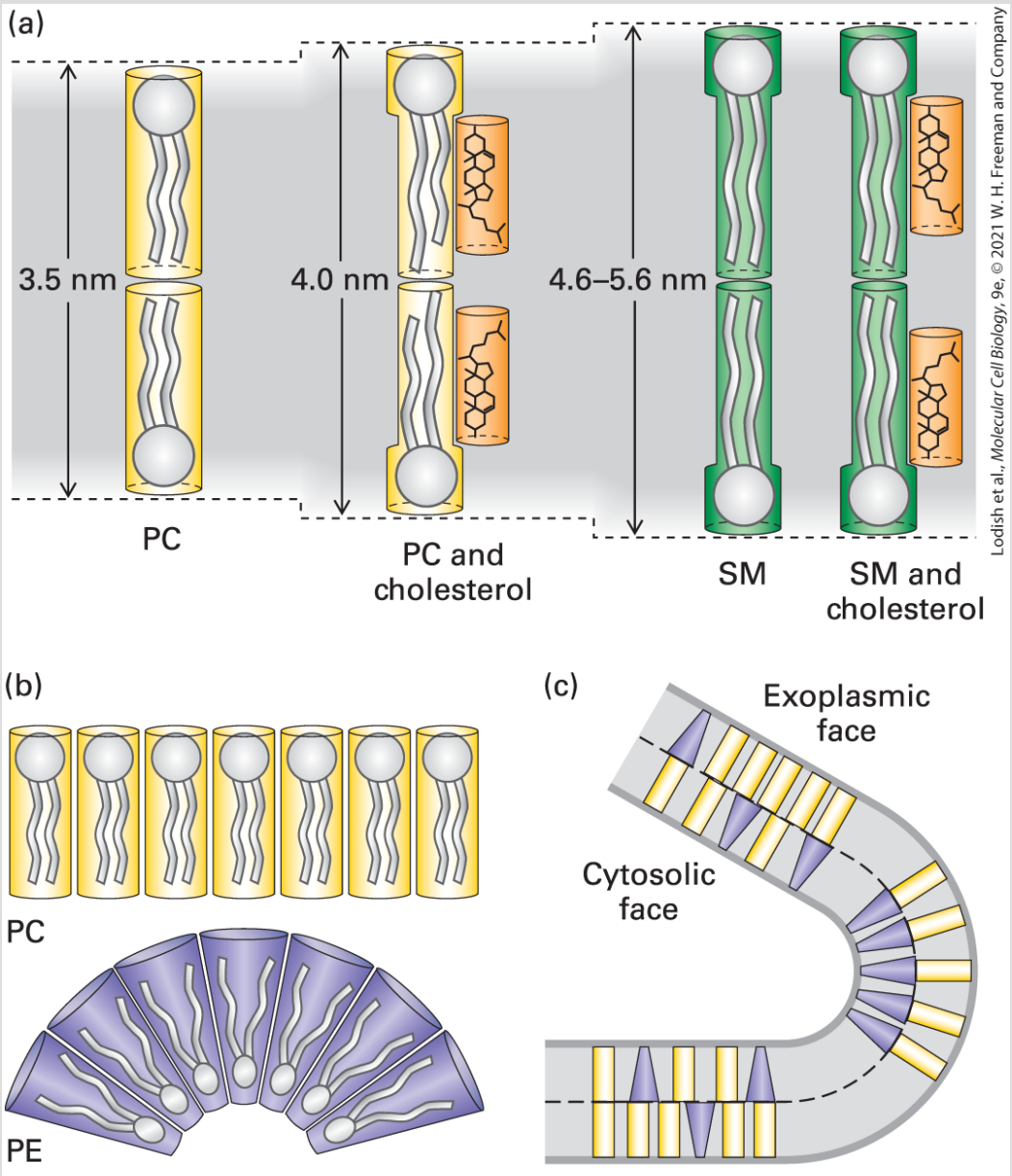


FIGURE 10-11 Effect of lipid composition on bilayer thickness and curvature. (a) A pure sphingomyelin (SM) bilayer is thicker than one formed from a phosphoglyceride such as phosphatidylcholine (PC). Cholesterol has a lipid-ordering effect on phosphoglyceride bilayers that increases their thickness, but it does not affect the thickness of the more ordered SM bilayer. (b) Phospholipids such as PC have a cylindrical shape and form essentially flat monolayers, whereas those with smaller head groups, such as phosphatidylethanolamine (PE), have a conical shape. (c) A bilayer enriched

with PC in the exoplasmic leaflet and with PE in the cytosolic face, as in many plasma membranes, would have a natural curvature.

[Information from H. Sprong et al., 2001, *Nat. Rev. Mol. Cell Biol.* **2**:504–513.]

Another property dependent on the lipid composition of a bilayer is its curvature, which depends on the relative sizes of the polar head groups and nonpolar tails of its constituent phospholipids. Lipids with long tails and large head groups are cylindrical in shape; those with small head groups are cone-shaped (Figure 10-11b). As a result, bilayers composed of cylindrical lipids are relatively flat, whereas those containing large numbers of cone-shaped lipids form curved bilayers (Figure 10-11c). This effect of lipid composition on bilayer curvature may play a role in the formation of highly curved membranes, such as sites of viral budding (see Figure 10-2) and of formation of internal vesicles from the plasma membrane (see Figure 10-6), and in specialized stable membrane structures such as microvilli. Several proteins, including *reticulons* (RTNs in vertebrates; reticulon-like proteins or RTNLs in other eukaryotes) bind to the surface of a phospholipid bilayer and cause the membrane to curve; such proteins are important in formation of transport vesicles that bud from a donor membrane (see Chapter 14).

Lipid Composition Is Different in the Exoplasmic and Cytosolic Leaflets

A characteristic of all biomembranes is an asymmetry in lipid composition across the bilayer. Although most phospholipids are present in both membrane leaflets, some are more abundant in one or the other leaflet. For instance, in plasma membranes from human erythrocytes and Madin-Darby canine kidney (MDCK) cells grown in culture, almost all the sphingomyelin and phosphatidylcholine, both of which form less fluid bilayers, are found in the exoplasmic leaflet. In contrast, phosphatidylethanolamine, phosphatidylserine, and phosphatidylinositol, which form more fluid bilayers, are preferentially located in the cytosolic leaflet. Because phosphatidylserine and phosphatidylinositol carry a net negative charge, the stretch of amino acids on the cytosolic face of a membrane protein with a single transmembrane segment is often enriched in positively charged (Lys, Arg) residues in close proximity to the transmembrane segment (this distribution of charged amino acids is referred to as the "inside positive" rule). This segregation of lipids across the bilayer may influence membrane curvature (see Figure 10-11c). Unlike particular phospholipids, cholesterol is relatively evenly distributed in both leaflets of cellular membranes.

The relative abundances of a particular phospholipid in the two leaflets of a plasma membrane can be determined experimentally on the basis of the susceptibility of phospholipids to hydrolysis by **phospholipases**, enzymes that cleave the ester bonds via which acyl chains and head groups are connected to the lipid molecule (Figure 10-12). When added to the external medium, phospholipases cannot

cross the membrane, and thus they cleave off the head groups of only those lipids present in the exoplasmic face; phospholipids in the cytosolic leaflet are resistant to hydrolysis because the enzymes cannot penetrate to the cytosolic face of the plasma membrane.

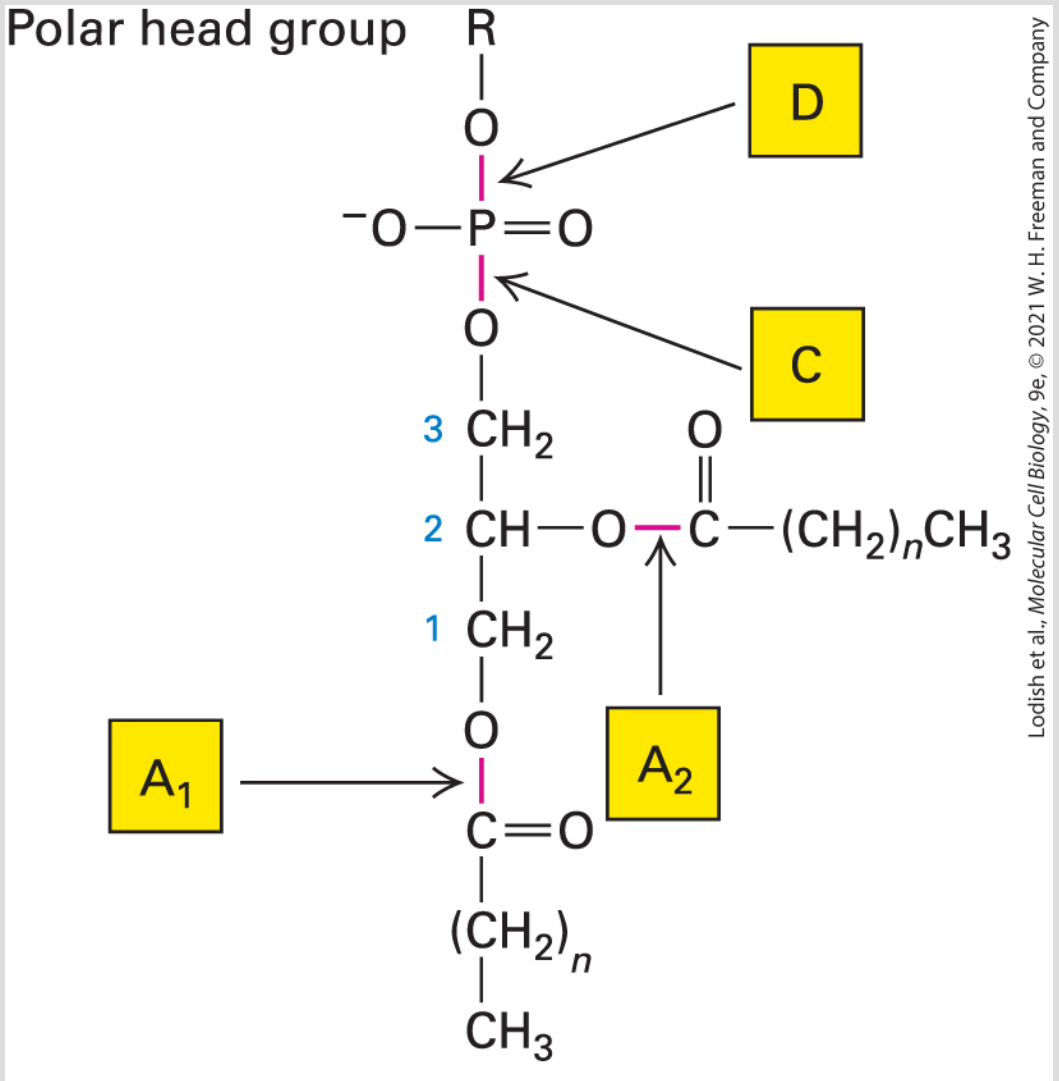


FIGURE 10-12 Specificity of phospholipases. Each type of phospholipase cleaves one of the susceptible bonds shown in red. The glycerol carbon atoms are indicated by small numbers. In intact cells, only phospholipids in the exoplasmic leaflet of the plasma membrane are cleaved by phospholipases in the surrounding medium. Phospholipase C, a cytosolic enzyme, cleaves certain phospholipids in the cytosolic leaflet of the plasma membrane.

How the asymmetric distribution of phospholipids in membrane leaflets arises and is maintained is still unclear. As noted, in pure bilayers, phospholipids do not spontaneously migrate, or flip-flop, from one leaflet to the other. In part, the asymmetry in phospholipid distribution may reflect where these lipids are synthesized in the ER and Golgi. Sphingomyelin is synthesized on the luminal (exoplasmic) face of the Golgi, which becomes the exoplasmic face of the plasma membrane. In contrast, phosphoglycerides are synthesized on the cytosolic face of the ER membrane, which is topologically equivalent to the cytosolic face of the plasma membrane (see Figure 10-5). Clearly, however, this explanation does not account for the preferential location of phosphatidylcholine (a phosphoglyceride) in the exoplasmic leaflet. Movement of this phosphoglyceride, and perhaps others, from one leaflet to the other in some natural membranes is most likely catalyzed by ATP-powered transport proteins called *flippases*, which are discussed in Chapter 11.

The preferential location of lipids on one face of the bilayer is necessary for a variety of membrane-based functions. For example, the head groups of all phosphorylated forms of phosphatidylinositol (see Figure 10-8b; PI), an important source of second messengers, face the cytosol. Stimulation of many cell-surface receptors by their corresponding ligands results in activation of the cytosolic enzyme phospholipase C, which can then hydrolyze the bond within PI

connecting the phosphoinositols to the diacylglycerol. As we will see in Chapter 15, both water-soluble phosphoinositols and membrane-embedded diacylglycerol participate in intracellular signaling pathways that affect many aspects of cellular metabolism. Phosphatidylserine, too, is normally most abundant in the cytosolic leaflet of the plasma membrane. In the initial stages of platelet stimulation by serum, phosphatidylserine is briefly translocated to the exoplasmic face, presumably by a flippase enzyme, where it activates enzymes participating in blood clotting. When cells die, lipid asymmetry is no longer maintained, and phosphatidylserine, normally enriched in the cytosolic leaflet, is increasingly found in the exoplasmic one. This increased exposure is detected experimentally by use of a labeled version of annexin V, a protein that specifically binds to phosphatidylserine, to measure the onset of programmed cell death (apoptosis). As discussed in Chapter 22, the increased exposure of phosphatidylserine on dying or dead cells is recognized by phagocytic cells, which initiate engulfment of such apoptotic bodies and thus ensure timely and safe disposal of cell remnants.

Cholesterol and Sphingolipids Cluster with Specific Proteins in Membrane Microdomains

Membrane lipids are not randomly distributed (evenly mixed) in each leaflet of a bilayer. One hint that lipids may be organized within the leaflets was the discovery that the lipids remaining after the extraction (solubilization) of plasma membranes with non-ionic detergents such as Triton X-100 predominantly contain two species: cholesterol and sphingomyelin. Because these two lipids are found in more ordered, less fluid bilayers, researchers hypothesized that they form microdomains, termed **lipid rafts**, surrounded by other, more fluid phospholipids that are more readily extracted by non-ionic detergents. (We discuss more fully the role of ionic and non-ionic detergents in extracting membrane proteins in Section 10.2.)

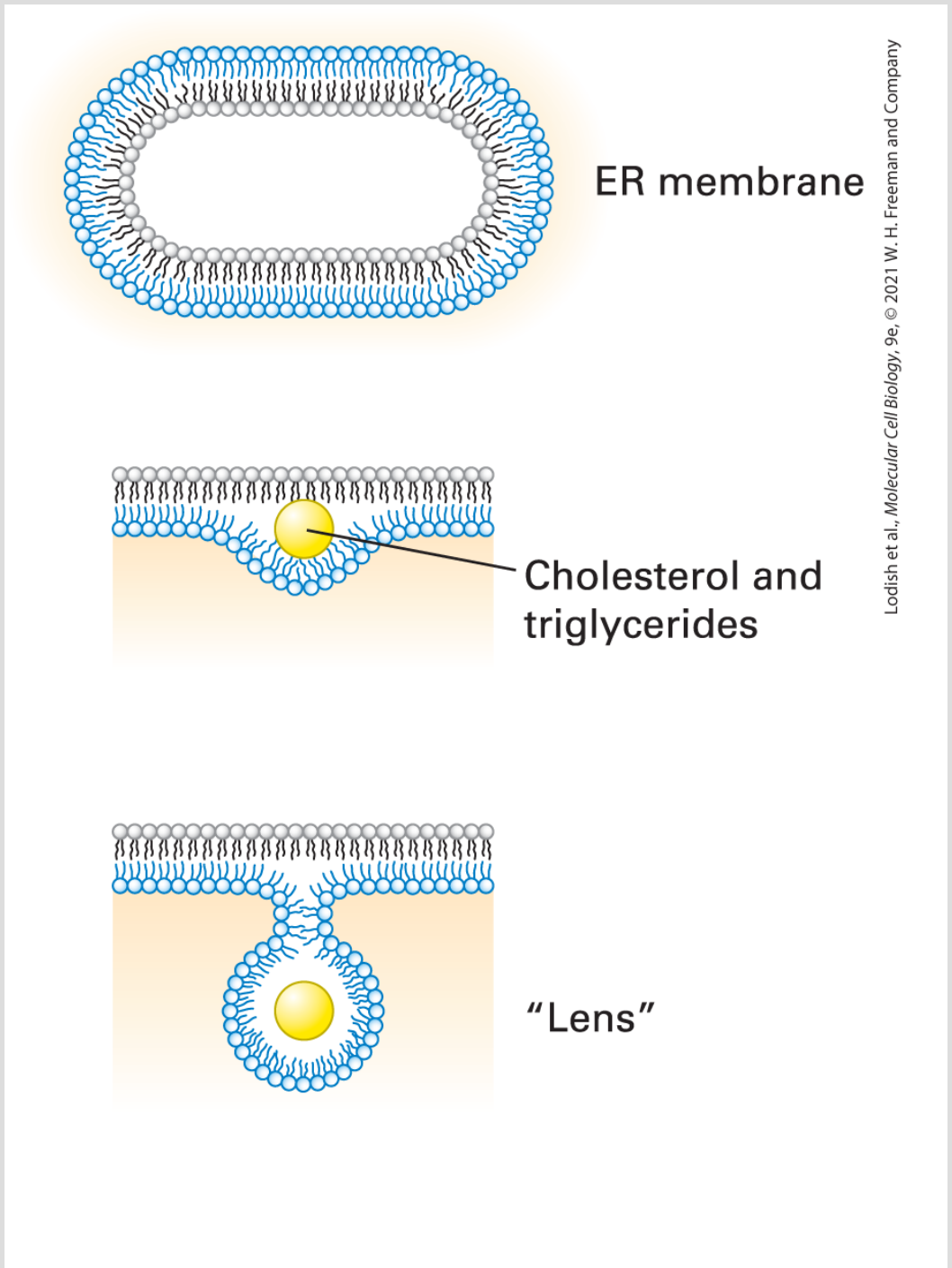
Some biochemical and microscopic evidence supports the existence of lipid rafts, which in natural membranes are typically 50 nm in diameter. Rafts can be disrupted by methyl-β-cyclodextrin, which specifically extracts cholesterol from membranes, or by antibiotics such as filipin that sequester cholesterol into aggregates within the membrane. Such findings indicate the importance of cholesterol in maintaining the integrity of lipid rafts. These raft fractions, defined by their insolubility in non-ionic detergents, contain a subset of plasma-membrane proteins, many of which are implicated in sensing extracellular signals and transmitting them into the cytosol. Because raft fractions are enriched in glycolipids, an important tool for microscopic visualization of raft-type structures in intact cells is the use of fluorescently labeled cholera toxin, a protein that specifically binds to certain gangliosides. By bringing many key proteins into close proximity and stabilizing their interactions, lipid

rafts may facilitate signaling by cell-surface receptors and the subsequent activation of cytosolic events. However, much remains to be learned about the structure and biological function of lipid rafts. The unique properties of some of the raft-associated lipids, such as glycolipids, may permit interactions of their tails across the hydrophobic core and help organize lipids of the cytosolic leaflet in the formation of signaling platforms.

Cells Store Excess Lipids in Lipid Droplets

Lipid droplets are vesicular structures, composed of triglycerides and cholesterol esters, that originate from the ER and serve a lipid-storage function. When a cell's supply of lipids exceeds the immediate need for membrane construction, excess lipids are relegated to these lipid droplets, readily visualized in live cells by staining with a lipophilic dye such as Congo red. Feeding cells with oleic acid, a type of fatty acid, enhances lipid droplet formation. Lipid droplets are not only storage compartments for triglycerides and cholesterol esters, but may also serve as platforms for storage of proteins targeted for degradation. The biogenesis of lipid droplets starts with delamination of the lipid bilayer of the ER through insertion of triglycerides and cholesterol esters that are more hydrophobic than cholesterol itself (Figure 10-13). The lipid "lens" continues to grow by insertion of more lipid, until finally a lipid

droplet is hatched by scission from the ER. The resulting cytoplasmic droplet is thereby surrounded by a phospholipid monolayer. The details of lipid droplet biogenesis, as well as its functions, remain to be defined more clearly.



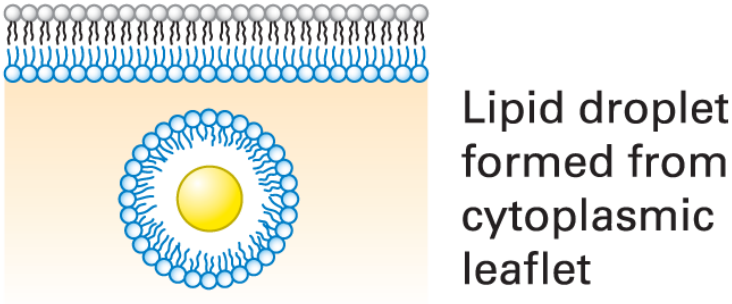


FIGURE 10-13 Lipid droplets form by budding and scission from the ER membrane. Lipid droplet formation begins with the accumulation of cholesterol esters and triglycerides within the hydrophobic core of the lipid bilayer. The resulting delamination of the two lipid monolayers causes a "lens" to form, the further growth of which creates a spherical droplet that is then released by scission at the neck. The newly formed droplet is surrounded by a lipid monolayer, derived from the cytosolic leaflet of the ER membrane.

KEY CONCEPTS OF SECTION 10.1

The Lipid Bilayer: Composition and Structural Organization

- Membranes are crucial to cell structure and function. The eukaryotic cell is demarcated from the external environment by the plasma membrane and organized into membrane-limited subcompartments (organelles and vesicles).
- The phospholipid bilayer, the basic structural unit of all biomembranes, is a lipid sheet two molecules thick, with hydrophilic faces and a hydrophobic core that is impermeable to water-soluble molecules and ions. Proteins embedded in the bilayer endow the membrane with specific functions (see Figure 10-1).
- Phospholipids spontaneously form bilayers and sealed compartments surrounding an aqueous space (see Figure 10-3).
- As bilayers, all biological membranes have an internal (cytosolic) face and an external (exoplasmic) face (see Figure 10-5). Some organelles, such as *E. coli*, are surrounded by two, rather than one, membrane bilayers.
- The primary lipid components of biomembranes are phosphoglycerides, sphingolipids, and sterols such as cholesterol. The term "phospholipid" applies to any amphipathic lipid molecule with a fatty acyl hydrocarbon tail and a phosphate-based polar head group (see Figure 10-8).

- Biomembranes can undergo phase transitions from fluidlike to gel-like states depending on the temperature and the composition of the membrane (see Figure 10-9).
- Most lipids and many proteins are laterally mobile in biomembranes (see Figure 10-10).
- Different cellular membranes vary in lipid composition (see Table 10-1). Phospholipids and sphingolipids are asymmetrically distributed in the two leaflets of the bilayer, whereas cholesterol is fairly evenly distributed in both leaflets.
- Natural biomembranes generally have a viscous consistency with fluidlike properties. In general, membrane fluidity is decreased by sphingolipids and cholesterol and increased by phosphoglycerides. The lipid composition of a membrane also influences its thickness and curvature (see Figure 10-11).
- Lipid rafts are microdomains containing cholesterol, sphingolipids, and certain membrane proteins that form in the plane of the bilayer. These lipid-protein aggregates might facilitate signaling by certain plasma-membrane receptors.
- Lipid droplets are storage vesicles for lipids, originating in the ER (see Figure 10-13).

10.2 Membrane Proteins: Structure and Basic Functions

Membrane proteins are defined by their location within or at the surface of a phospholipid bilayer. Although every biological membrane has the same basic bilayer structure, in which the proteins associated with a particular membrane are responsible for its distinctive activities. The kinds and amounts of proteins associated with biomembranes vary depending on cell type and subcellular location. For example, the inner mitochondrial membrane is 76 percent protein, but the myelin membrane that surrounds nerve axons is only 18 percent. The high phospholipid content of myelin allows it to electrically insulate the nerve from its environment, as we discuss in Chapter 23. The importance of membrane proteins is evident from the finding that approximately a third of all yeast genes encode a membrane protein. The relative abundance of genes for membrane proteins is even greater in multicellular organisms, in which they have additional functions in cell adhesion and in communication between different cell types (cell-cell interactions). Indeed, of the ~20,000 genes in the human genome, ~7000 encode membrane proteins.

The lipid bilayer presents a distinctive two-dimensional hydrophobic environment for membrane proteins. Some proteins contain segments that are embedded within the hydrophobic core of the phospholipid bilayer; other proteins are associated with the

exoplasmic or cytosolic leaflet of the bilayer. Protein domains on the extracellular surface of the plasma membrane generally bind to extracellular molecules, including external signaling proteins, ions, and small metabolites (e.g., glucose, fatty acids), as well as proteins on other cells or in the external environment. Segments of proteins within the plasma membrane perform multiple functions, such as forming the channels and pores through which molecules and ions move into and out of cells. These intramembrane segments also serve to organize multiple membrane proteins into larger assemblies within the plane of the membrane. Domains lying along the cytosolic face of the plasma membrane have a wide range of functions, from anchoring cytoskeletal proteins to the membrane to triggering intracellular signaling pathways or synthesis of ATP.

In many cases, the function of a membrane protein and the topology of its polypeptide chain in the membrane can be predicted on the basis of its similarity with other well-characterized proteins. In this section, we examine the characteristic structural features of membrane proteins and some of their basic functions. We describe the structures of several proteins to help you get a feel for how membrane proteins interact with membranes. More complete characterizations of the properties of various types of membrane proteins are presented in later chapters that focus on their structures and activities in the context of their cellular functions.

Proteins Interact with Membranes in Three Different Ways

Membrane proteins can be classified into three categories — integral, lipid-anchored, and peripheral — on the basis of their position with respect to the membrane (see Figure 10-1). **Integral membrane proteins**, also called *transmembrane proteins*, span a phospholipid bilayer and comprise three domains. The cytosolic and exoplasmic domains have hydrophilic exterior surfaces that interact with the aqueous environment on the cytosolic and exoplasmic faces of the membrane. These domains resemble segments of other water-soluble proteins in their amino acid composition and structure. In contrast, the membrane-spanning segments usually contain many hydrophobic amino acids whose side chains protrude outward and interact with the hydrophobic hydrocarbon core of the phospholipid bilayer. In all transmembrane proteins examined to date, the membrane-spanning domains consist of one or more α helices or of multiple β strands. We discuss the ribosomal synthesis and post-translational processing of soluble cytosolic proteins in Chapters 5 and 9; the process by which integral membrane proteins are inserted into membranes as part of their synthesis is discussed in Chapter 13.

Lipid-anchored membrane proteins are bound covalently to one or more lipid molecules. The hydrophobic tail of the attached lipid is embedded in one leaflet of the membrane and anchors the protein

to the membrane. The polypeptide chain itself does not enter the phospholipid bilayer.

Peripheral membrane proteins (see Figure 10-1) do not directly contact the hydrophobic core of the phospholipid bilayer. Instead, they are bound to the membrane either indirectly by interactions with integral or lipid-anchored membrane proteins or directly by interactions with lipid head groups. Peripheral proteins can be bound to either the cytosolic or the exoplasmic face of the plasma membrane. In addition to these proteins, which are closely associated with the bilayer, cytoskeletal filaments can be more loosely associated with the cytosolic face, usually through one or more peripheral adapter proteins. Such associations with the cytoskeleton provide support for various cellular membranes, helping to determine the cell's shape and mechanical properties, and play a role in the two-way communication between the cell interior and the exterior, as we explore in Chapter 17. Finally, peripheral proteins on the outer surface of the plasma membrane and the exoplasmic domains of integral membrane proteins are often attached to components of the extracellular matrix or to the cell walls surrounding bacterial and plant cells, providing a crucial interface between the cell and its environment.

Most Transmembrane Proteins Have Membrane-Spanning α Helices

Soluble proteins exhibit hundreds of distinct localized folded structures, or motifs (see Figure 3-7). In comparison, the repertoire of folded structures in the transmembrane domains of integral membrane proteins is quite limited, with the hydrophobic α helix predominating. Proteins containing membrane-spanning α-helical domains are stably embedded in membranes. This is due to energetically favorable hydrophobic and van der Waals interactions of the hydrophobic side chains in the domain with specific lipids and probably also because of ionic interactions with the polar head groups of the phospholipids.

A single α-helical domain is sufficient to incorporate an integral membrane protein into a membrane. However, many such proteins have more than one transmembrane α helix. Typically, a membrane-embedded α helix is composed of a continuous segment of 20–25 hydrophobic (uncharged) amino acids (see Figure 2-14). The predicted length of such an α helix (3.75 nm) is just sufficient to span the hydrocarbon core of a phospholipid bilayer. In many membrane proteins, these helices are perpendicular to the plane of the membrane, whereas in others, the helices traverse the membrane at an oblique angle. The hydrophobic side chains protrude outward from the helix and form van der Waals interactions with the fatty acyl chains in the bilayer. In contrast, the hydrophilic amide peptide bonds are in the interior of the α helix (see Figure 3-4); each carbonyl (C═O) group forms a hydrogen bond with the amide hydrogen atom of the amino acid four residues toward the C-

terminus of the helix. These polar groups are shielded from the hydrophobic interior of the membrane.

To help you get a better sense of the structures of proteins with α-helical domains, we will briefly discuss four different kinds of such proteins: glycophorin A, G protein–coupled receptors, aquaporins (water/glycerol channels), and the T-cell receptor for antigen.

Glycophorin A, the major protein in the erythrocyte plasma membrane, is a representative *single-pass* transmembrane protein, which contains only one membrane-spanning α helix (Figure 10-14a). The 23-residue membrane-spanning α helix is composed of amino acids with hydrophobic (uncharged) side chains, which interact with the fatty acyl chains in the surrounding bilayer. In cells, glycophorin A typically forms dimers: the transmembrane helix of one glycophorin A polypeptide associates with the corresponding transmembrane helix in a second glycophorin A to form a coiled-coil structure (Figure 10-14b). Such interactions of membrane-spanning α helices are a common mechanism for creating dimeric membrane proteins, and many membrane proteins form oligomers (two or more polypeptides bound together noncovalently) by interactions between their membrane-spanning helices.

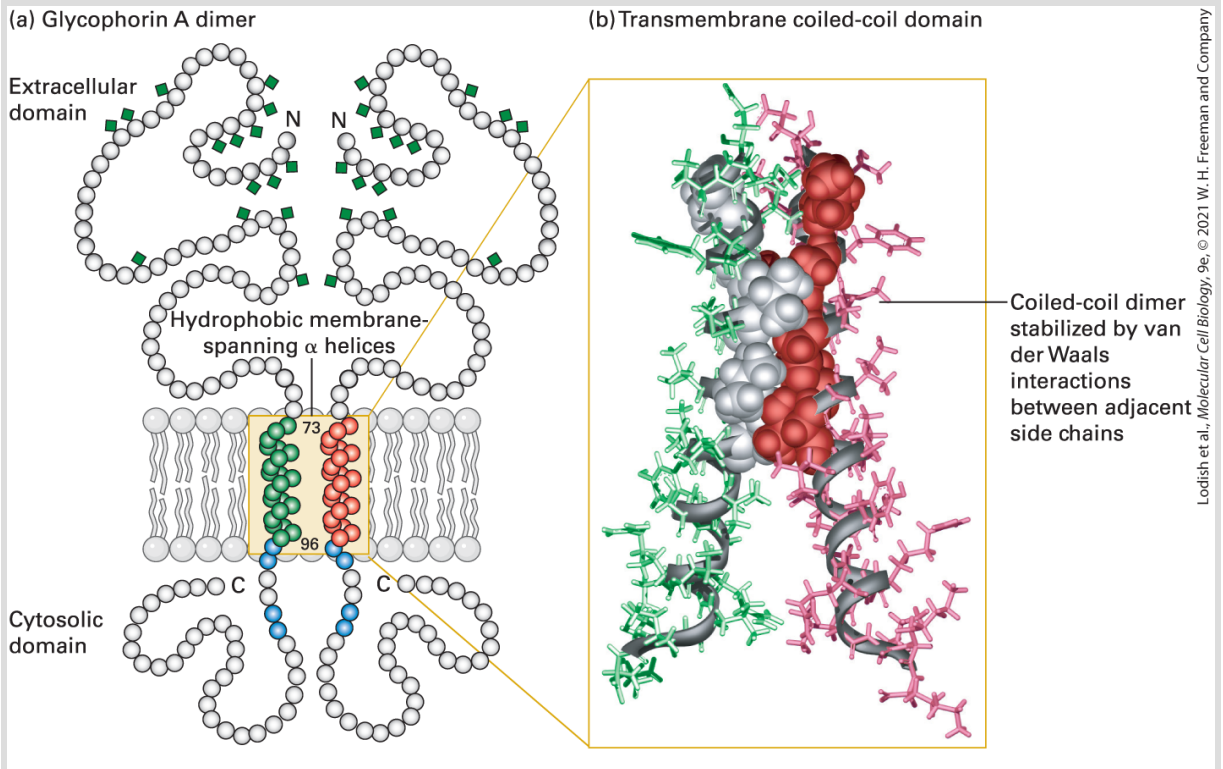


FIGURE 10-14 Structure of glycophorin A, a typical single-pass transmembrane protein. (a) Diagram of dimeric glycophorin, showing its major sequence features and its relation to the membrane. The single 23-residue membrane-spanning α helix in each monomer is composed of amino acids with hydrophobic (uncharged) side chains (red and green spheres). By binding negatively charged phospholipid head groups, the positively charged arginine and lysine residues (blue spheres) near the cytosolic side of the helix help anchor glycophorin in the membrane. Both the extracellular and the cytosolic domains are rich in charged residues and polar uncharged residues; the extracellular domain is heavily glycosylated, with carbohydrate chains (green diamonds) attached to specific serine, threonine, and asparagine residues. (b) Molecular model of the transmembrane domain of dimeric glycophorin A corresponding to residues 73–96. The hydrophobic side chains of the α helix in one monomer are shown in pink; those of the other monomer, in green. Residues depicted as space-filling structures participate in van der Waals interactions that stabilize the coiled-coil dimer. Note how the hydrophobic side chains project outward from the helix, toward what would be the surrounding fatty acyl chains.

[Part (b) Data from K. R. MacKenzie, J. H. Prestegard, and D. M. Engelman, 1997, *Science* **276**:131, PDB ID 1afo.]

A large and important group of integral membrane proteins is defined by the presence of seven membrane-spanning α helices. This group includes the large family of G protein–coupled cell-surface receptors discussed in Chapter 15, many of which have been crystallized. One such *multipass* transmembrane protein of known structure is bacteriorhodopsin, a protein found in the membranes of certain photosynthetic bacteria; it illustrates the general structure of all these proteins (Figure 10-15a). Absorption of light by the retinal group covalently attached to this protein causes a conformational change in the protein that results in the pumping of protons from the cytosol across the bacterial membrane to the extracellular space. The proton concentration gradient thus generated across the membrane is used to synthesize ATP during photosynthesis (see Chapter 12). In the high-resolution structure of bacteriorhodopsin, the positions of all the individual amino acids, the retinal group, and the surrounding lipids are clearly defined. As might be expected, virtually all the amino acids on the exterior of the membrane-spanning segments of bacteriorhodopsin are hydrophobic, permitting energetically favorable interactions with the hydrocarbon core of the surrounding lipid bilayer.

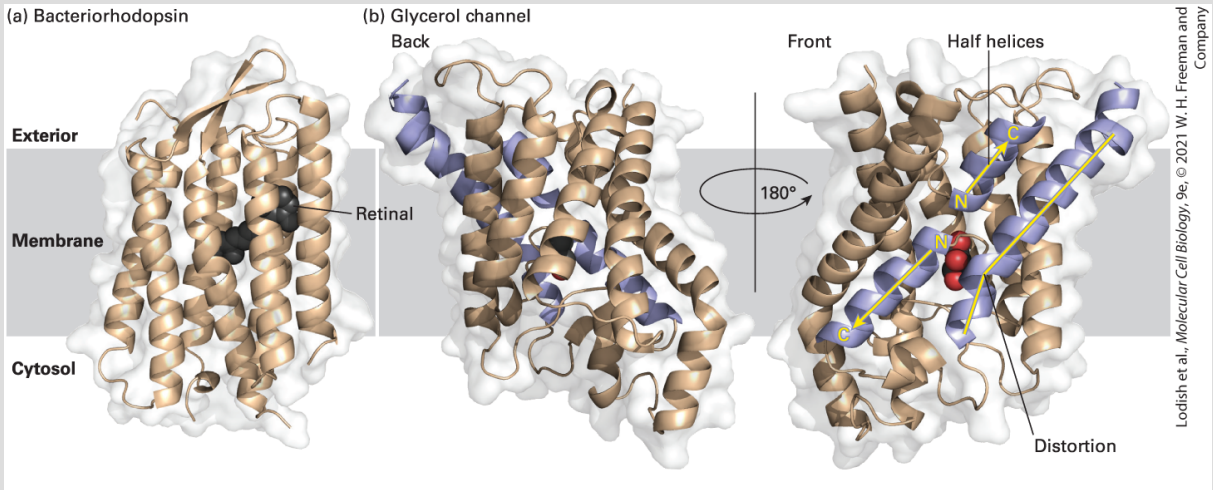


FIGURE 10-15 Structural models of two multipass transmembrane proteins. (a) The seven hydrophobic α helices in bacteriorhodopsin (a photoreceptor in certain bacteria) traverse the lipid bilayer roughly perpendicular to the plane of the membrane. A retinal molecule (black) covalently attached to one helix absorbs light. The large class of G protein–coupled receptors in eukaryotic cells also has seven membrane-spanning α helices; their three-dimensional structure is similar to that of bacteriorhodopsin. (b) Two views of the glycerol channel Glpf, rotated 180° with respect to each other along an axis perpendicular to the plane of the membrane. Note the several membrane-spanning α helices that are at oblique angles, the two helices that penetrate only halfway through the membrane (purple with yellow arrows), and the one long membrane-spanning helix with a "break" or distortion in the middle (purple with yellow line). The glycerol molecule in the hydrophilic "core" is colored red. The protein structure was approximately positioned in the hydrocarbon core of the membrane by finding the most hydrophobic 3-μm slab of the protein perpendicular to the membrane plane.

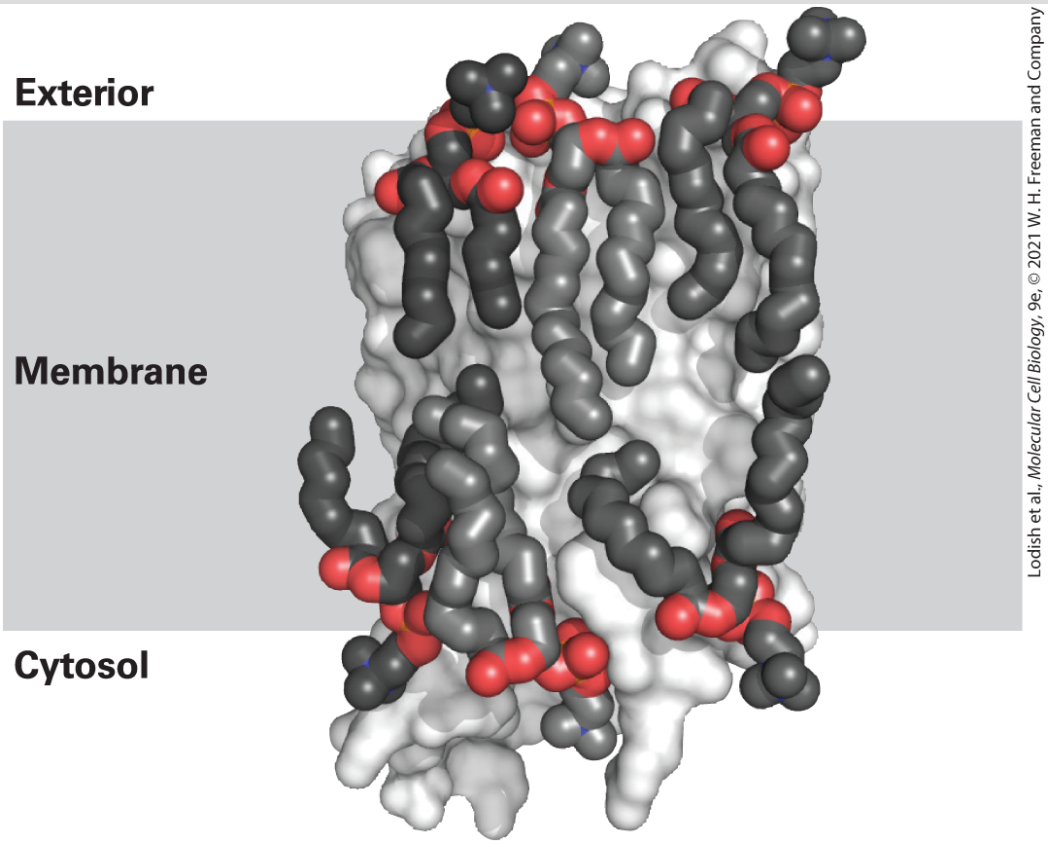
[Part (a) Data from H. Luecke et al., 1999, *J. Mol. Biol.* **291**:899. Part (b) Data from J. Bowie, 2005, *Nature* **438**:581–589, PDB ID 1c3w; and D. Fu et al., 2000, *Science* **290**:481–486, PDB ID 1fx8.]

The aquaporins are a large family of highly conserved proteins that transport water, glycerol, and other hydrophilic molecules across biomembranes. They illustrate several aspects of the structure of multipass transmembrane proteins. Aquaporins are tetramers of

four identical subunits. Each of the four subunits has six membrane-spanning α helices, some of which traverse the membrane at oblique angles rather than perpendicularly. Because all aquaporins have similar structures, we will focus on one, the glycerol channel Glpf, whose structure has been especially well defined by x-ray diffraction studies (Figure 10-15b). This aquaporin has one long transmembrane helix with a bend in the middle, and more strikingly, there are two α helices that penetrate only halfway through the membrane. The N-termini of these helices face each other (yellow N's in the figure), and together they span the membrane at an oblique angle. Thus some membrane-embedded helices — and other, nonhelical, structures we will encounter later — do not traverse the entire bilayer. As we will see in Chapter 11, these short helices in aquaporins form part of the glycerol/water-selective pore in the middle of each subunit (see Figure 11-8). This structure highlights the considerable diversity in the ways membrane-spanning α helices interact with the lipid bilayer and with other segments of the protein.

The specificity of phospholipid-protein interactions is evident from the structure of a different aquaporin, aquaporin 0 (Figure 10-16). Aquaporin 0 is the most abundant protein in the plasma membrane of the fiber cells that make up the bulk of the lens of the mammalian eye. Like other aquaporins, it is a tetramer of identical subunits. The protein's surface is not covered by a set of uniform binding sites for phospholipid molecules. Instead, fatty acyl chains pack tightly against the irregular hydrophobic outer surface of the protein. The

lipids involved in this interaction are referred to as *annular phospholipids* because they form a tight ring (annulus) of lipids around the protein that are not easily exchanged with bulk phospholipids in the bilayer. Some of the fatty acyl chains are straight, in the trans conformation (see Chapter 2), whereas others are kinked in order to interact with bulky hydrophilic side chains on the surface of the protein. Some of the lipid head groups are parallel to the surface of the membrane, as is the case in pure phospholipid bilayers. Others, however, are oriented almost at right angles to the plane of the membrane. Thus there can be specific interactions between phospholipids and membrane-spanning proteins, and the function of many membrane proteins can be affected by the specific types of phospholipids present in the bilayer.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 10-16 Annular phospholipids. Side view of the three- dimensional structure of one subunit of the lens-specific aquaporin 0 homotetramer, crystallized in the presence of the phospholipid dimyristoylphosphatidylcholine, a phospholipid with 14 carbon-saturated fatty acyl (myristic acid) chains. Note the lipid molecules forming a bilayer shell around the protein, which is shown as a surface plot (the lighter background molecule). The lipid molecules are shown in space-filling format; the polar lipid head groups (gray and red) and the lipid fatty acyl chains (the extended black and gray structures) form a bilayer with almost uniform thickness around the protein. Presumably, in the membrane, lipid fatty acyl chains cover the whole of the hydrophobic surface of the protein; only the most ordered of the lipid molecules would be resolved in the crystallographic structure.

[Data from A. Lee, 2005, *Nature* **438**:569–570; and T. Gonen et al., 2005, *Nature* **438**:633–688, PDB ID 2b6o.]

In addition to the predominantly hydrophobic (uncharged) residues that serve to embed integral membrane proteins in the bilayer, many α-helical transmembrane segments contain polar or charged residues. Their amino acid side chains can be used to guide the assembly and stabilization of multimeric membrane proteins. The T-cell receptor for antigen is a case in point: it is composed of four separate dimers, the interactions of which are driven by charge-charge interactions between α helices at the appropriate "depth" in the hydrocarbon core of the lipid bilayer (Figure 10-17). The electrostatic attraction of positive and negative charges on each dimer helps the dimers to find each other. Thus charged residues in otherwise hydrophobic transmembrane segments can help guide assembly of multimeric membrane proteins.

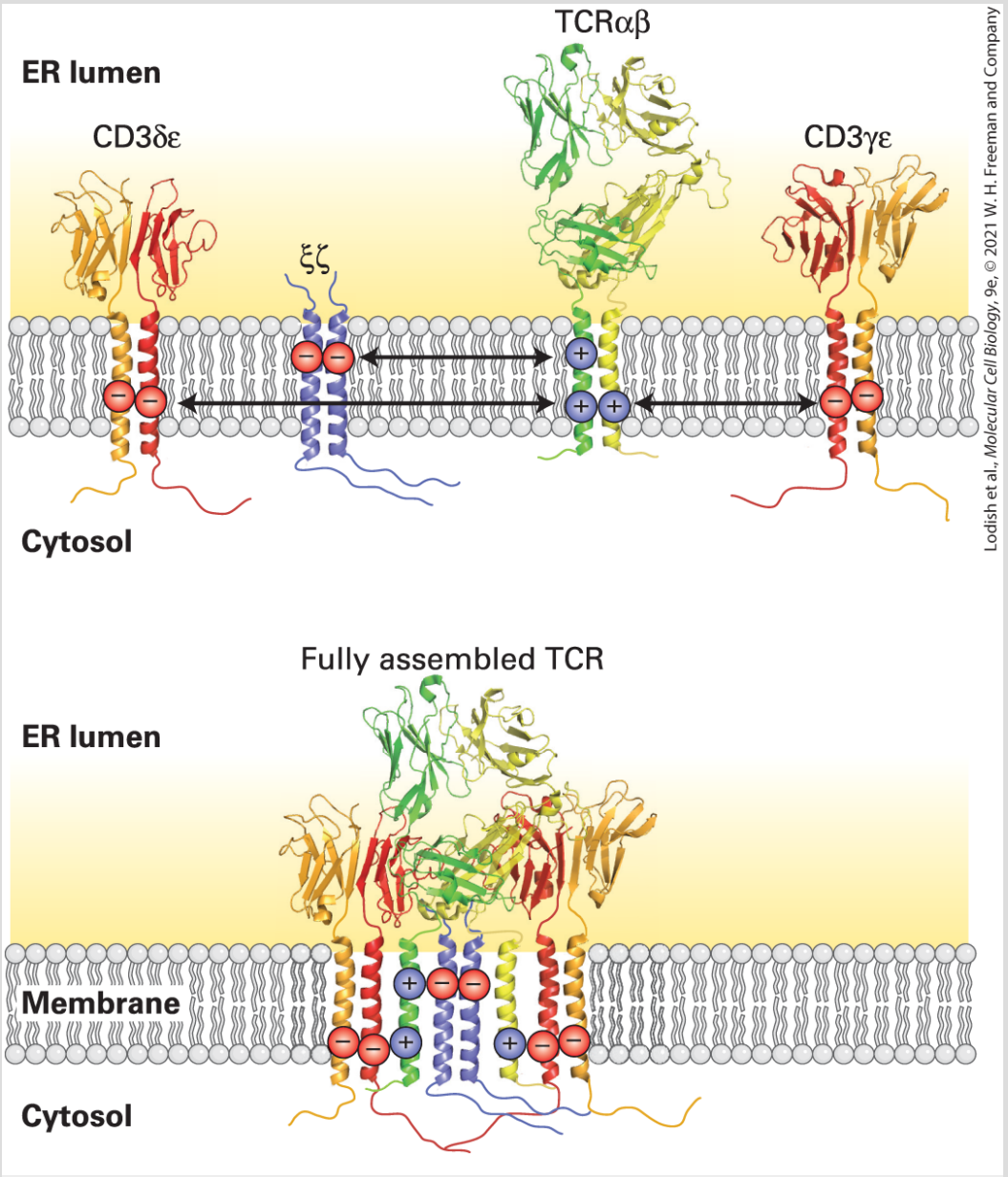


FIGURE 10-17 Charged residues can orchestrate the assembly of multimeric membrane proteins. The T-cell receptor (TCR) for antigen is composed of four separate dimers: an αβ pair directly responsible for antigen recognition and accessory subunits collectively referred to as the *CD3 complex*. These accessory subunits include the γ, δ, ε, and ζ subunits. The ζ subunits form a disulfide-linked homodimer. The γ and δ subunits occur in complex with an ε subunit to generate a γε and a δε pair. The transmembrane segments of the TCR α and β chains each contain positively charged residues (blue). These residues allow recruitment of corresponding δε and γε heterodimers, which carry negative

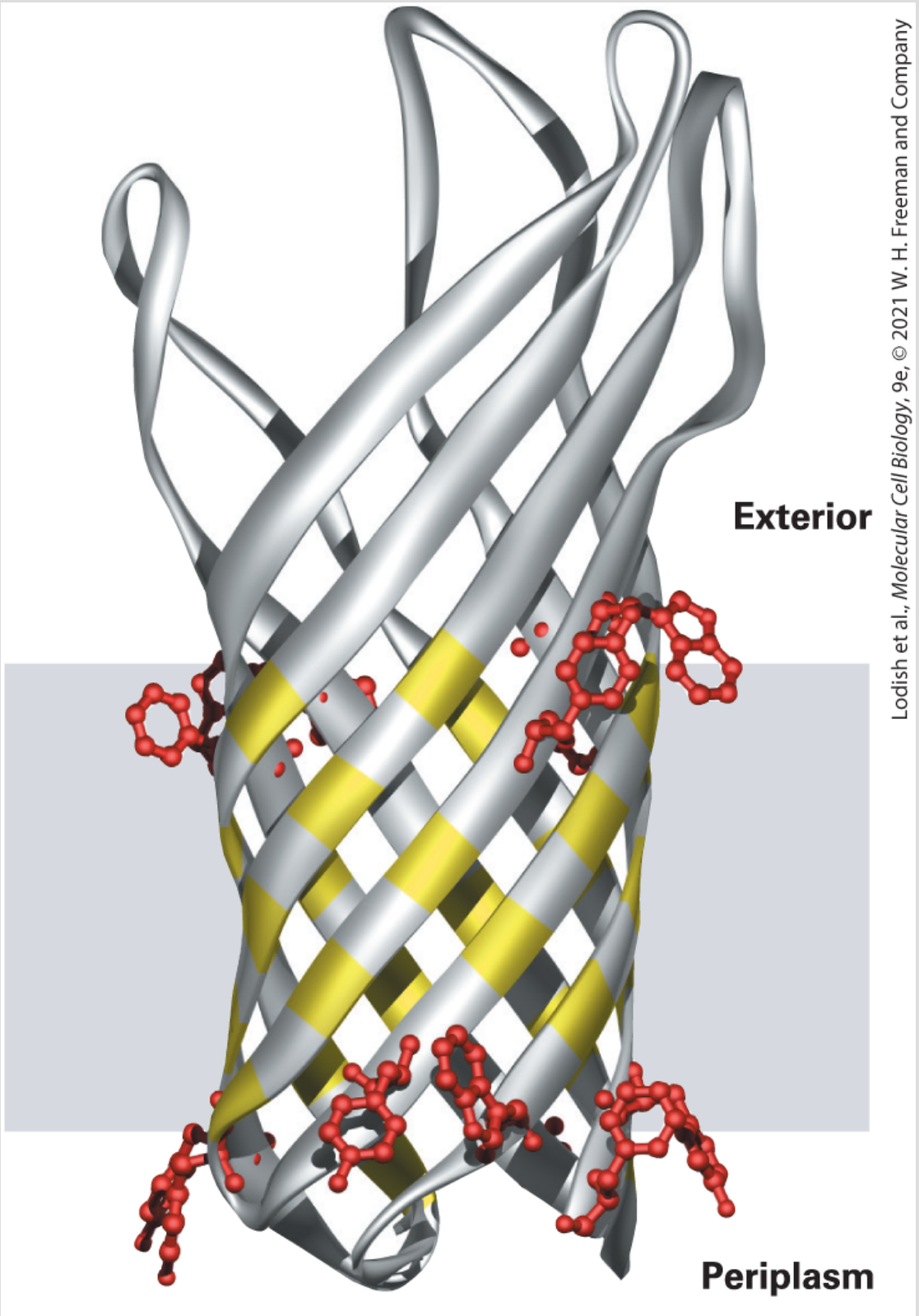
charges (red) at the appropriate depth in the hydrophobic core of the bilayer. The ζ homodimer docks onto the charges in the TCR α chain (dark green), while the γε and δε subunit pairs find their corresponding partners deeper down in the hydrophobic core on both the TCR α and TCR β chain (light green). Charged residues in otherwise nonpolar transmembrane segments can thus guide assembly of higher order structures.

[Data from K. W. Wucherpfennig et al., 2010, *Cold Spring Harb. Perspect. Biol.* **2**:a005140, PDB ID 1xmw; M. E. Call et al., 2006, *Cell* **127**:355, PDB ID 2hac; and L. Kjer-Nielsen et al., 2003, *Immunity* **18**:53, PDB ID 1mi5.]

Multiple β Strands in Porins Form Membrane-Spanning "Barrels"

The **porins** are a class of transmembrane proteins whose structure differs radically from that of other integral membrane proteins based on α-helical transmembrane domains. (Note that the aquaporins discussed above, despite their name, are not porins and contain multiple transmembrane α helices.) Several types of porins are found in the outer membranes of gram-negative bacteria such as *E. coli* (see Figure 1-12a) and in the outer membranes of mitochondria and chloroplasts. The outer membrane protects an intestinal bacterium from harmful agents (e.g., antibiotics, bile salts, and proteases) but permits the uptake and disposal of small hydrophilic molecules, including nutrients and waste products. Different types of porins in the outer membrane of an *E. coli* cell provide channels for the passage of specific types of disaccharides or other small molecules as well as for ions such as phosphate. The

amino acid sequences of porins contain none of the long, continuous hydrophobic segments typical of integral membrane proteins with α-helical membrane-spanning domains. Rather, the entire outer surface of the fully folded porin displays its hydrophobic character to the hydrocarbon core of the lipid bilayer. X-ray crystallography shows that porins are trimers of identical subunits. In each subunit, 16 β strands form a sheet that twists into a barrel-shaped structure with a pore in the center (Figure 10-18). Unlike a typical water-soluble globular protein, a porin has a hydrophilic interior and a hydrophobic exterior; in this sense, porins are inside out. In a porin monomer, the outward-facing side chains on each of the β strands are hydrophobic and form a nonpolar ribbonlike band that encircles the outside of the barrel. This hydrophobic band interacts with the fatty acyl groups of the membrane lipids or with other porin monomers. The side chains facing the inside of a porin monomer are predominantly hydrophilic; they line the pore through which small water-soluble molecules cross the membrane.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 10-18 Structural model of one subunit of OmpX, a porin found in the outer membrane of *E. coli*. All porins are trimeric transmembrane proteins. Each subunit

is barrel shaped, with β strands forming the wall and a transmembrane pore in the center. A band of aliphatic (hydrophobic and noncyclic) side chains (yellow) and a border of aromatic (ring-containing) side chains (red) position the protein in the bilayer.

[Data from J. Vogt and G. E. Schulz, 1999, *Structure* **7**:1301, PDB ID 1qj8.]

Recent advances in structural biology, notably the application of cryo-electron microscopy (see Chapter 3) have made it possible to study the structure of membrane proteins, including their transmembrane segments, at near-atomic resolution. Examples include members of the ATP-binding cassette family of transporters (ABC transporters).

Covalently Attached Lipids Anchor Some Proteins to Membranes

In eukaryotic cells, covalently attached lipids anchor some otherwise typically water-soluble proteins to one or the other leaflet of the membrane. In such lipid-anchored proteins, the hydrocarbon chains of the lipid anchor are embedded in the bilayer, but the protein itself does not enter the bilayer. The lipid anchors used to anchor proteins to the cytosolic face are not used for the exoplasmic face, and vice versa.

One group of cytosolic proteins is anchored to the cytosolic face of a membrane by a fatty acyl group (e.g., myristate or palmitate)

covalently attached to an N-terminal glycine residue, a modification called *acylation* (Figure 10-19a). Retention of such proteins at the membrane by the N-terminal acyl anchor may play an important role in a membrane-associated function. For example, v-Src, a mutant form of a cellular tyrosine kinase, induces abnormal cellular growth that can lead to cancer, but does so only when it has a myristylated N-terminus (see Chapter 25).

A second group of cytosolic proteins are anchored to membranes by a hydrocarbon chain attached to a cysteine residue at or near the C-terminus, a modification called *prenylation* (Figure 10-19b). Prenyl anchors are built from 5-carbon isoprene units, which, as detailed in Section 10.3, are also used in the synthesis of cholesterol. In prenylation, a 15-carbon farnesyl or 20-carbon geranylgeranyl group is bound by a thioether bond to the —SH group of a C-terminal cysteine residue of the protein, usually part of a C-terminal Cys-Ala-Ala-X (X = any of a number of amino acids) or CAAX box. Once prenylation has occurred, the C-terminal Ala-Ala-X motif may be removed by proteolysis. In some cases, a second geranylgeranyl group or a fatty acyl palmitate group is linked to a nearby cysteine residue. The additional hydrocarbon anchor is thought to reinforce the attachment of the protein to the membrane. For example, Ras, a GTPase superfamily protein that functions in intracellular signaling (see Chapter 16), is recruited to the cytosolic face of the plasma membrane by such a double anchor. Rab proteins, which also belong to the GTPase superfamily, are similarly bound to the

cytosolic surface of intracellular vesicles by prenyl anchors; these proteins are required for the fusion of vesicles with their target membranes (see Chapter 14). Yet other proteins are palmitoylated on membrane-proximal cysteine residues in the absence of other lipid modifications.

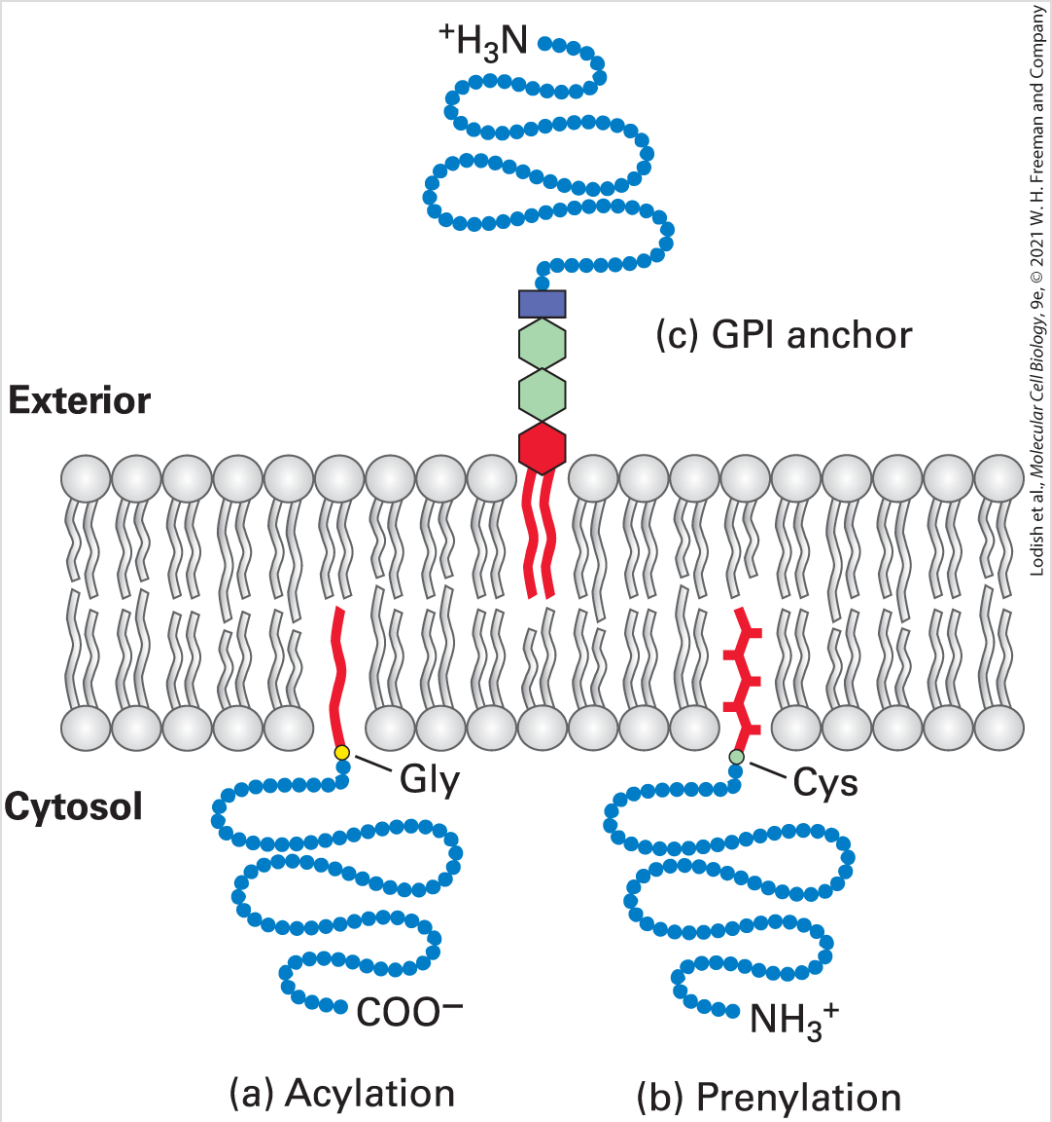


FIGURE 10-19 Anchoring of plasma-membrane proteins to the phospholipid bilayer by covalently linked hydrocarbon groups. (a) Cytosolic proteins such as v-Src are associated with the plasma membrane through a single fatty acyl

chain attached to the N-terminal glycine (Gly) residue of the polypeptide. Myristate (C14) and palmitate (C16) are common acyl anchors. (b) Other cytosolic proteins (e.g., Ras and Rab proteins) are anchored to the membrane by prenylation of one or two cysteine (Cys) residues at or near the C-terminus. The anchors are farnesyl (C15) and geranylgeranyl (C20) groups, both of which are unsaturated. (c) The lipid anchor on the exoplasmic surface of the plasma membrane is glycosylphosphatidylinositol (GPI). The phosphatidylinositol part (red) of this anchor contains two fatty acyl chains that extend into the bilayer. The phosphoethanolamine unit (purple) in the anchor links it to the protein. The two green hexagons represent sugar units, which vary in number, nature, and arrangement in different GPI anchors. The complete structure of a yeast GPI anchor is shown in Figure 13-15.

[Information from H. Sprong et al., 2001, *Nature Rev. Mol. Cell Biol.* **2**:504–513.]

Some cell-surface proteins, and some specialized proteins with distinctive covalently attached polysaccharides called *proteoglycans* (see Chapter 20), are bound to the exoplasmic face of the plasma membrane by a third anchor group, glycosylphosphatidylinositol (GPI) anchors. The exact structures of GPI anchors vary greatly among different cell types, but they always contain phosphatidylinositol, whose two fatty acyl chains extend into the lipid bilayer just like those of typical membrane phospholipids; phosphoethanolamine, which covalently links the anchor to the C-terminus of a protein; and several sugar residues (Figure 10-19c). Therefore, GPI anchors are glycolipids. The GPI anchor is both necessary and sufficient for binding proteins to the membrane. For instance, treatment of cells with phospholipase C, which cleaves the phosphate-glycerol bond in phospholipids and in GPI anchors (see

Figure 10-12), releases GPI-anchored proteins such as Thy-1 and placental alkaline phosphatase (PLAP) from the cell surface.

All Transmembrane Proteins and Glycolipids Are Asymmetrically Oriented in the Bilayer

Every type of transmembrane protein has a specific orientation, known as its *topology*, with respect to the membrane faces. Its cytosolic segments always face the cytosol, and its exoplasmic segments always face the opposite side of the membrane. This asymmetry in protein orientation gives the two membrane faces their different properties. The orientations of different types of transmembrane proteins are established during their synthesis, as we will see in Chapter 13. Membrane proteins have never been observed to flip-flop across a membrane; this would require a transient movement of hydrophilic amino acid residues through the hydrophobic interior of the membrane, which would be energetically unfavorable. Accordingly, the asymmetric topology of a transmembrane protein is maintained throughout the protein's lifetime. As Figure 10-6 shows, membrane proteins retain their asymmetric orientation during membrane budding and fusion events; the same segment always faces the cytosol and the same segment is always exposed to the exoplasmic face. In multipass membrane proteins, the orientation of individual transmembrane

segments can be affected by changes in the membrane's phospholipid composition.

Many transmembrane proteins contain carbohydrate chains covalently linked to serine, threonine, or asparagine side chains of the polypeptide. Such transmembrane **glycoproteins** are always oriented so that all the carbohydrate chains are in the exoplasmic domain (see Figure 10-14 for the example of glycophorin A). Likewise, glycolipids, in which a carbohydrate chain is attached to the glycerol or sphingosine backbone of a membrane lipid, are always located in the exoplasmic leaflet, with the carbohydrate chain protruding from the membrane surface. The biosynthetic basis for the asymmetric glycosylation of proteins is described in Chapters 13 and 14. Both glycoproteins and glycolipids are especially abundant in the plasma membranes of eukaryotic cells and in the membranes of the intracellular compartments that establish the secretory and endocytic pathways; they are absent from the inner mitochondrial membrane, chloroplast lamellae, and several other intracellular membranes. Because the carbohydrate chains of glycoproteins and glycolipids in the plasma membrane extend into the extracellular space, they are available to interact with components of the extracellular matrix as well as with **lectins** (proteins that bind specific sugars), growth factors, and antibodies.

One important consequence of interactions involving membrane glycoproteins and glycolipids is illustrated by the ABO

blood-group antigens. These three structurally related oligosaccharide components of certain glycoproteins and glycolipids are expressed on the surfaces of human red blood cells and many other cell types (Figure 10-20). All humans have the enzymes for synthesizing O antigen. Persons with type A blood also have a glycosyltransferase enzyme that adds an extra modified monosaccharide called N-*acetylgalactosamine* to O antigen to form A antigen. Those with type B blood have a different transferase that adds an extra galactose to O antigen to form B antigen. People with both transferases produce both A and B antigen (AB blood type); those who lack these transferases produce only O antigen (O blood type).

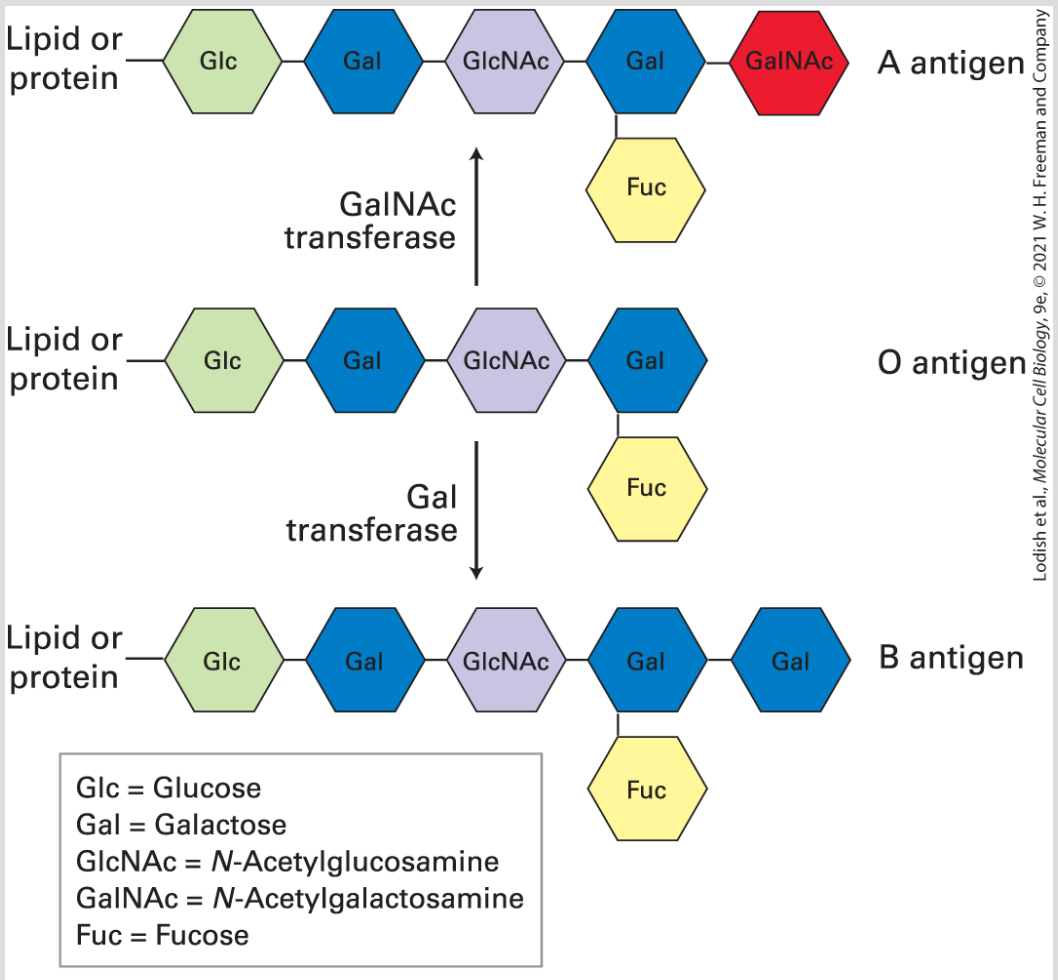


FIGURE 10-20 Human ABO blood-group antigens. These antigens are oligosaccharide chains covalently attached to glycolipids or glycoproteins in the plasma membrane. The terminal oligosaccharide sugars distinguish the three antigens. The presence or absence of the glycosyltransferases that add galactose (Gal) or *N*-acetylgalactosamine (GalNAc) to O antigen determine a person's blood type.

People whose erythrocytes lack the A antigen, the B antigen, or both on their surface normally have antibodies against the missing antigen(s) in their serum. Thus if a type A or O person receives a transfusion of type B blood, antibodies against the B antigen will bind to the introduced red cells and trigger their destruction. To

prevent such harmful reactions, blood-group typing and appropriate matching of blood donors and recipients are required in all transfusions (Table 10-2). ■

TABLE 10-2 • ABO Blood Groups

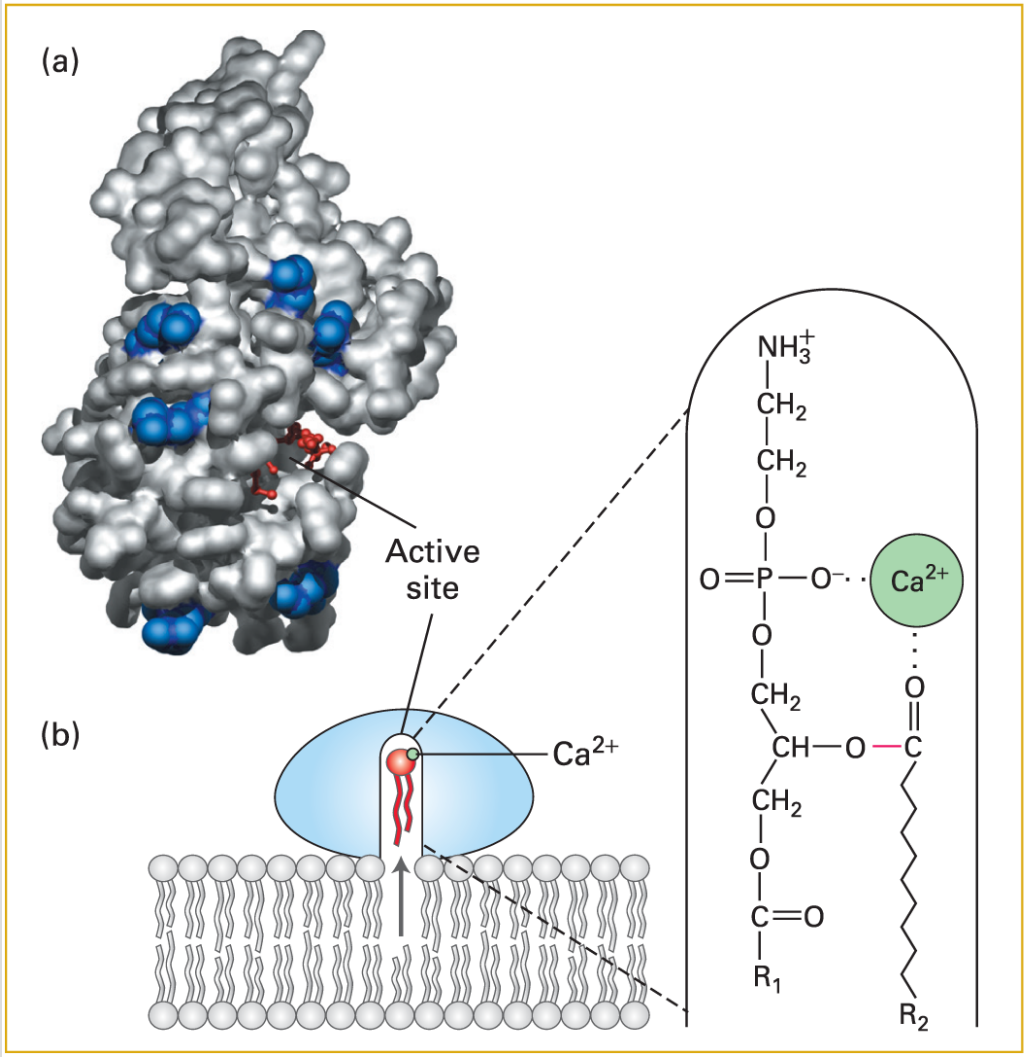
Blood Group	Antigens on RBCs[i]	Serum Antibodies	Can Receive Blood Types
A	A	Anti-B	A and O
B	B	Anti-A	B and O
AB	A and B	None	All
O	O	Anti-A and anti-B	O

[i]See Figure 10-20 for antigen structures.

Lipid-Binding Motifs Help Target Peripheral Proteins to the Membrane

Many water-soluble enzymes whose substrates are phospholipids must bind to membrane surfaces. Phospholipases, for example, hydrolyze various bonds in the head groups of phospholipids (see Figure 10-12) and thereby play a variety of roles in cells — helping to degrade damaged or aged cellular membranes, generating precursors to signaling molecules, and even serving as the active

components in many snake venoms. Many peripheral proteins, including phospholipases, initially bind to the polar head groups of membrane phospholipids to carry out their catalytic functions. The mechanism of action of phospholipase A_2 illustrates how such enzymes can reversibly interact with membranes and catalyze reactions at the interface of an aqueous solution and a lipid surface. When this enzyme is in aqueous solution, its Ca^{2+}-containing active site is buried in a channel lined with hydrophobic amino acids. The enzyme binds with greatest affinity to bilayers composed of negatively charged phospholipids (e.g., phosphatidylserine). This observation suggests that the rim of positively charged lysine and arginine residues around the entrance to the catalytic channel is particularly important in binding (Figure 10-21a) and constitutes a lipid binding motif. Binding induces a conformational change in phospholipase A_2 that strengthens its binding to the phospholipid heads and opens the hydrophobic channel. As a phospholipid molecule moves from the bilayer into the channel, the enzyme-bound Ca^{2+} binds to the phosphate in the head group, thereby positioning the ester bond to be cleaved in the catalytic site (Figure 10-21b), releasing the acyl chain.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 10-21 Lipid-binding surface and mechanism of action of phospholipase A_2. (a) A structural model of the enzyme, showing the surface that interacts with a membrane. This lipid-binding surface contains a rim of positively charged arginine and lysine residues (blue), surrounding the cavity of the catalytic active site, in which a substrate lipid (red ball-and-stick structure) is bound. (b) Diagram of catalysis by phospholipase A_2. When docked on a lipid membrane, positively charged residues of the binding site bind to negatively charged polar groups at the membrane surface. This binding triggers a small conformational change, opening a channel lined with hydrophobic amino acids that leads from the bilayer to the catalytic site. As a phospholipid moves into the channel, an enzyme-bound Ca^{2+} ion (green) binds to the head group, positioning the ester bond to be cleaved (red) next to the catalytic site. See D.

Blow, 1991, *Nature* **351**:444; and M. H. Gelb, W. Cho, and D. C. Wilton, 1999, *Curr. Opin. Struct. Biol.* **9**:428.

[Part (a) Data from D. L. Scott et al., 1990, *Science* **250**:1563, PDB ID 1poc.]

Proteins Can Be Removed from Membranes by Detergents or High-Salt Solutions

Membrane proteins are often difficult to purify and study, mostly because of their tight association with membrane lipids and other membrane proteins. *Detergents,* which are amphipathic molecules that disrupt membranes by intercalating into phospholipid bilayers, can be used to solubilize lipids and many membrane proteins. The hydrophobic part of a detergent molecule is attracted to the phospholipid hydrocarbons and mingles with them readily; the hydrophilic part is strongly attracted to water. Some detergents, such as bile salts, are natural products, but most are synthetic molecules developed for cleaning and for dispersing mixtures of oil and water in the food industry (e.g., to keep salad dressings from separating). Ionic detergents, such as sodium deoxycholate (a bile salt) and sodium dodecylsulfate (SDS), contain a charged group; non-ionic detergents, such as Triton X-100 and octylglucoside, lack a charged group (Figure 10-22). At very low concentrations, detergents dissolve in pure water as isolated molecules. As the concentration increases, the molecules begin to form micelles — small spherical

aggregates in which the hydrophilic parts of the molecules face outward and the hydrophobic parts cluster in the center (see Figure 10-3c). The *critical micelle concentration (CMC)* at which micelles form is characteristic of each detergent and is a function of the structures of its hydrophobic and hydrophilic parts.

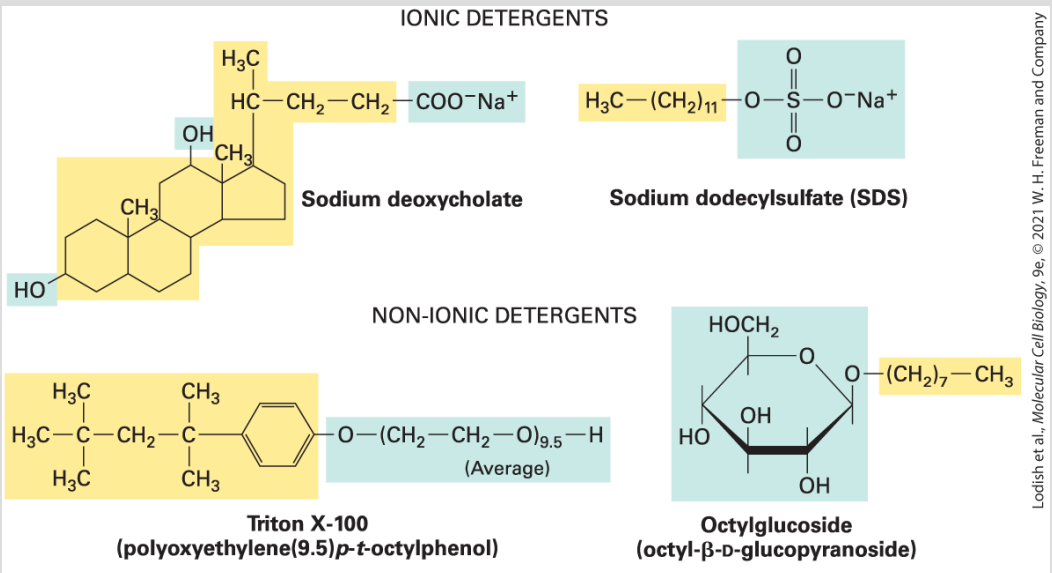


FIGURE 10-22 Structures of four common detergents. The hydrophobic part of each molecule is shown in yellow; the hydrophilic part, in blue. The bile salt sodium deoxycholate is a natural product; the others are synthetic. Although ionic detergents commonly cause denaturation of proteins, non-ionic detergents do not and are thus useful in solubilizing integral membrane proteins.

Ionic and non-ionic detergents interact differently with proteins and have different uses in the lab. Ionic detergents bind to the exposed hydrophobic regions of membrane proteins as well as to the hydrophobic cores of water-soluble proteins. Because of their charge, these detergents can disrupt ionic and hydrogen bonds. At

high concentrations, sodium dodecylsulfate, for example, completely denatures proteins by binding to every side chain, a property that is exploited in SDS gel electrophoresis (see Figure 3-41). Non-ionic detergents generally do not denature proteins and are thus useful in extracting proteins in their folded and active form from membranes before the proteins are purified. Protein-protein interactions, especially the weaker ones, can be sensitive to both ionic and non-ionic detergents.

At high concentrations (above the CMC), non-ionic detergents solubilize biological membranes by forming mixed micelles of detergent, phospholipids, and integral membrane proteins — bulky hydrophobic structures that do not dissolve in water (Figure 10-23, *top*). At low concentrations (below the CMC), these detergents bind to the hydrophobic regions of most integral membrane proteins, but without forming micelles, allowing the proteins to remain soluble (Figure 10-23, *bottom*). Creating such an aqueous solution of integral membrane proteins is a necessary first step in protein purification.

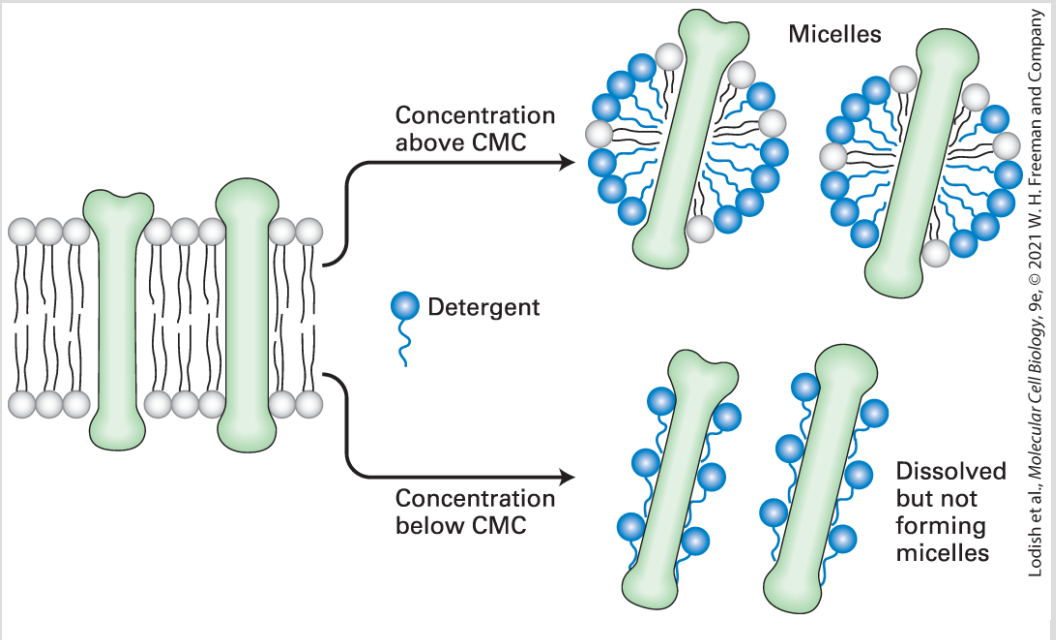


FIGURE 10-23 Solubilization of integral membrane proteins by non-ionic detergents. At a concentration higher than its critical micelle concentration (CMC), a detergent solubilizes lipids and integral membrane proteins, forming mixed micelles containing detergent, protein, and lipid molecules. At concentrations below the CMC, non-ionic detergents (e.g., octylglucoside, Triton X-100) can dissolve membrane proteins without forming micelles by coating the membrane-spanning regions.

Treatment of cultured cells with a buffered salt solution containing a non-ionic detergent such as Triton X-100 extracts water-soluble proteins as well as integral membrane proteins from cellular membranes. As noted earlier, the exoplasmic and cytosolic domains of integral membrane proteins are generally hydrophilic and soluble in water. The membrane-spanning domains, however, are rich in hydrophobic and uncharged residues (see Figure 10-14). When separated from membranes, these exposed hydrophobic segments tend to interact with one another, causing the protein molecules to

aggregate and precipitate from aqueous solutions. The hydrophobic parts of non-ionic detergent molecules preferentially bind to the hydrophobic segments of transmembrane proteins, preventing protein aggregation and allowing the proteins to remain in aqueous solution. Detergent-solubilized transmembrane proteins can then be purified by affinity chromatography and other techniques used in purifying water-soluble proteins (see Chapter 3).

As discussed previously, most peripheral membrane proteins are bound to specific transmembrane proteins or membrane phospholipids by ionic or other weak noncovalent interactions. Generally peripheral proteins can be removed from the membrane by solutions of high ionic strength (high salt concentrations), which disrupt ionic bonds, or by chemicals that bind divalent cations such as Mg^{2+}. Unlike integral membrane proteins, most peripheral proteins are soluble in water and do not need to be solubilized by non-ionic detergents.

The localization of membrane proteins to the phospholipid bilayer is not necessarily permanent. Phosphatidylinositide-specific phospholipase C hydrolyzes the membrane anchor of proteins attached to the membrane via this hydrophobic tail. Similarly, proteases can “shave off” the extracellular portions of single-pass membrane proteins. This can be done experimentally, for example, by exposing cells or membranes derived from them to a protease such as papain, and often allows recovery of the released extracellular portion for structural and functional studies. These

and similar methods are commonly used to detach cultured mammalian cells from the solid supports on which they are often grown or to detach cells from tissues to generate single cell suspensions. However, many single-pass membrane proteins are natural substrates for *sheddases*: members of the metalloprotease family of proteases that are discussed in Chapter 16. The extracellular domains released in this fashion may be found in the circulation and exert biological functions distinct from their membrane-anchored equivalents.

KEY CONCEPTS OF SECTION 10.2

Membrane Proteins: Structure and Basic Functions

- Biological membranes usually contain integral (transmembrane) proteins as well as lipid-anchored proteins and peripheral membrane proteins, which do not enter the hydrophobic core of the bilayer (see Figure 10-1).
- Most integral membrane proteins contain one or more membrane-spanning hydrophobic α helices, which are bracketed by hydrophilic domains that extend into the aqueous environment surrounding the cytosolic and exoplasmic faces of the membrane (see Figures 10-14, 10-15, and 10-17).
- Fatty acyl side chains as well as the polar head groups of membrane lipids pack tightly and irregularly around the hydrophobic segments of integral membrane proteins (see Figure 10-16).
- The porins, unlike other integral membrane proteins, contain membrane-spanning β sheets that form a barrel-like channel through the bilayer (see Figure 10-18).
- Lipids attached to certain amino acids anchor some proteins to one or the other membrane leaflet (see Figure 10-19).
- All transmembrane proteins and glycolipids are asymmetrically oriented in the bilayer. Invariably, carbohydrate chains are present only on the exoplasmic surface of a glycoprotein or glycolipid.
- Many water-soluble enzymes (e.g., phospholipases) whose substrates are phospholipids must bind to the membrane surface to carry out their function. Such

binding is often due to the attraction between positive charges on lysine and arginine residues in the protein and negative charges on phospholipid head groups in the bilayer.

- Transmembrane proteins can be selectively extracted from membranes with the use of non-ionic detergents.

10.3 Phospholipids, Sphingolipids, and Cholesterol: Synthesis and Intracellular Movement

In this section, we consider some of the special challenges that a cell faces in synthesizing and transporting lipids, which are poorly soluble in the aqueous interior of a cell. The focus will be the biosynthesis and movement of the major lipids found in cellular membranes — phospholipids, sphingolipids, and cholesterol — and their precursors. In lipid biosynthesis, water-soluble precursors are assembled into membrane-associated intermediates that are then converted into membrane lipid products. The movement of these lipids, especially membrane components, between different organelles is critical for maintaining the proper composition and properties of membranes and overall cell structure.

A fundamental principle of membrane biosynthesis is that cells synthesize new membranes only by the expansion of existing membranes. [The one exception may be autophagy, in which new membrane is formed first through the formation of an autophagic crescent, the construction of which involves modification of phosphatidylethanolamine with the ubiquitin-like modifier Atg8 (see Figure 14-34).] Although some early steps in the synthesis of membrane lipids take place in the cytosol, the final steps are catalyzed by enzymes bound to preexisting cellular membranes, and

the products are incorporated into the membranes as they are generated. Evidence for this process is seen when cells are briefly exposed to radioactively labeled precursors (e.g., phosphates or fatty acids): all the phospholipids and sphingolipids incorporating these labeled precursors are associated with intracellular membranes; as expected from the hydrophobicity of the fatty acyl chains, none are found free in the cytosol.

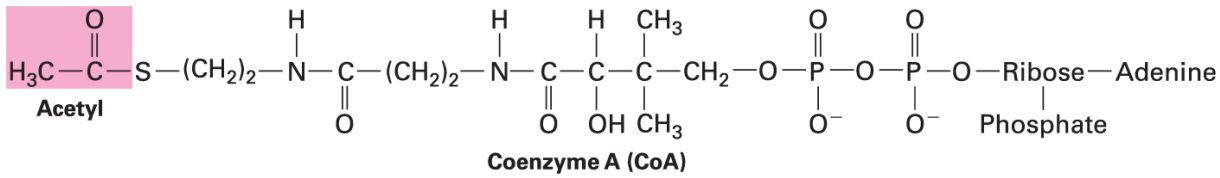
After they are formed, membrane lipids must be distributed appropriately both between the leaflets of a given membrane and among the independent membranes of different organelles in eukaryotic cells, as well as the plasma membrane. Here we consider how this precise lipid distribution is accomplished; in Chapters 13 and 14, we discuss how membrane proteins are inserted into cellular membranes and trafficked to their appropriate location within the cell.

Fatty Acids Are Assembled from Two-Carbon Building Blocks by Several Important Enzymes

Fatty acids (see Chapter 2) play a number of important roles in cells. In addition to being a cellular fuel source (see the discussion of aerobic oxidation in Chapter 12), fatty acids are key components of both the phospholipids and the sphingolipids that make up cellular

membranes; they also anchor some proteins to cellular membranes (see Figure 10-19). Thus the regulation of fatty acid synthesis plays a key role in the regulation of membrane synthesis as a whole. The major fatty acids in phospholipids contain 14, 16, 18, or 20 carbon atoms and include both saturated and unsaturated chains. The fatty acyl chains found on sphingolipids can be longer than those in the phosphoglycerides, containing up to 26 carbon atoms, and may bear other chemical modifications (e.g., hydroxylation) as well.

Fatty acids are synthesized from the two-carbon building block acetate (CH_3COO^-). In cells, both acetate and the intermediates in fatty acid biosynthesis are esterified to a large water-soluble molecule called *coenzyme A (CoA)*, as exemplified by the structure of **acetyl CoA** shown here.

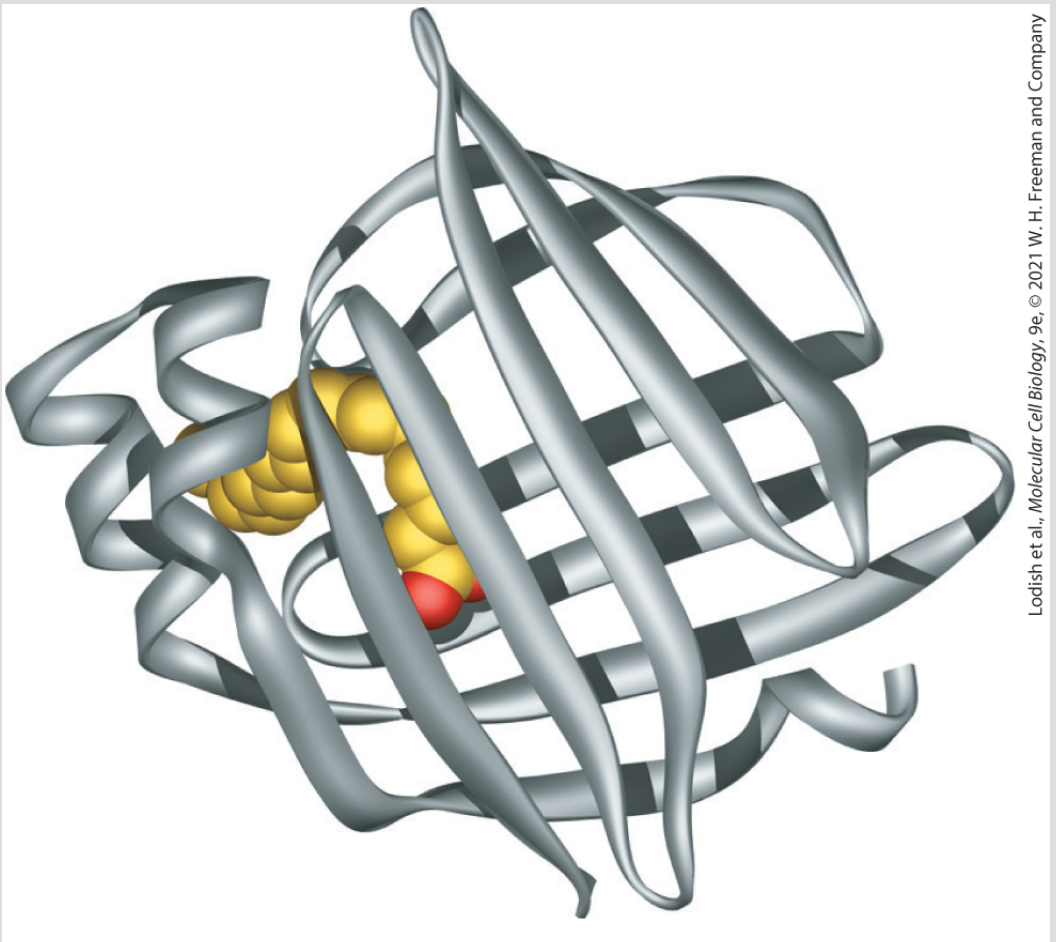


Acetyl CoA is an important intermediate in the metabolism of glucose, fatty acids, and many amino acids, as detailed in Chapter 12. It also contributes acetyl groups in many biosynthetic pathways. **Saturated** fatty acids (with no carbon-carbon double bonds) containing 14 or 16 carbon atoms are made from acetyl CoA by two enzymes, acetyl-CoA carboxylase *and* fatty acid synthase. In animal cells, these enzymes are found in the cytosol; in plants, they are

found in chloroplasts. Palmitoyl CoA (a 16-carbon fatty acyl group linked to CoA) can be elongated to 18–24 carbons by the sequential addition of two-carbon units in the endoplasmic reticulum (ER) or sometimes in the mitochondrion. Desaturase enzymes, also located in the ER, introduce double bonds at specific positions in some fatty acids, yielding **unsaturated** fatty acids. Oleyl CoA (oleate linked to CoA; see Table 2-4), for example, is formed by removal of two H atoms from stearyl CoA. In contrast to free fatty acids, fatty acyl CoA derivatives are soluble in aqueous solutions because of the hydrophilicity of the CoA segment.

Small Cytosolic Proteins Facilitate Movement of Fatty Acids

In order to be transported through the cell cytosol, free, or unesterified, fatty acids (those unlinked to CoA) are commonly bound by *fatty acid–binding proteins (FABPs)*, which belong to a group of small cytosolic proteins that act as chaperones to facilitate the intracellular movement of many lipids. These proteins contain a hydrophobic pocket lined by β sheets (Figure 10-24). A long-chain fatty acid can fit into this pocket and interact noncovalently with the surrounding protein.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 10-24 Binding of a fatty acid to the hydrophobic pocket of a fatty acid–binding protein (FABP). The crystal structure of adipocyte FABP (ribbon diagram) shows that the hydrophobic binding pocket is generated from two β sheets that are nearly at right angles to each other, forming a clam shell–like structure. A fatty acid (carbons yellow; oxygens red) interacts noncovalently with hydrophobic amino acid residues within this pocket.

[Data from Z. Xu, D. A. Bernlohr, and L. J. Banaszak, 1993, *J. Biol. Chem.* **268**:7874, PDB ID 1lid.]

The expression of cellular FABPs is regulated coordinately with cellular requirements for the uptake and release of fatty acids. Thus FABP levels are high in active muscles that are using fatty acids for generation of ATP, and in adipocytes (fat-storing cells) when they

are either taking up fatty acids to be stored as triglycerides or releasing fatty acids for use by other cells. The importance of FABPs in fatty acid metabolism is highlighted by the observations that they can compose as much as 5 percent of all cytosolic proteins in the liver and that genetic inactivation of cardiac muscle FABP converts the heart from a muscle that primarily burns fatty acids for energy into one that primarily burns glucose.

Fatty Acids Are Incorporated into Phospholipids Primarily on the ER Membrane

Fatty acids are not directly incorporated into phospholipids; rather, in eukaryotic cells, they are first converted into CoA esters. The subsequent synthesis of phospholipids such as the phosphoglycerides is carried out by enzymes associated with the cytosolic face of the ER membrane, usually the smooth ER, in animal cells; through a series of steps, fatty acyl CoAs, glycerol 3-phosphate, and polar head group precursors are linked together and then inserted into the ER membrane (Figure 10-25). The fact that the enzymes involved in this process are located on the cytosolic side of the membrane means that there is an inherent asymmetry in membrane biogenesis: new membranes are initially synthesized only on one leaflet — a fact with important consequences for the asymmetric distribution of lipids in membrane leaflets. Once

synthesized on the ER, phospholipids are transported to other organelles and to the plasma membrane. Mitochondria synthesize some of their own membrane lipids and import others.

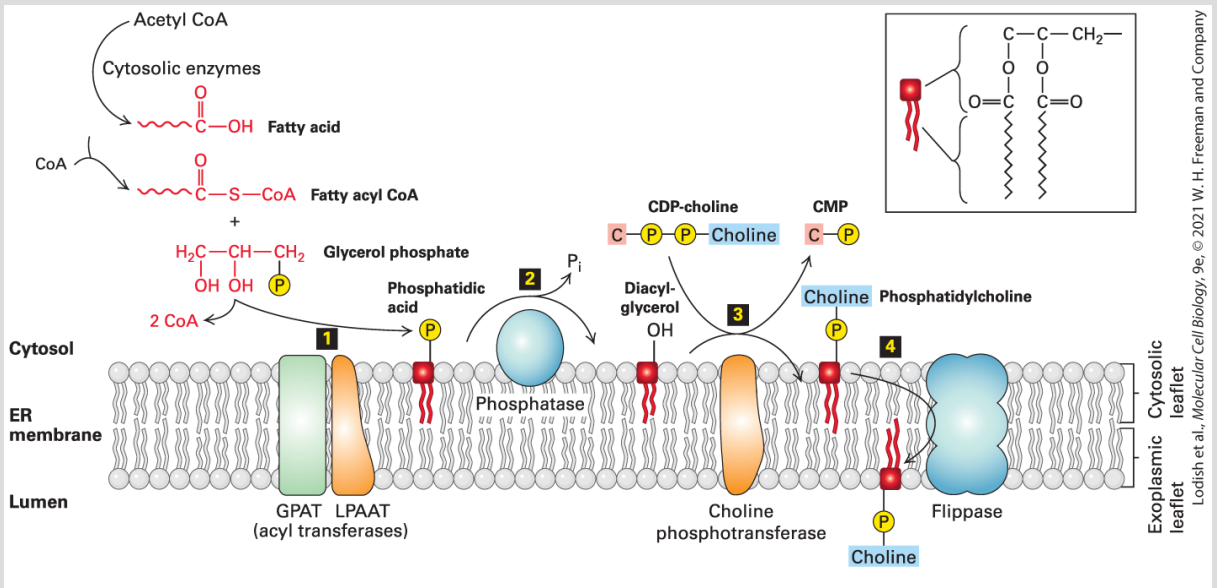


FIGURE 10-25 Phospholipid synthesis in the ER membrane. Because phospholipids are amphipathic molecules, the last steps of their multistep synthesis take place at the interface between a membrane and the cytosol and are catalyzed by membrane-associated enzymes. Step 1: Two fatty acids from fatty acyl CoA are esterified to the phosphorylated glycerol backbone, forming phosphatidic acid, whose two long hydrocarbon chains anchor the molecule to the membrane. Step 2: A phosphatase converts phosphatidic acid into diacylglycerol. Step 3: A polar head group (e.g., phosphorylcholine) is transferred from cytosine diphosphocholine (CDP-choline) to the exposed hydroxyl group. Step 4: Flippase proteins catalyze the movement of phospholipids from the cytosolic leaflet in which they are initially formed to the exoplasmic leaflet.

Sphingolipids are also synthesized indirectly from multiple precursors. Sphingosine, the building block of these lipids, is made in the ER, beginning with the coupling of a palmitoyl group from palmitoyl CoA to serine; the subsequent addition of a second fatty

acyl group to form *N*-acyl sphingosine (ceramide) also takes place in the ER. Later, in the Golgi, a polar head group is added to ceramide, yielding *sphingomyelin*, whose head group is phosphorylcholine, and various *glycosphingolipids*, in which the head group may be a monosaccharide or a more complex oligosaccharide (see Figure 10-8b). Some sphingolipid synthesis can also take place in mitochondria. In addition to serving as the backbone for sphingolipids, ceramide and its metabolic products are important signaling molecules that can influence cell growth, proliferation, endocytosis, resistance to stress, and programmed cell death (apoptosis).

After their synthesis is completed in the Golgi, sphingolipids are transported to other cellular compartments through vesicle-mediated mechanisms similar to those for the transport of proteins, discussed in Chapter 14. Any type of vesicular transport results in movement not only of the protein payload but also of the lipids that compose the vesicular membrane. Phospholipids such as phosphoglycerides, as well as cholesterol, can move between organelles by additional mechanisms, described below.

Flippases Move Phospholipids from One Membrane Leaflet to the Opposite Leaflet

Even though phospholipids are initially incorporated into the cytosolic leaflet of the ER membrane, various phospholipids are asymmetrically distributed in the two leaflets of the ER membrane and of other cellular membranes. As noted above, phospholipids do not readily flip-flop from one leaflet to the other. For the ER membrane to expand by growth of both leaflets and to have asymmetrically distributed phospholipids, its phospholipid components must be able to move from one leaflet to the other. Although the mechanisms employed to generate and maintain membrane phospholipid asymmetry are not well understood, it is clear that flippases play a key role. As will be described in Chapter 11, these integral membrane proteins use the energy of ATP hydrolysis to facilitate the movement of phospholipid molecules from one leaflet to the other. When cells die, this lipid asymmetry is lost and results in increased surface exposure of those phospholipids mostly confined to the cytoplasmic leaflet, such as phosphatidylserine.

Cholesterol Is Synthesized by Enzymes in the Cytosol and ER Membrane

Next we focus on cholesterol, the principal sterol in animal cells. Cholesterol is synthesized mainly in the liver. The first steps of cholesterol synthesis (Figure 10-26) — conversion of three acetyl groups linked to CoA (acetyl CoA) forming the six-carbon molecule

β-hydroxy-β-methylglutaryl linked to CoA (HMG-CoA) — take place in the cytosol. The conversion of HMG-CoA into mevalonate, the key rate-controlling step in cholesterol biosynthesis, is catalyzed by *HMG-CoA reductase*, an ER integral membrane protein, even though both its substrate and its product are water soluble. The water-soluble catalytic domain of HMG-CoA reductase extends into the cytosol, but its eight transmembrane α helices firmly embed the enzyme in the ER membrane. Five of the transmembrane α helices compose the so-called *sterol-sensing domain* and regulate enzyme stability. When levels of cholesterol in the ER membrane are high, binding of cholesterol to this domain causes the enzyme to bind to two other integral ER membrane proteins, Insig-1 and Insig-2. This binding, in turn, induces ubiquitylation (see Figure 3-32) of HMG-CoA reductase and its degradation by the proteasome pathway, reducing the production of mevalonate, the key intermediate in cholesterol biosynthesis.

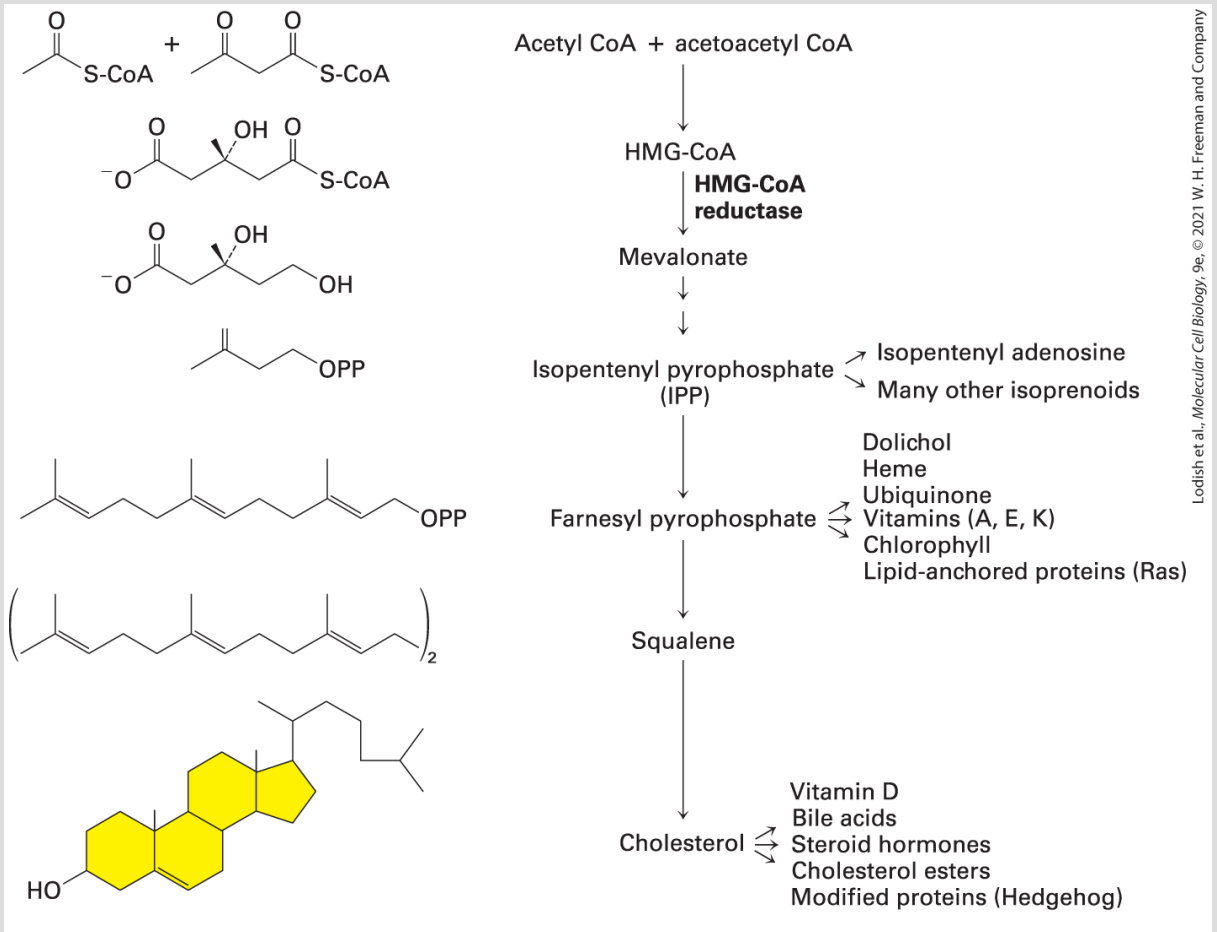


FIGURE 10-26 Cholesterol biosynthetic pathway. The regulated rate-controlling step in cholesterol biosynthesis is the conversion of β-hydroxy-β-methylglutaryl CoA (HMG-CoA) into mevalonate by HMG-CoA reductase, an ER-membrane protein. Mevalonate is then converted into isopentenyl pyrophosphate (IPP), which has the basic five-carbon isoprenoid structure. IPP can be converted into cholesterol and many other lipids, often through the polyisoprenoid intermediates shown. Some of the numerous compounds derived from isoprenoid intermediates and cholesterol itself are indicated.

Atherosclerosis, frequently called *cholesterol-dependent clogging of the arteries*, is characterized by the progressive deposition of cholesterol and other lipids, cells, and extracellular matrix material in the inner layer of the wall of an artery. The resulting distortion of the artery's wall can lead, either alone or in combination with a

blood clot, to major blockage of blood flow. Atherosclerosis accounts for 75 percent of deaths due to cardiovascular disease in the United States.

Perhaps the most successful anti-atherosclerosis medications are the *statins*. These drugs bind to HMG-CoA reductase and directly inhibit its activity, thereby lowering cholesterol biosynthesis. As a consequence, the concentration of low-density lipoproteins (see Figure 14-27) — the small, membrane-enveloped particles containing cholesterol esterified to fatty acids that often and rightly are called "bad cholesterol" — drops in the blood, reducing the formation of atherosclerotic plaques. ■

Mevalonate, the six-carbon product formed by HMG-CoA reductase, is converted in several steps into the five- carbon isoprenoid compound isopentenyl pyrophosphate (IPP) and its stereoisomer, dimethylallyl pyrophosphate (DMPP) (see Figure 10-26). These reactions are catalyzed by cytosolic enzymes, as are the subsequent reactions in the cholesterol synthesis pathway, in which six IPP units condense to yield squalene, a branched-chain 30-carbon intermediate. Enzymes bound to the ER membrane catalyze the multiple reactions that convert squalene into cholesterol in mammals or into related sterols in other species. One of the intermediates in this pathway, farnesyl pyrophosphate, is the precursor of the prenyl lipid that anchors Ras and related proteins to

the cytosolic surface of the plasma membrane (see Figure 10-19) as well as other important biomolecules (see Figure 10-26).

Cholesterol and Phospholipids Are Transported Between Organelles by Several Mechanisms

As already noted, the final steps in the synthesis of cholesterol and phospholipids take place primarily in the ER. Thus the plasma membrane and the membranes bounding other organelles must obtain these lipids by means of one or more intracellular transport processes. Membrane lipids can and do accompany both soluble and membrane proteins along the secretory pathway described in Chapter 14: membrane vesicles bud from the ER and fuse with membranes in the Golgi complex, and other membrane vesicles bud from the Golgi complex and fuse with the plasma membrane (Figure 10-27a). However, evidence suggests that there is substantial inter-organelle movement of cholesterol and phospholipids through other mechanisms. For example, chemical inhibitors of the classic secretory pathway and mutations that impede vesicular traffic along this pathway do not prevent cholesterol or phospholipid transport between membranes.

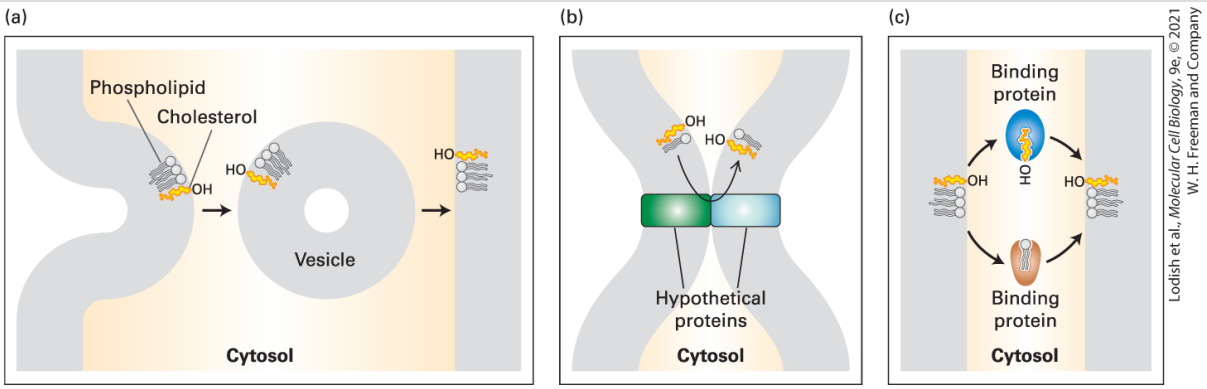


FIGURE 10-27 Proposed mechanisms of transport of cholesterol and phospholipids between membranes. In mechanism (a), vesicles transfer lipids between membranes. In mechanism (b), lipid transfer is a consequence of direct contact between membranes that is mediated by membrane-embedded proteins. In mechanism (c), transfer is mediated by small, soluble lipid-transfer proteins.

[Information from F. R. Maxfield and D. Wustner, 2002, *J. Clin. Invest.* **110**:891.]

A second proposed mechanism of lipid movement entails direct protein-mediated contact of ER or ER-derived membranes with the membranes of other organelles (Figure 10-27b). In a third proposed mechanism, small lipid-transfer proteins facilitate the exchange of phospholipids or cholesterol between different membranes (Figure 10-27c). Although such lipid-transfer proteins have been identified in assays in vitro, their role in the intracellular movements of most phospholipids is not well defined. For instance, mice with a knockout mutation in the gene encoding the phosphatidylcholine-transfer protein appear to be normal in most respects, indicating that this protein is not essential for cellular phospholipid metabolism.

As noted earlier, the lipid compositions of different organelle membranes vary considerably (see Table 10-1). Some of these differences are due to different sites of synthesis. For example, a phospholipid called *cardiolipin*, which is localized to the mitochondrial membrane, is made only in mitochondria, and little of it is transferred to other organelles. Differential transport of lipids also plays a role in determining the lipid compositions of different cellular membranes. For instance, even though cholesterol is made in the ER, the cholesterol concentration (cholesterol-to-phospholipid molar ratio) is 1.5–13-fold higher in the plasma membrane than in other organelles (ER, Golgi, mitochondrion, lysosome). Although the mechanisms responsible for establishing and maintaining these differences are not well understood, we have seen that the distinctive lipid composition of each membrane has a major influence on its physical and biological properties.

KEY CONCEPTS OF SECTION 10.3

Phospholipids, Sphingolipids, and Cholesterol: Synthesis and Intracellular Movement

- Saturated and unsaturated fatty acids of various chain lengths are components of phospholipids and sphingolipids.
- Fatty acids are synthesized from acetyl CoA by water-soluble enzymes and modified by elongation and desaturation in the endoplasmic reticulum (ER).
- Free fatty acids are transported within cells by fatty acid–binding proteins (FABPs).
- Fatty acids are incorporated into phospholipids through a multistep process. The final steps in the synthesis of phosphoglycerides and sphingolipids are catalyzed by membrane-associated enzymes primarily on the cytosolic face of the ER (see Figure 10-25).

- Each type of newly synthesized lipid is incorporated into the preexisting membranes on which it is made; thus membranes are themselves the platform for the synthesis of new membrane material.
- Most membrane phospholipids are preferentially distributed in either the exoplasmic or the cytosolic leaflet. This asymmetry results in part from the action of flippases, which flip phospholipids from one leaflet to the other.
- The initial steps in cholesterol biosynthesis take place in the cytosol, whereas the last steps are catalyzed by enzymes associated with the ER membrane.
- The rate-controlling step in cholesterol biosynthesis is catalyzed by HMG-CoA reductase, whose transmembrane segments are embedded in the ER membrane and contain a sterol-sensing domain.
- Considerable evidence indicates that vesicular transport, direct protein-mediated contacts between different membranes, soluble lipid-transfer proteins, or all three may account for some inter-organelle transport of cholesterol and phospholipids (see Figure 10-27).

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

acetyl CoA
amphipathic
aquaporin
cholesterol
cilium
cytoskeleton
cytosolic face
exoplasmic face
flagellum
glycolipid
glycoprotein
hydrophilic

hydrophobic
integral membrane protein
lectin
lipid-anchored membrane protein
lipid raft
liposome
membrane transport protein
micelle
peripheral membrane protein
phosphoglyceride
phospholipase
phospholipid bilayer
plasma membrane
porin
receptor protein
saturated
sphingolipid
unsaturated

Review the Concepts

1. When viewed by electron microscopy, the lipid bilayer is often described as looking like a railroad track. Explain how the structure of the bilayer creates this image.
2. Explain the following statement: The structure of all biomembranes depends on the chemical properties of phospholipids, whereas the function of each specific

biomembrane depends on the specific proteins associated with that membrane.

3. Biomembranes contain many different types of lipid molecules. What are the three main types of lipid molecules found in biomembranes? How are the three types similar, and how are they different?
4. Lipid bilayers are said to behave like two-dimensional fluids. What does this mean? What drives the movement of lipid molecules and proteins within the bilayer? How can such movement be measured? What factors affect the degree of membrane fluidity?
5. Why are water-soluble substances unable to freely cross the lipid bilayer of the plasma membrane? How does the cell overcome this permeability barrier?
6. Name the three groups into which membrane-associated proteins may be classified. Explain the mechanism by which each group associates with a biomembrane.
7. Identify the following membrane-associated proteins based on their structure: (a) tetramers of identical subunits, each with six membrane-spanning α helices; (b) trimers of identical subunits, each with 16 β sheets forming a barrel-like structure.
8. Proteins may be bound to the exoplasmic or cytosolic face of the plasma membrane by covalently attached lipids. What are the three types of lipid anchors responsible for tethering proteins to the plasma-membrane bilayer? Which type is used by cell-surface proteins that face the external medium? By glycosylated proteoglycans?

9. Although both faces of a biomembrane are composed of the same general types of macromolecules, principally lipids and proteins, the two faces of the bilayer are not identical. What accounts for the asymmetry between the two faces?
10. What are detergents? How do ionic and non-ionic detergents differ in their ability to disrupt biomembrane structure?
11. What is the likely identity of these membrane-associated proteins: (a) a protein that is released from a membrane treated with a high-salt solution, which causes disruption of ionic linkages; (b) a protein that is not released from the membrane upon its exposure to a high-salt solution alone, but is released when the membrane is incubated with an enzyme that cleaves phosphate-glycerol bonds and covalent linkages are disrupted; (c) a protein that is not released from the membrane upon exposure to a high-salt solution, but is released after the addition of the detergent sodium dodecylsulfate (SDS). Will the activity of the protein released in part (c) be preserved following its release?
12. Following the production of membrane extracts using the non-ionic detergent Triton X-100, you analyze the membrane lysates via mass spectrometry and note a high content of cholesterol and sphingolipids. Furthermore, biochemical analysis of the lysates shows potential kinase activity. What have you probably isolated?
13. Phospholipid biosynthesis at the interface between the endoplasmic reticulum (ER) and the cytosol presents a

number of challenges that must be solved by the cell. Explain how each of the following is handled.

a. The substrates for phospholipid biosynthesis are all water soluble, yet the end products are not.
b. The immediate site of incorporation of all newly synthesized phospholipids is the cytosolic leaflet of the ER membrane, yet phospholipids must be incorporated into both leaflets.
c. Many membrane systems in the cell, such as the plasma membrane, are unable to synthesize their own phospholipids, yet these membranes must also expand if the cell is to grow and divide.

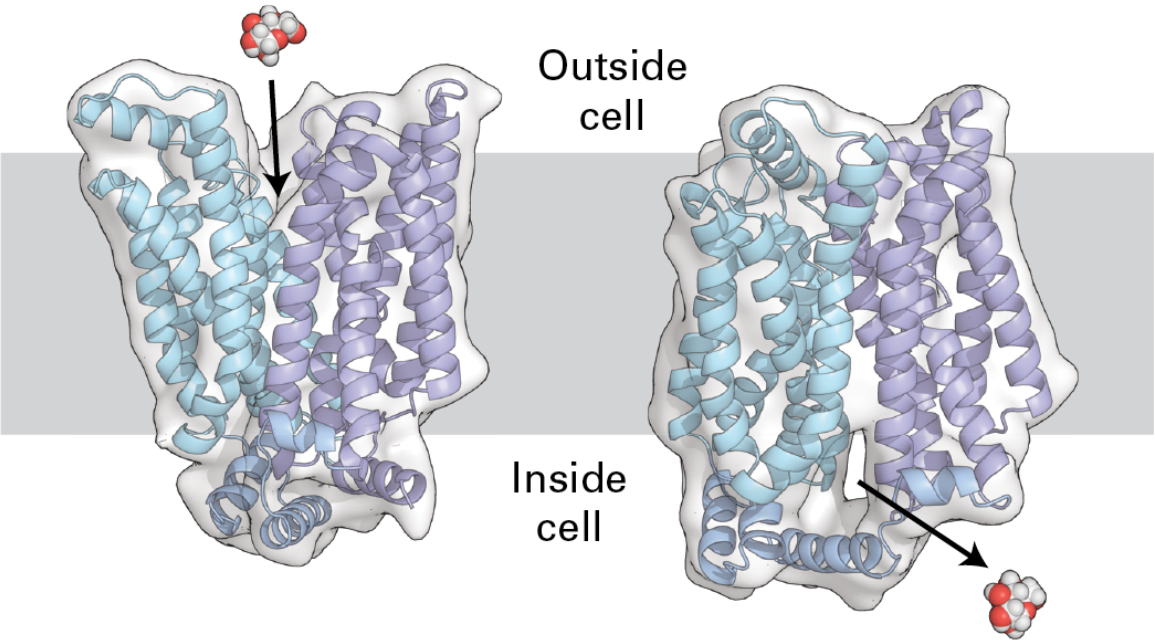
14. What are the common fatty acid chains in phosphoglycerides, and why do these fatty acid chains differ in their number of carbon atoms by multiples of 2?
15. Fatty acids must associate with lipid chaperones in order to move within the cell. Why are these chaperones needed, and what is the name given to a group of proteins that are responsible for this intracellular trafficking of fatty acids? What is the key distinguishing feature of these proteins that allows fatty acids to move within the cell?
16. What is the key enzyme in the highly regulated process of cholesterol biosynthesis? This enzyme is subject to feedback inhibition. What is feedback inhibition? How does this enzyme sense cholesterol levels in a cell?
17. Phospholipids and cholesterol must be transported from their site of synthesis to various membrane systems within cells. One way of doing this is through vesicular transport,

as is the case for many proteins in the classic secretory pathway (see Chapter 14). However, phospholipid and cholesterol membrane-to-membrane transport in cells does not occur solely by vesicular transport. What is the evidence for this? What appear to be the major mechanisms for phospholipid and cholesterol transport?

18. Explain the mechanism by which statins lower “bad” cholesterol.

CHAPTER 11

Transmembrane Transport of Ions and Small Molecules



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Transmembrane, side view of the human glucose transporter GLUT1. Like other glucose transporters, GLUT1 transports glucose across the cell's plasma membrane by alternating between an outward-facing and an inward-facing state. In the outward-facing state, a molecule of glucose binds to GLUT1. Glucose binding changes the conformation of the transporter such that it is now open to the inside of the cell, releasing glucose into the cytoplasm.

[Data from D. Deng et al., 2014 *Nature* **510**:121–125, PDB ID 4PYP.]

OUTLINE

11.1 Overview of Transmembrane Transport
11.2 Facilitated Transport of Glucose and Water
11.3 ATP-Powered Pumps and the Intracellular Ionic Environment
11.4 Nongated Ion Channels and the Resting Membrane Potential
11.5 Cotransport by Symporters and Antiporters
11.6 Transcellular Transport

In all cells, the plasma membrane forms the barrier that separates the cytoplasm from the exterior environment, thus defining a cell's physical and chemical boundaries. By preventing the unimpeded movement of molecules and ions into and out of the cell, the plasma membrane maintains essential differences between the composition of the fluid inside the cell (the cytosol) and the fluid outside the cell (the extracellular fluid). For example, the concentration of sodium ion (Na^+) in the cytosol of animal cells is about 15 mM, while the concentration of sodium chloride (NaCl) in extracellular fluids, including blood, of animals is tenfold higher, about 150 mM. Conversely, the potassium ion (K^+) concentration is tenfold higher in the cytosol of animal cells than in the extracellular fluid. As we describe in Section 11.4, these differences in the distribution of ions across the plasma membrane are critical to establishing and maintaining the membrane potential of the cell.

Organelle membranes, which separate the cytosol from the interior of the organelle, also form permeability barriers. For example, the proton concentration in the lysosome interior, pH 5, is about a hundredfold greater than that in the cytosol, and many specific metabolites accumulate at higher concentrations in the interior of

other organelles, such as the endoplasmic reticulum or the Golgi complex, than in the cytosol.

All cellular membranes, both plasma membranes and organelle membranes, consist of a bilayer of phospholipids in which other lipids and specific types of proteins are embedded. It is this combination of lipids and proteins that gives cellular membranes their distinctive permeability qualities. If cellular membranes were pure phospholipid bilayers (see Figure 9-5), they would be excellent chemical barriers, impermeable to virtually all ions, amino acids, sugars, and other water-soluble molecules. In fact, only a few gases and small, uncharged, water-soluble molecules can readily diffuse across a pure phospholipid bilayer (Figure 11-1). Cellular membranes must serve not only as barriers but also as conduits, selectively transporting molecules and ions from one side of the membrane to the other. Energy-rich glucose, for example, must be imported into the cell, and wastes must be shipped out.

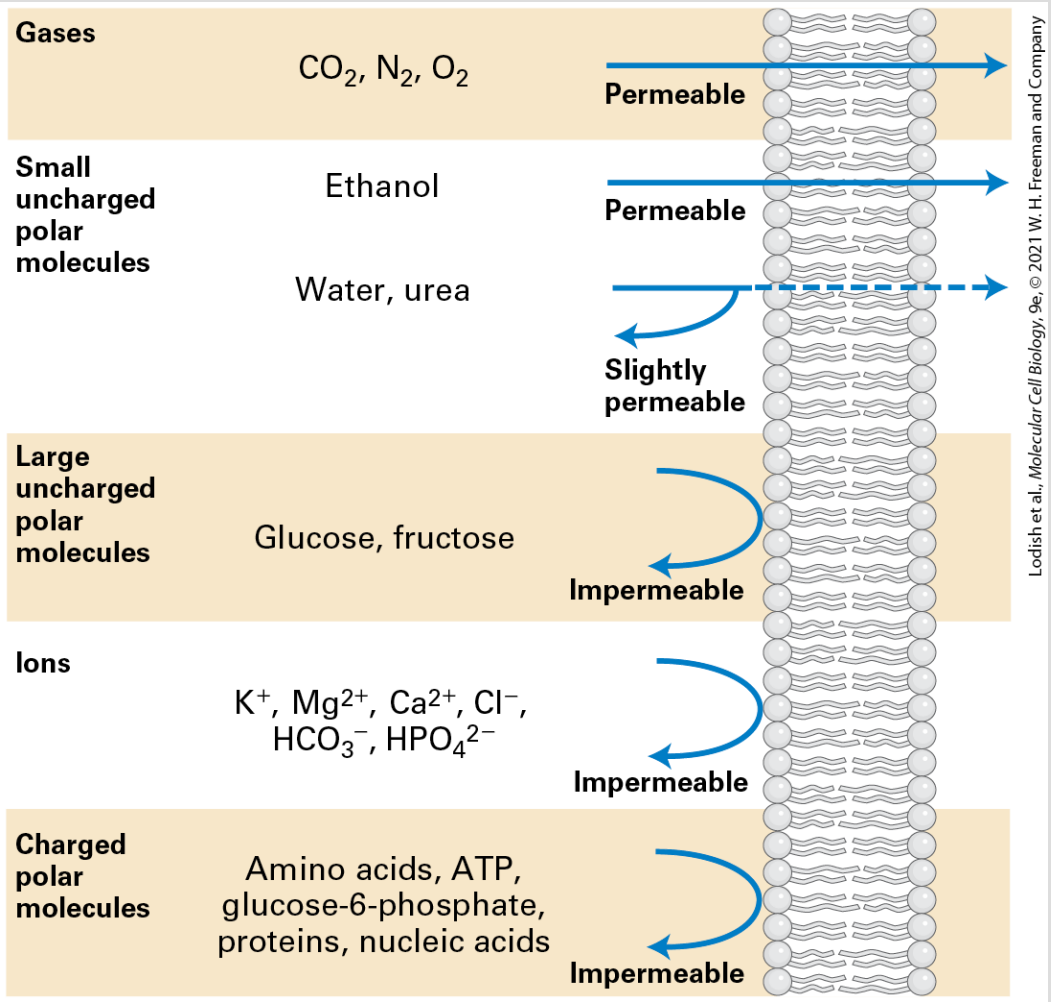


FIGURE 11-1 Relative permeability of a pure phospholipid bilayer to various molecules and ions. A pure phospholipid bilayer is permeable to many gases and to small, uncharged, water-soluble (polar) molecules. It is slightly permeable to water, and essentially impermeable to ions and to large polar molecules.

Movement of virtually all small molecules and ions across cellular membranes is mediated by **membrane transport proteins** — integral membrane proteins with multiple transmembrane domains embedded in cellular membranes. These membrane-spanning proteins act as shuttles, channels, or pumps for transporting

molecules and ions through a membrane's hydrophobic interior. In some cases, these proteins transport molecules or ions from a higher to a lower concentration, a thermodynamically favored process powered by an increase in entropy. Examples include the transport of water or glucose from the blood into most body cells. In other cases, a transport protein pumps molecules or ions from a lower to a higher concentration, a thermodynamically unfavorable process that can occur only when an external source of energy is available to push the molecules *uphill* against a concentration gradient. An example of such a process is the movement of protons into lysosomes to generate a low pH in the lumen. Often the required energy is provided by mechanistic coupling of the energy-releasing hydrolysis of ATP with the movement of a molecule or ion across the membrane. Other transport proteins couple the movement of one molecule or ion against its concentration gradient with the movement of another down its gradient, using the energy released by the *downhill* movement of one molecule or ion to drive the uphill movement of another. Proper functioning of any cell relies on a precise balance between the import and export of various molecules and ions.

We begin our discussion of membrane transport proteins by reviewing some of the general principles of transport across membranes and distinguishing among three major classes of such proteins. In subsequent sections, we describe the structure and operation of specific examples of each class and show how members

of families of homologous transport proteins have different properties that enable different cell types to function appropriately. We also explain how specific combinations of transport proteins in both the plasma membrane and organelle membranes enable cells to carry out essential physiological processes, including the maintenance of cytosolic pH, the direction of the flow of water in both plant and animal cells, and the transport of sugar and amino acids into cells.

The identification of many disease-causing mutations in membrane transporters underscores the importance of this class of molecules to human health. Mutations in these proteins cause a range of diseases, including cystic fibrosis and long QT syndrome (a potentially fatal form of cardiac arrhythmia) as well as specific forms of epilepsy, migraine, ataxia, polycystic kidney disease, and congenital hearing loss. As highlighted throughout this chapter, channels, transporters, and pumps are the targets of a growing number of therapeutic drugs. Recent advances in powerful new imaging technologies, such as cryo-electron microscopy (cryo-EM, see Figure 4-34), have given rise to an explosion of discoveries about the structure and function of many of these membrane proteins. These discoveries in turn promise to lead to the development of new, lifesaving therapies. Our increased understanding of channel structure and function has also led to the development of heterologously expressed channels that can be activated by light (optogenetic) or drugs (chemogenetic). These have become powerful

tools to control biochemical processes (described in Chapter 4) and neural circuits (described in Chapter 23).

Note that in this chapter we cover only transport of small molecules and ions; transport of larger molecules, such as proteins and oligosaccharides, is covered in Chapters 13 and 14.

11.1 Overview of Transmembrane Transport

In this section, we briefly describe the factors that influence the permeability of lipid membranes to molecules, then briefly describe the three major classes of membrane transport proteins that allow molecules and ions to cross them. Different kinds of membrane-embedded proteins accomplish the task of moving molecules and ions in different ways.

Only Gases and Small Uncharged Molecules Cross Membranes by Simple Diffusion

With its dense hydrophobic core, a phospholipid bilayer is largely impermeable to water-soluble molecules and ions. Only gases, such as O_2 and CO_2, and small uncharged polar molecules, such as urea and ethanol, can readily move across an artificial membrane composed of pure phospholipid or of phospholipid and cholesterol (see Figure 11-1). Such molecules can also diffuse across cellular membranes without the aid of transport proteins. No metabolic energy is expended during simple diffusion because movement is from a high to a low concentration of the molecule, down its chemical concentration gradient. Such movements are spontaneous

because they have a positive ΔS value (increase in entropy) and thus a negative ΔG (decrease in free energy). The diffusion rate of any substance across a pure phospholipid bilayer is proportional to its concentration gradient across the bilayer and to its hydrophobicity and size.

If a substance carries a net charge, its movement across a membrane is influenced by both its concentration gradient and the **membrane potential**, the electric potential (voltage) across the membrane. The combination of these two forces, called the **electrochemical gradient**, determines the energetically favorable direction of movement of a charged molecule across a membrane (Figure 11-2). The electric potential that exists across most cellular membranes results from a small imbalance in the concentrations of positively and negatively charged ions on the two sides of the membrane. In Section 11.4, we discuss how this ionic imbalance, and the resulting potential, arise and are maintained.

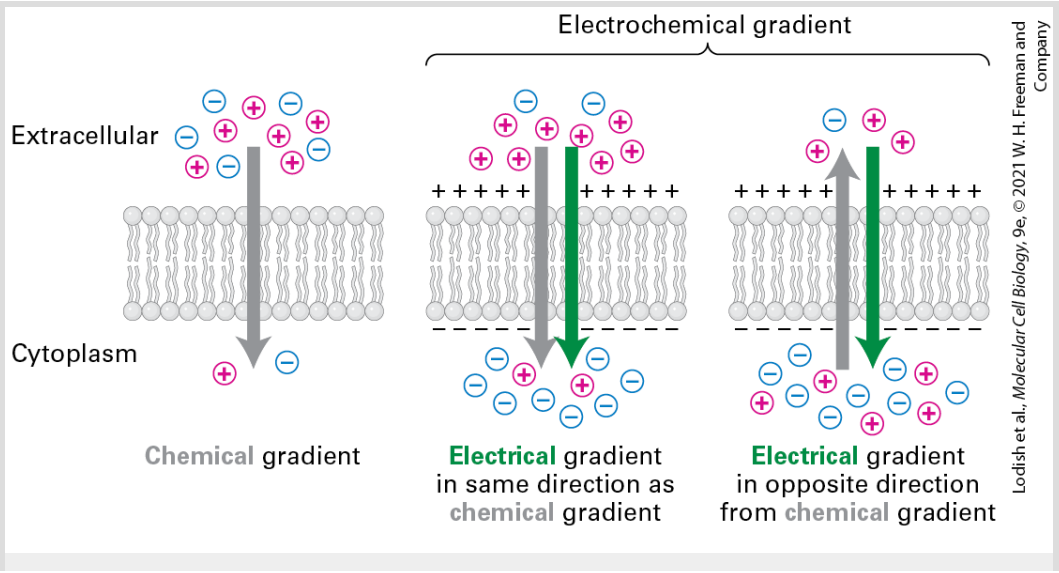


FIGURE 11-2 Electrochemical gradients. Ions form a chemical gradient when there is a concentration difference across the membrane. Ions also form an electrical gradient across the membrane, created by the asymmetric distribution of positively and negatively charged ions across the membrane. The electrical gradient can be in the same or opposite direction from the chemical gradient. Together, they form an electrochemical gradient that serves as a force driving the transport of ions across the membrane.

Three Main Classes of Membrane Proteins Transport Molecules and Ions Across Cellular Membranes

As is evident from Figure 11-1, very few molecules and no ions can cross a pure phospholipid bilayer at appreciable rates by simple diffusion. Thus transport of most molecules into and out of cells requires the assistance of specialized membrane proteins. Even in the cases of some small uncharged polar molecules (e.g., urea, fatty acids) and certain gases, such as CO_2 (carbon dioxide) and NH_3 (ammonia), transport is frequently accelerated by specific proteins because simple diffusion is too slow to meet cellular needs.

All membrane transport proteins are transmembrane proteins containing multiple membrane-spanning segments that are generally α helices. By forming a protein-lined pathway across the membrane, transport proteins are thought to allow hydrophilic substances to move through the membrane without coming into

contact with its hydrophobic interior. Here we introduce the three main types of membrane transport proteins covered in this chapter (Figure 11-3).

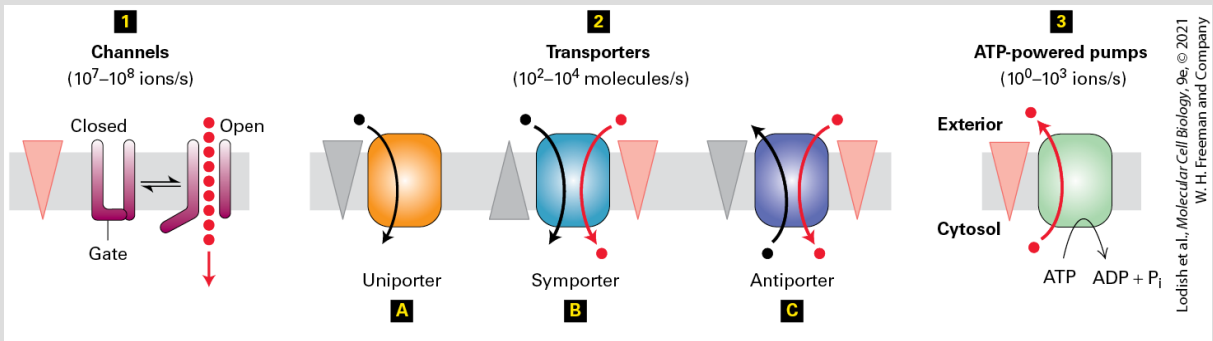


FIGURE 11-3 Overview of membrane transport proteins. Gradients are indicated by triangles with the tip pointing toward lower concentration, electric potential, or both. **1** Channels permit movement of specific ions (or water) down their electrochemical gradient. **2** Transporters, which fall into three groups, facilitate movement of specific small molecules or ions. Uniporters transport a single type of molecule down its concentration gradient **2A**. Cotransport proteins (symporters, **2B**, and antiporters, **2C**) catalyze the movement of one molecule *against* its concentration gradient (black circles), driven by movement of one or more ions down an electrochemical gradient (red circles). **3** Pumps use the energy released by ATP hydrolysis to power movement of specific ions or small molecules (red circles) against their electrochemical gradient. Differences in the mechanisms of transport by these three major classes of proteins account for their varying rates of solute movement.

Channels transport water, specific ions, or small hydrophilic molecules across membranes *down* their concentration or electric potential gradients. Because this process requires transport proteins but not energy, it is sometimes referred to as *passive transport* or *facilitated diffusion*, but it is more properly called **facilitated transport**. Channels form a hydrophilic tube or passageway across

the membrane through which multiple water molecules or ions move simultaneously, single file, at a very rapid rate. Some channels are open much of the time; they are referred to as *nongated channels*. Most ion channels, however, open only in response to specific chemical or electrical signals. These channels are referred to as *gated channels* because a protein "gate" alternatively blocks the channel or moves out of the way to open the channel (see Figure 11-3). Like all transport proteins, channels are very selective for the type of molecule they transport.

Transporters (also called *carriers*) move a wide variety of ions and molecules across cellular membranes, but at a much slower rate than channels. Three types of transporters have been identified. *Uniporters* transport a single type of molecule *down* its concentration gradient. Glucose and amino acids cross the plasma membrane into most mammalian cells with the aid of uniporters. Collectively, channels and uniporters are sometimes called *facilitated transporters*, indicating movement down a concentration or electrochemical gradient.

In contrast, **antiporters** and **symporters** couple the movement of one type of ion or molecule *against* its concentration gradient with the movement of one or more different ions *down* its concentration gradient, in the same (symporter) or different (antiporter) directions. These proteins are often called **cotransporters** because

of their ability to transport two or more different solutes simultaneously.

ATP-powered pumps (or simply *pumps*) are ATPases, enzymes that catalyze hydrolysis of the terminal phosphoanhydride bond of ATP to produce ADP and Pi. These pumps use the energy of ATP hydrolysis to move ions or small molecules across a membrane *against* a chemical concentration gradient, an electric potential, or both. This process, referred to as **active transport**, is an example of a coupled chemical reaction (see Chapter 2). In this case, transport of ions or small molecules uphill against an electrochemical gradient, which requires energy, is coupled to the hydrolysis of ATP, which releases energy. The overall reaction — ATP hydrolysis and the uphill movement of ions or small molecules — is energetically favorable.

Both cotransporters and pumps rely on coupled reactions in which an energetically unfavorable reaction (i.e., the uphill movement of one type of molecule or ion) is coupled to an energetically favorable reaction (i.e., the downhill movement of another). Note, however, that the nature of the energy-supplying reaction driving active transport by these two classes of proteins differs. ATP pumps use energy from hydrolysis of ATP, whereas cotransporters use the energy stored in an electrochemical gradient. The latter process is sometimes referred to as *secondary active transport*.

Conformational changes are essential to the function of all transport proteins. Whereas channels switch between open and closed states, transporters and ATP-powered pumps undergo a cycle of conformational change exposing a binding site (or sites) to one side of the membrane in one conformation and to the other side in a second conformation, a process that has been termed the *alternating access model*. Determination of the high-resolution structures of transporters has revealed distinct molecular mechanisms by which alternating access is achieved. Because each cycle of conformational change within transporters and pumps results in the movement of only one substrate molecule (or, at most, a few), these proteins are characterized by relatively slow rates of transport, ranging from 10^0 to 10^4 ions or molecules per second (see Figure 11-3). In contrast, many ions can pass through an open channel without any further conformational change. For this reason, channels are characterized by very fast rates of transport, up to 10^8 ions per second.

Usually several different types of transport proteins work in concert within the plasma membrane of cells to achieve a physiological function. An example is seen in Figure 11-4, where an ATPase pumps Na^+ out of the cell and K^+ inward; this pump, which is found in virtually all metazoan cells, establishes the oppositely directed concentration gradients of Na^+ and K^+ ions across the plasma membrane (relatively high concentrations of K^+ inside and of Na^+ outside cells). These concentration gradients are used, for example, to power the import of amino acids by cotransporters. The human

genome encodes hundreds of different types of transport proteins that use the energy stored across the plasma membrane in the Na^+ concentration gradient and its associated electric potential to transport a wide variety of molecules into cells against their concentration gradients.

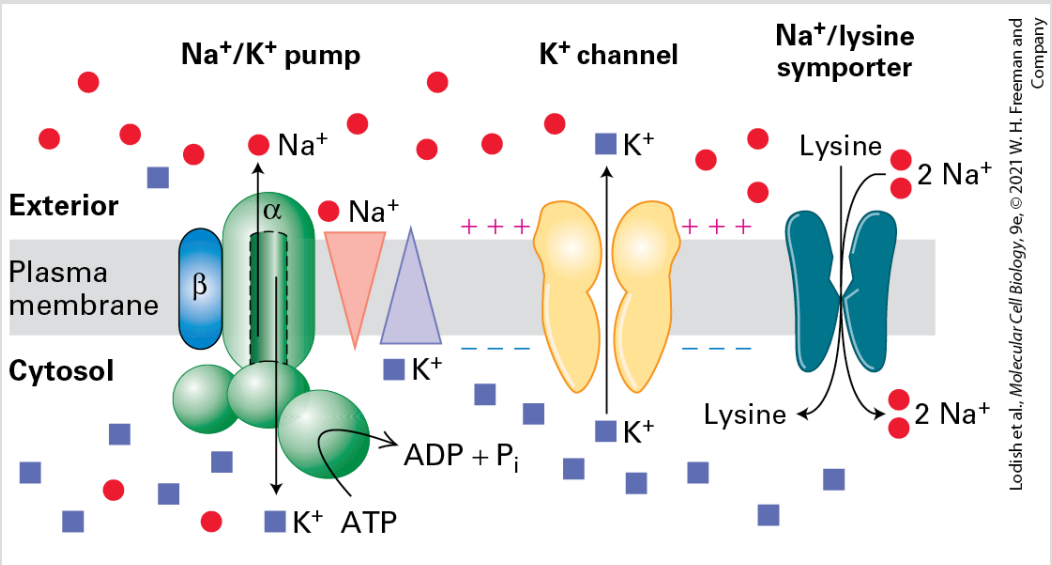


FIGURE 11-4 Multiple membrane transport proteins function together in the plasma membrane of metazoan cells. Gradients are indicated by triangles with the tip pointing toward lower concentration. The Na^+/K^+ ATPase in the plasma membrane uses energy released by ATP hydrolysis to pump Na^+ (red circles) out of the cell and K^+ (blue squares) into the cell; its action creates a concentration gradient of Na^+ that is greater outside the cell than inside, and one of K^+ that is greater inside than outside. Movement of positively charged K^+ ions out of the cell through membrane K^+ channels creates an electric potential across the plasma membrane — the cytosolic face is negative with respect to the extracellular face. A Na^+/lysine transporter, a typical sodium/amino acid cotransporter, moves two Na^+ ions together with one lysine from the extracellular medium into the cell. "Uphill" movement of the amino acid is powered by "downhill" movement of Na^+ ions, which in turn is powered both by the outside-greater-than-inside Na^+ concentration gradient and by the negative charge on the inside of the plasma membrane, which

attracts the positively charged Na^+ ions. The ultimate source of the energy to power amino acid uptake comes from the ATP hydrolyzed by the Na^+/K^+ ATPase, since this pump creates both the Na^+ ion concentration gradient and, via the K^+ channels, the membrane potential, which together power the influx of Na^+ ions.

Table 11-1 summarizes the four mechanisms by which small molecules and ions are transported across cellular membranes. In the next section, we consider some of the simplest membrane transport proteins, the uniporter responsible for the transport of glucose and the channel responsible for the transport of water.

TABLE 11-1 • Mechanisms for Transporting Ions and Small Molecules Across Cellular Membranes

Property	Simple Diffusion	Facilitated Transport	Cotransport[i]	Active Transport
Requires specific protein	–	+	+	+
Solute transported against its gradient	–	–	+	+
Coupled to ATP hydrolysis	–	–	–	+
Driven by	–	–	+	–

movement of a cotransported ion down its gradient				
Examples of molecules transported	O_2, CO_2, steroid hormones, many drugs	Glucose and amino acids (uniporters); ions and water (channels)	Glucose and amino acids (symporters); various ions and sucrose (antiporters)	Ions, small hydrophilic molecules, lipids (ATP-powered pumps)

[i]Also called *secondary active transport.*

KEY CONCEPTS OF SECTION 11.1

Overview of Transmembrane Transport

- Cellular membranes regulate the traffic of molecules and ions into and out of cells and their organelles. The rate of simple diffusion of a substance across a membrane is proportional to its concentration gradient and hydrophobicity.
- With the exception of gases (e.g., O_2 and CO_2) and small, uncharged, water-soluble molecules, most molecules cannot diffuse across a pure phospholipid bilayer at rates sufficient to meet cellular needs.
- Membrane transport proteins provide a hydrophilic passageway for molecules and ions to travel through a membrane that has a hydrophobic interior.
- Three classes of transmembrane proteins mediate transport of ions, sugars, amino acids, and other metabolites across cellular membranes: channels, transporters, and ATP-powered pumps (see Figure 11-2).
- Channels form a hydrophilic tube through which water or ions move *down* a concentration gradient, a process known as *facilitated transport*.
- Some channels are ungated, or open all the time, while other channels are gated, and open only in response to specific electrical or chemical signals (or to light in the case of the light-activated channels used in optogenetics, described in Chapters 4 and 23).

- Transporters fall into three groups. Uniporters transport a molecule down its concentration gradient (facilitated transport); symporters and antiporters couple movement of a substrate against its concentration gradient to the movement of a second substrate down its concentration gradient, a process known as *secondary active transport* or *cotransport* (see Table 11-1).
- ATP-powered pumps couple the movement of a substrate *against* its concentration gradient to ATP hydrolysis, a process known as *active transport*.
- Conformational changes are essential to the function of all membrane transport proteins; speed of transport depends on the number of substrate molecules or ions that can pass through a protein at once.
- The alternating access model describes the conformational changes that many transporters and pumps use to transport molecules across the membrane.

11.2 Facilitated Transport of Glucose and Water

Most animal cells use glucose as a substrate for ATP production; they usually employ a glucose uniporter to take up glucose from the blood or other extracellular fluid. Many cells use channel-like membrane transport proteins called *aquaporins* to increase the rate of water movement across their plasma membranes. Here we discuss the structure and function of these and other facilitated transporters.

Uniport Transport Is Faster and More Specific than Simple Diffusion

The transport of a single type of molecule, such as glucose or another small hydrophilic molecule, down a concentration gradient across a cellular membrane through a transport protein is known as **uniport**. Several features distinguish uniport from simple diffusion:

1. The rate of substrate movement by uniporters is far higher than simple diffusion through a pure phospholipid bilayer.
2. Because the transported molecule never enters the hydrophobic core of the phospholipid bilayer, its solubility in the lipid membrane is irrelevant.

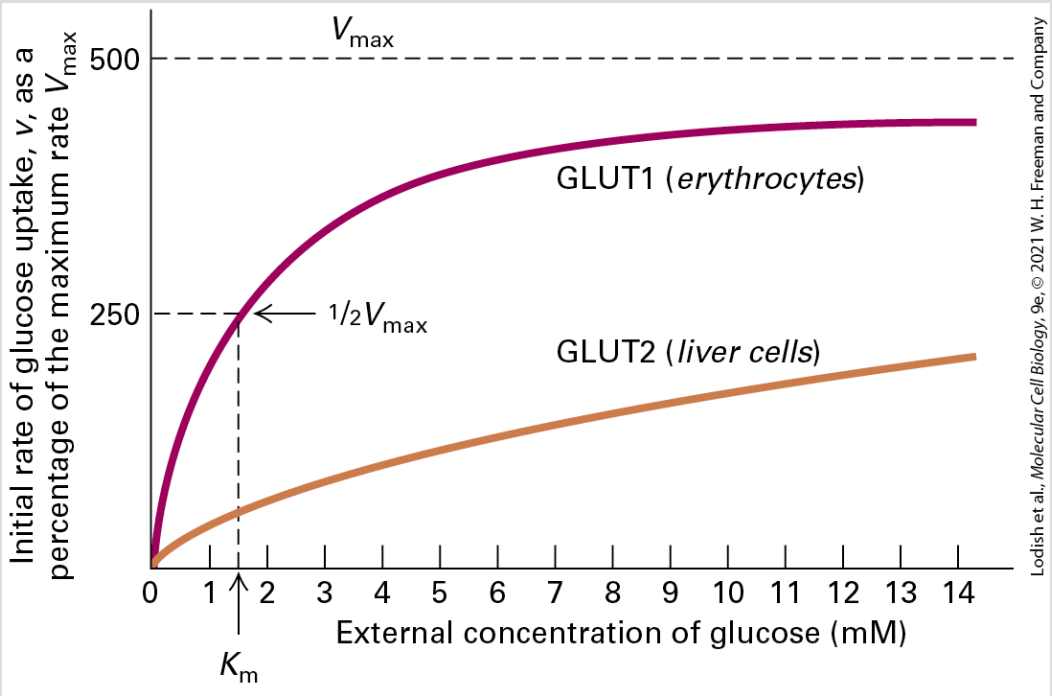
3. Transport occurs via a limited number of uniporter molecules. Consequently, there is a maximum transport rate, V_{max}, which depends on the number of uniporters in the membrane. V_{max} is achieved when the concentration gradient across the membrane is very large and each uniporter is working at its maximal rate.
4. Transport is reversible, and the direction of transport changes if the direction of the concentration gradient changes.
5. Transport is specific. Each uniporter transports only a single type of molecule or a single group of closely related molecules. A measure of the affinity of a transporter for its substrate is the Michaelis constant, K_m, which is the concentration of substrate at which transport is half V_{max}.

These properties also apply to transport mediated by the other classes of proteins depicted in Figure 11-3.

One of the best understood uniporters is the glucose transporter called *GLUT1*, found in the plasma membrane of most mammalian cells. GLUT1 is especially abundant in the erythrocyte (red blood cell) plasma membrane. Because erythrocytes have a single membrane and no nucleus or other internal organelles (see Figure 10-7), it is relatively simple to isolate and purify the transport proteins in their plasma membranes. As a result, the properties of GLUT1 and many other transport proteins from mature erythrocytes have been extensively studied. In addition, the three-dimensional structure of human GLUT1 was solved in 2014, providing further

insights into the molecular details of GLUT1 function (see chapter opening image and Figure 11-6).

Figure 11-5 shows that glucose uptake by erythrocytes and liver cells exhibits kinetics similar to those of a simple enzyme-catalyzed reaction involving a single substrate. The kinetics of transport reactions mediated by other types of proteins are more complicated than those for uniporters. Nonetheless, all protein-assisted transport reactions occur faster than simple diffusion across the bilayer, are substrate-specific, and exhibit a maximal rate (V_{max}).



EXPERIMENTAL FIGURE 11-5 Cellular uptake of glucose mediated by GLUT proteins exhibits simple enzyme kinetics. The initial rate of glucose uptake, *v* (measured as micromoles per milliliter of cells per hour), in the first few seconds is plotted as a percentage of the maximum rate, V_{max}, against increasing glucose concentration in the extracellular medium. In this

experiment, the initial concentration of glucose in the cells is always zero. Both GLUT1, expressed by erythrocytes, and GLUT2, expressed by liver cells, catalyze glucose uptake. Like enzyme-catalyzed reactions, GLUT-facilitated uptake of glucose exhibits a maximum rate (V_{max}). K_{m} is the concentration at which the rate of glucose uptake is half maximal. GLUT2, with a K_{m} of about 20 mM (not shown), has a much lower affinity for glucose than GLUT1, with a K_{m} of about 1.5 mM.

The Low K_{m} of the GLUT1 Uniporter Enables It to Transport Glucose into Most Mammalian Cells

Like other uniporters, GLUT1 alternates between two conformational states: in one, a glucose-binding site faces the outside of the cell; in the other, a glucose-binding site faces the cytosol. The latter conformation has been solved at high resolution, as shown Figure 11-6a. Since the glucose concentration is usually higher in the extracellular medium (blood, in the case of erythrocytes) than in the cell, the GLUT1 uniporter generally catalyzes the net import of glucose from the extracellular medium into the cell. Figure 11-6b depicts the sequence of events during the unidirectional transport of glucose from the cell exterior inward to the cytosol as described by the alternating access model; note the conformational changes in several of the membrane-spanning α helices during this process. GLUT1 can also catalyze the net export of glucose from the cytosol to the extracellular medium when the

glucose concentration is higher inside the cell than outside. In the transition between the inward- and outward-facing states, the transporter is in an occluded conformation, in which the glucose-binding site is inaccessible to either the cytoplasm or the extracellular space.

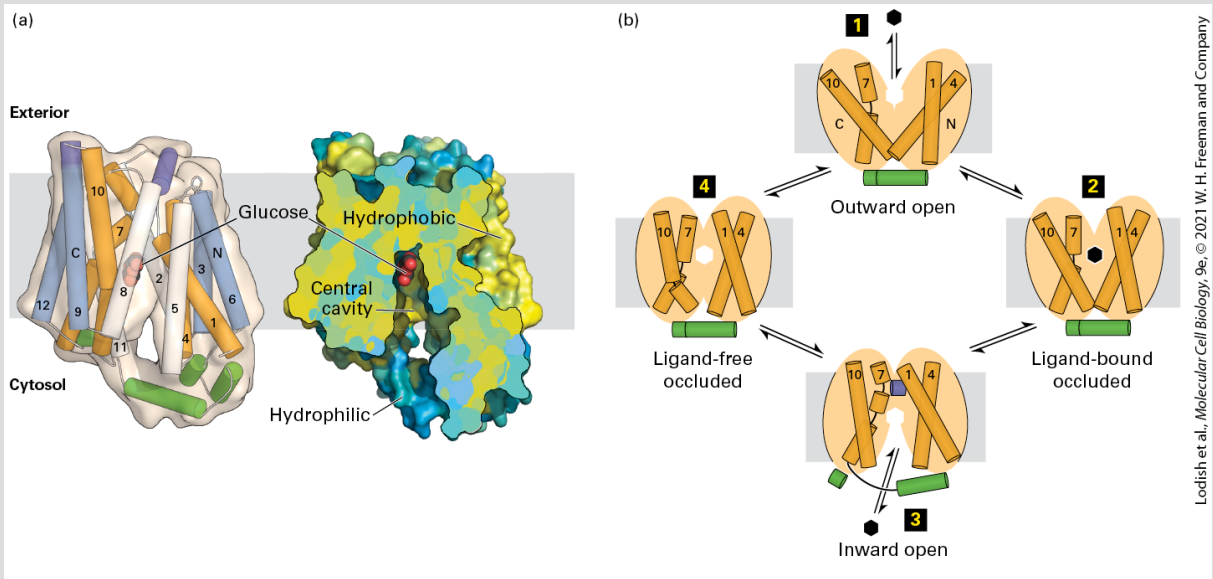


FIGURE 11-6 The human GLUT1 uniporter transports glucose across cellular membranes. (a) Structural model (side view) of the full-length human GLUT1 protein in an inward-open conformation. The transporter consists of 12 transmembrane α-helical segments, which are organized into amino-terminal and carboxy-terminal domains, each of which consists of a pair of three transmembrane α helices, colored in orange, white, and blue. The amino-terminal and carboxy-terminal domains are connected by intracellular and extracellular α helices, which are colored green and purple, respectively. A section of a cut-open view of the surface electrostatic potential highlights the central cavity that transports glucose (red) across the membrane. The colors represent the hydrophobicity of the amino acids, with hydrophobic in yellow and hydrophilic in blue. (b) A working model for GLUT1. In this alternating access model, the outward-open conformation of GLUT1 binds glucose (step 1) and moves to a ligand-bound occluded conformation (step 2) before changing to its inward-open conformation (step 3) when it delivers glucose to the cytoplasm, then moves through a ligand-free occluded conformation (step 4) before beginning another round of glucose transport from outside to inside the cell. If the concentration of glucose is higher

inside the cell than outside, the cycle will work in reverse (step 4 → step 1), resulting in net movement of glucose out of the cell. The orange helices correspond to helices 1, 4, 7, and 10 in (a). The actual conformational changes are probably smaller than those depicted here.

[Part (a) data from D. Deng et al., 2014, *Nature* **510**:121–125, PDB ID 4pyp.]

The kinetics of the unidirectional transport of glucose from the outside of a cell inward via GLUT1 can be described by the same type of equation used to describe a simple enzyme-catalyzed chemical reaction. For simplicity, let's assume that the substrate (glucose), S, is present initially only on the outside of the cell; this can be achieved experimentally by first incubating cells in a medium lacking glucose so that their internal stores are depleted. In this case, we can write

$$\mathrm{S_{out}} + \mathrm{GLUT1} \overset{K_\mathrm{m}}{\rightleftharpoons} \mathrm{S_{out}} - \mathrm{GLUT1} \overset{V_\mathrm{max}}{\rightleftharpoons} \mathrm{S_{in}} + \mathrm{GLUT1}$$

where $\mathrm{S_{out}} - \mathrm{GLUT1}$ represents GLUT1 in the outward-facing conformation with a bound glucose. This equation is similar to the one describing the path of a simple enzyme-catalyzed reaction in which the protein binds a single substrate and then transforms it into a different molecule. Here, however, no chemical modification of the GLUT1-bound glucose molecule occurs; rather, it is moved across a cellular membrane. Nonetheless, the kinetics of this transport reaction are similar to those of simple enzyme-catalyzed reactions, and we can use the same derivation as that of the

Michaelis–Menten equation in Chapter 3 to derive the following expression for v_0, the initial transport rate for S into the cell catalyzed by GLUT1:

$$v_0 = \frac{V_{\max}}{1 + \frac{K_{\mathrm{m}}}{C}} \tag{11-1}$$

where C is the concentration of $\mathrm{S}_{\mathrm{out}}$ (initially, the concentration of $\mathrm{S}_{\mathrm{in}} = 0$). $V_{\max}$, the rate of transport when all molecules of GLUT1 contain a bound S, occurs at an infinitely high $\mathrm{S}_{\mathrm{out}}$ concentration. The lower the value of K_{m}, the more tightly the substrate binds to the transporter. Equation 11-1 describes the curve for glucose uptake by erythrocytes shown in Figure 11-4 as well as similar curves for other uniporters.

For GLUT1 in the human erythrocyte membrane, the K_{m} for glucose transport is 1.5 mM. Thus when the extracellular glucose concentration is 1.5 mM, roughly half the GLUT1 transporters with outward-facing binding sites will have a bound glucose, and transport will occur at 50 percent of the maximal rate. Blood glucose is normally 5 mM, so the erythrocyte glucose transporter is usually functioning at 77 percent of its maximal rate, as can be seen from Equation 11-1. The GLUT1 transporter (or the very similar GLUT3 glucose transporter) is expressed by all body cells that need to take up glucose from the blood continuously at high rates. The rate of

glucose uptake by such cells remains high regardless of small changes in the concentration of blood glucose because the blood concentration remains much higher than the K_m and the intracellular glucose concentration is kept low as a result of active metabolism. Of note, GLUT1 expression is elevated in many malignant cancers, consistent with the need for glucose as a source of energy for rapidly dividing cancer cells.

In addition to transporting glucose, GLUT1 transports the isomeric sugars D-mannose and D-galactose, which differ from D-glucose in their configuration at only one carbon atom, at measurable rates. However, the K_m for glucose (1.5 mM) is much lower than it is for D-mannose (20 mM) or D-galactose (30 mM). Thus GLUT1 is quite specific, having a much higher affinity (indicated by a lower K_m) for its normal substrate D-glucose than for other substrates.

GLUT1 accounts for 2 percent of the protein in the plasma membrane of erythrocytes. After glucose is transported into the erythrocyte, it is rapidly phosphorylated, forming glucose-6-phosphate, which cannot leave the cell. Because this reaction, the first step in the metabolism of glucose (see Figure 12-3), is rapid and occurs at a constant rate, the intracellular concentration of glucose is kept low even when glucose is imported from the extracellular environment. Consequently, the concentration gradient of glucose (outside greater than inside the cell) is kept sufficiently high to

support continuous, rapid import of additional glucose molecules and provide sufficient glucose for cellular metabolism.

The Human Genome Encodes a Family of Sugar-Transporting GLUT Proteins

The human genome encodes at least 14 highly homologous *GLUT proteins*, GLUT1–GLUT14. All are thought to contain 12 membrane-spanning α helices, suggesting that they evolved from a single ancestral transport protein. In the human GLUT1 protein, the transmembrane α helices are predominantly hydrophobic; several helices, however, bear amino acid residues (e.g., serine, threonine, asparagine, and glutamine) whose side chains can form hydrogen bonds with the hydroxyl groups on glucose. These residues are thought to form the inward-facing and outward-facing glucose-binding sites in the interior of the protein (see Figure 11-5).

The structures of all GLUT isoforms are thought to be quite similar, and all of them transport sugars. Nonetheless, the isoforms differ in their expression in various cell types, in the regulation of their numbers on cell surfaces, and in their functional properties. By using isoforms with different properties, different body cells are able to regulate glucose metabolism differently and at the same time allow a constant concentration of glucose in the blood to be maintained. For instance, GLUT3 is found in neuronal cells of the

brain. Neurons depend on a constant influx of glucose for metabolism, and the low K_m of GLUT3 for glucose (1.5 mM), like that of GLUT1, ensures that these cells incorporate glucose from brain extracellular fluids at a high and constant rate.

GLUT2, expressed in liver cells and in the insulin-secreting β islet cells of the pancreas, has a K_m of ~20 mM, about 13 times higher than the K_m of GLUT1. As a result, when blood glucose rises after a meal from its basal level of 5 mM to 10 mM or so, the rate of glucose influx will almost double in GLUT2-expressing cells, whereas it will increase only slightly in GLUT1-expressing cells (see Figure 11-5). In the liver, the excess glucose brought into the cell is stored as the polymer glycogen. In β islet cells, the rise in glucose triggers secretion of the hormone insulin (see Figure 21-1b), which in turn lowers blood glucose by increasing glucose uptake and metabolism in muscle (see Figure 15-32) and by inhibiting glucose production in the liver. Indeed, inactivation of GLUT2 in pancreatic β islet cells prevents insulin secretion even when levels of glucose are high and disrupts the regulated expression of glucose-sensitive genes in liver cells (hepatocytes).

Another GLUT isoform, GLUT4, is expressed only in fat and muscle cells. These cells respond to insulin by increasing their uptake of glucose, thereby removing glucose from the blood. In the absence of insulin, GLUT4 resides in intracellular membranes, not the plasma membrane, and is unable to facilitate glucose uptake from the

extracellular fluid. By a process detailed in Figure 21-2, insulin causes these GLUT4-rich internal membranes to fuse with the plasma membrane, increasing the number of GLUT4 molecules present on the cell surface and thus the rate of glucose uptake. This is one principal mechanism by which insulin lowers blood glucose; defects in the movement of GLUT4 to the plasma membrane are one of the causes of adult-onset, or type II, diabetes, a disease marked by continuously high blood glucose.

GLUT5 is the only GLUT protein with a high specificity (preference) for fructose. Its principal site of expression is the apical membrane of intestinal epithelial cells, where it transports dietary fructose from the intestinal lumen to the inside of the cells.

Transport Proteins Can Be Studied Using Artificial Membranes and Recombinant Cells

There are a variety of approaches to studying the intrinsic properties of transport proteins, such as the V_{max} and K_{m} parameters and the key residues responsible for binding. Most cellular membranes contain many different types of transport proteins but a relatively low concentration of any particular one, which makes functional studies of a single protein difficult. To facilitate such studies, researchers use two approaches to enrich a

transport protein of interest so that it predominates in the membrane: purification and insertion into artificial membranes and overexpression in recombinant cells.

In the first approach, a specific transport protein is extracted from its membrane with detergent and purified. Although transport proteins can be isolated from membranes and purified, their functional properties (i.e., their role in the movement of substrates across membranes) can be studied only when they are associated with a membrane. Thus the purified proteins are usually reincorporated into membranes composed of a pure phospholipid bilayer, such as liposomes (see Figure 10-3). The rate of substrate transport across such membranes can be readily measured.

Alternatively, the gene encoding a specific transport protein can be expressed at high levels in a cell type that normally does not express it. The difference between the transport rate of a substance by the transfected cells and by control nontransfected cells will be due to the expressed transport protein. In these systems, the functional properties of the various membrane proteins can be examined without ambiguity caused, for instance, by partial protein denaturation during isolation and purification procedures. As an example, overexpressing GLUT1 in lines of cultured fibroblasts increases their rate of uptake of glucose severalfold, and expression of mutant GLUT1 proteins with specific amino acid alterations can identify residues important for substrate binding.

Osmotic Pressure Causes Water to Move Across Membranes

The life of all organisms depends on the movement of water into and out of cells. The **aquaporins** are a family of membrane channel proteins that allow water and a few other small uncharged molecules, such as glycerol, to cross cellular membranes efficiently. Before discussing these transport proteins, we need to review osmosis, the force that powers the movement of water across membranes.

Water spontaneously moves downhill across a semipermeable membrane from a solution of lower solute concentration (relatively high water concentration) to one of higher solute concentration (relatively low water concentration), a process termed **osmosis**, or *osmotic flow*. In effect, osmosis is equivalent to diffusion of water across a semipermeable membrane. *Osmotic pressure* is defined as the hydrostatic pressure required to stop the net flow of water across a membrane separating solutions of different water concentrations (Figure 11-7). In other words, osmotic pressure balances the entropy-driven thermodynamic force of the water concentration gradient. In this context, a "membrane" may be a layer of cells or a plasma membrane that is permeable to water but not to the solutes it contains. The osmotic pressure is directly proportional to the difference in the concentrations of the total numbers of solute molecules on the two sides of the membrane. For example, a 0.5 M

NaCl solution is actually 0.5 M Na^+ ions and 0.5 M Cl^- ions and has a similar osmotic pressure as a 1 M solution of glucose or sucrose.

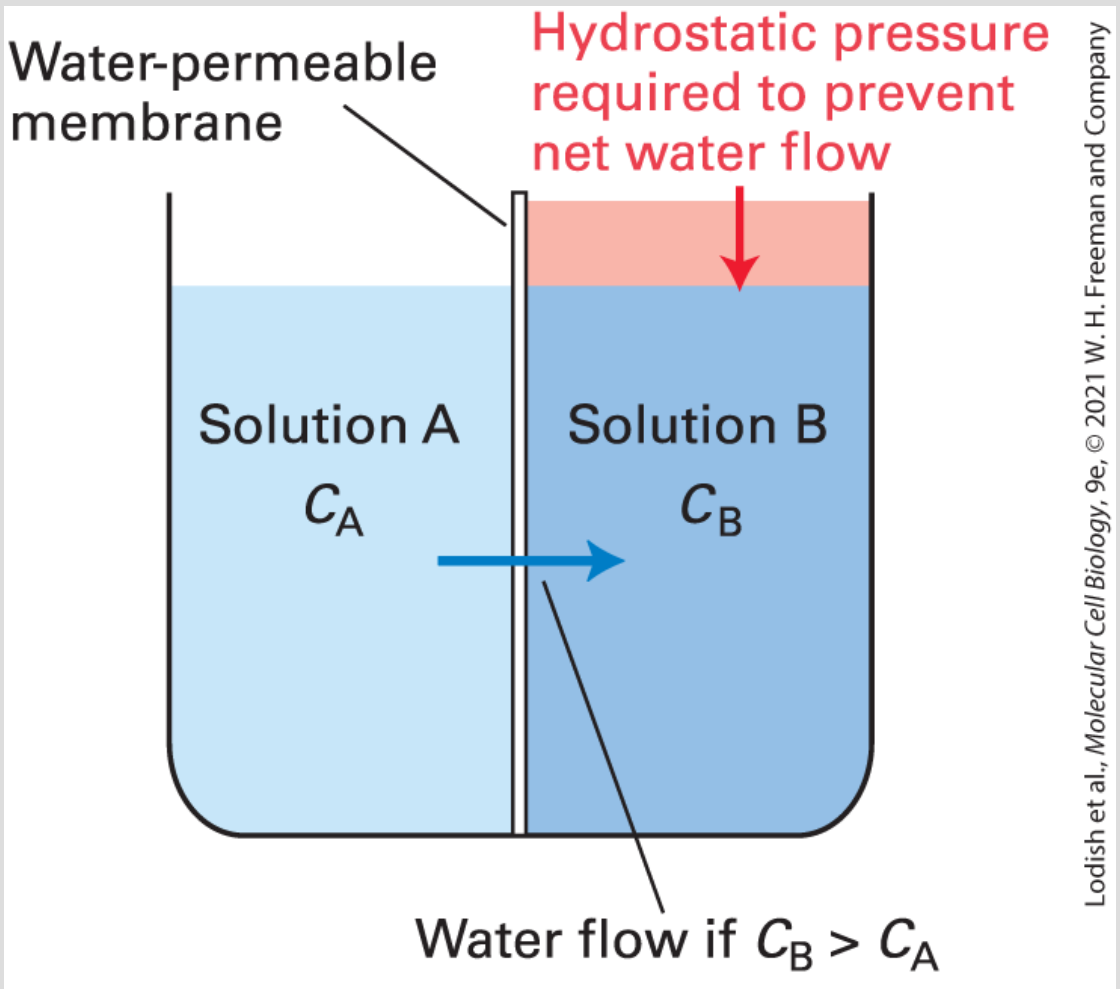


FIGURE 11-7 Osmotic pressure. Solutions A and B are separated by a membrane that is permeable to water but impermeable to all solutes. If C_B (the total concentration of solutes in solution B) is greater than C_A, water will tend to flow across the membrane from solution A to solution B. Osmotic pressure π is the hydrostatic pressure that would have to be applied to solution B to prevent this water flow. From the van't Hoff equation, osmotic pressure is given by $\pi = RT(C_B - C_A)$, where *R* is the gas constant and *T* is the absolute temperature.

The movement of water across the plasma membrane determines the volume of an individual cell, which must be regulated to avoid damage to the cell. Small changes in extracellular osmotic conditions cause most animal cells to swell or shrink rapidly. When placed in a *hypotonic* solution (i.e., one in which the concentration of non-membrane-penetrating solutes is lower than in the cytosol), animal cells swell owing to the osmotic flow of water inward. Conversely, when placed in a *hypertonic* solution (i.e., one in which the concentration of non-membrane-penetrating solutes is higher than in the cytosol), animal cells shrink as cytosolic water leaves the cell by osmotic flow. Consequently, cultured animal cells must be maintained in an *isotonic* medium, which has a solute concentration, and thus osmotic strength, similar to that of the cell cytosol.

In vascular plants, water and minerals are absorbed from the soil by the roots and move up the plant through conducting tubes (the xylem); water loss from the plant, mainly by evaporation from the leaves, drives this movement of water. Unlike animal cells, plant, algal, fungal, and bacterial cells are surrounded by a rigid cell wall, which resists the expansion of the volume of the cell when the intracellular osmotic pressure increases. Without such a wall, animal cells expand when internal osmotic pressure increases; if that pressure rises too much, the cells burst like overinflated balloons. Because of the cell wall, the osmotic influx of water that occurs when plant cells are placed in a hypotonic solution (even pure water) leads to an increase in intracellular pressure, but not in cell volume.

In plant cells, the concentration of solutes (e.g., sugars and salts) is usually higher in the vacuole (see Figure 1-13a) than in the cytosol, which in turn has a higher solute concentration than the extracellular space. The osmotic pressure generated by the entry of water into the cytosol and then into the vacuole, called *turgor pressure*, pushes the cytosol and the plasma membrane against the resistant cell wall. Plant cells can harness this pressure to help them stand upright and grow. Cells elongate during growth by means of a process induced by hormones: a defined region of the cell wall loosens and then water flows into the vacuole, increasing its size and thus the size of the cell. ■

Although most protozoans (like animal cells) do not have a rigid cell wall, many contain a *contractile vacuole* that permits them to avoid bursting from osmotic pressure. A contractile vacuole takes up water from the cytosol and, unlike a plant vacuole, periodically discharges its contents through fusion with the plasma membrane. Thus even though water continuously enters the protozoan cell by osmotic flow, the contractile vacuole prevents too much water from accumulating in the cell and swelling it to the bursting point.

Aquaporins Increase the Water Permeability of Cellular Membranes

The natural tendency of water to flow across cellular membranes as a result of osmotic pressure raises an obvious question: Why don't the cells of freshwater animals burst in water? Frogs, for example, lay their eggs in pond water (a hypotonic solution), but frog oocytes and eggs do not swell with water even though their internal salt (mainly KCl) concentration is comparable to that of other cells ($\sim$150 mM KCl). These observations were what first led investigators to suspect that the plasma membranes of most cell types, but not of frog oocytes, contain water-channel proteins that accelerate the osmotic flow of water. The finding that injection of the mRNA encoding the aquaporin channel protein into frog eggs caused them to burst in fresh water demonstrated that aquaporins are sufficient to increase the permeability of cells to water.

Aquaporin channels are selective for water molecules, which they transport in either direction depending on the osmotic gradient. A tetramer of identical 28-kDa subunits (Figure 11-8a), each subunit contains six membrane-spanning α helices that form a central pore through which water can move (Figure 11-8b, c). The ∼2-nm-long channel, or pore, at the center of each monomer is only 0.28 nm in diameter — only slightly larger than the diameter of a water molecule. The molecular sieving properties of the channel are determined by several conserved, hydrophilic amino acid residues whose side-chain and carbonyl groups extend into the middle of the channel and by a relatively hydrophobic wall that lines one side of the channel. Several water molecules can move simultaneously through the channel, each molecule sequentially forming specific hydrogen bonds with the channel-lining amino acids and displacing another water molecule downstream. Aquaporins do not undergo conformational changes during water transport, so they transport water orders of magnitude faster than GLUT1 transports glucose. The formation of hydrogen bonds between the oxygen atom of water and the amino groups of two amino acid side chains ensures that only uncharged water (i.e., H_2O, but not H_3O^+) passes through the channel; the orientations of the water molecules in the channel prevent protons from jumping from one to the next and thus prevent the net movement of protons through the channel. As a consequence, ionic gradients are maintained across membranes even when water is flowing across them through aquaporins.

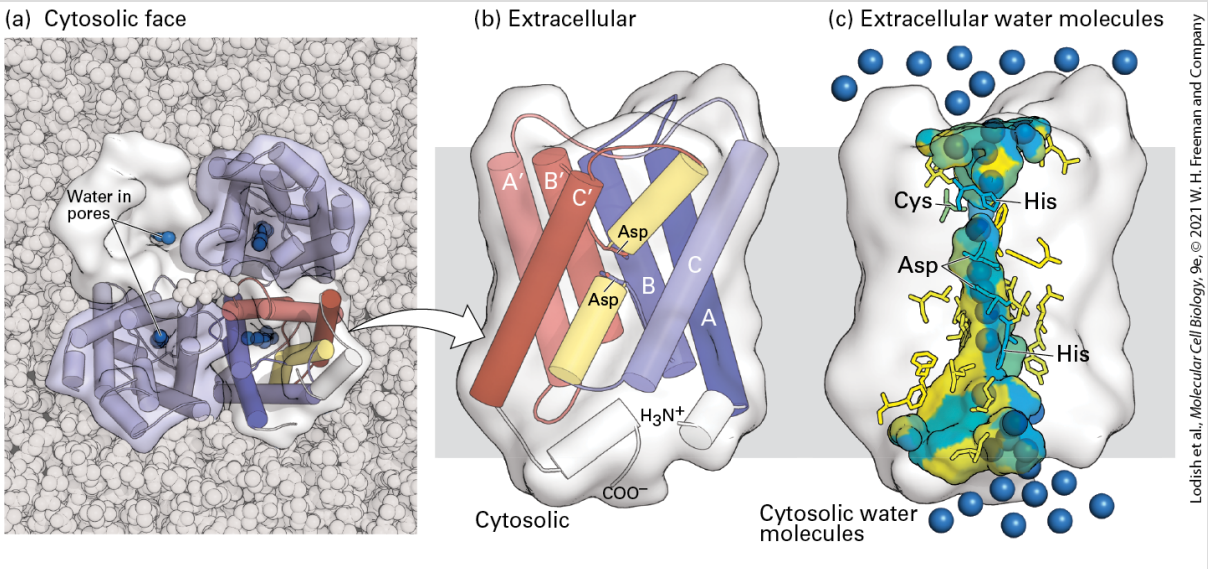


FIGURE 11-8 Structure of an aquaporin. (a) Structural model of the tetrameric protein comprising four identical subunits. Each subunit forms a water channel, as seen in this view looking down on the protein from the exoplasmic side. One of the monomers is shown as a water-accessible surface model, in which the pore entrance can be seen. (b) Schematic diagram of the topology of a single aquaporin subunit in relation to the membrane. Three pairs of homologous transmembrane α helices (A and A', B and B', and C and C') are oriented in the opposite direction with respect to the membrane and are connected by two hydrophilic loops containing short non-membrane-spanning helices and conserved asparagine (Asn) residues. The loops bend into the cavity formed by the six transmembrane helices, meeting in the middle to form part of the water-selective gate. (c) Side view of the pore in a single aquaporin subunit, in which several water molecules (blue spheres) are seen within the 2-nm-long water-selective gate that separates the water-filled cytosolic and extracellular vestibules. The gate contains highly conserved hydrophilic amino acid residues whose side chains form hydrogen bonds with transported water molecules. The amino acids lining the pore are colored from hydrophilic (blue) to hydrophobic (yellow). The arrangement of these hydrogen bonds and the narrow pore diameter of 0.28 nm prevent passage of protons (i.e., H_3O^+) or other ions.

[Data from H. Sui et al., 2001, *Nature* **414**:872, PDB ID 1j4n.]

Mammals express a family of aquaporins; 14 such genes are known in humans. Aquaporin 1 is expressed in abundance in

erythrocytes, and the homologous aquaporin 2 is found in the kidney epithelial cells that resorb water from the urine, thus controlling the amount of water in the body. The activity of aquaporin 2 is regulated by vasopressin, also called *antidiuretic hormone*, in a manner that resembles the regulation of GLUT4 activity in fat and muscle. When the cells are in their resting state and water is being excreted to form urine, aquaporin 2 is sequestered in intracellular vesicle membranes and so is unable to import water into the cell. When the polypeptide hormone vasopressin binds to the cell-surface vasopressin receptor, it activates a signaling pathway using cAMP as the intracellular signal (detailed in Chapter 15) that causes these aquaporin 2–containing vesicles to fuse with the plasma membrane, increasing the rate of water uptake and return to the circulation. Inactivating mutations in either the vasopressin receptor or the aquaporin 2 gene cause *diabetes insipidus*, a disease marked by excretion of large volumes of dilute urine. This finding demonstrates that the level of aquaporin 2 is rate limiting for water resorption from urine being formed by the kidney. ■

Other members of the aquaporin family transport hydroxyl-containing molecules such as glycerol rather than water. Human aquaporin 3, for instance, transports glycerol and is similar in amino acid sequence and structure to the *Escherichia coli* glycerol transport protein GlpF.

KEY CONCEPTS OF SECTION 11.2

Facilitated Transport of Glucose and Water

- Protein-catalyzed transport of biological solutes across a membrane occurs much faster than simple diffusion, exhibits $V_{\max}$ when the limited number of transporter molecules are saturated with substrate, and is highly specific for substrate (see Figure 11-4).
- Uniport proteins, such as the glucose transporters (GLUTs), are thought to shuttle between two conformational states, one in which the substrate-binding site faces outward and one in which the binding site faces inward (see Figure 11-5).
- All members of the GLUT protein family transport sugars and have similar structures. Differences in their K_{m} values, expression in different cell types, and substrate specificities are important for proper sugar metabolism in the body.
- Two common experimental systems for studying the functions of transport proteins are liposomes containing a purified transport protein and cells transfected with the gene encoding a particular transport protein.
- Most cellular membranes are semipermeable, more permeable to water than to ions or most other solutes. Water moves by osmosis across membranes from a solution of lower solute concentration to one of higher solute concentration.
- The rigid cell wall surrounding plant cells prevents their swelling and leads to generation of turgor pressure in response to the osmotic influx of water.
- Aquaporins are water-channel proteins that specifically increase the permeability of cellular membranes to water (see Figure 11-8).
- Aquaporin 2 in the plasma membrane of certain kidney cells is essential for resorption of water from urine being formed; the absence of aquaporin 2 leads to the medical condition diabetes insipidus.

11.3 ATP-Powered Pumps and the Intracellular Ionic Environment

In the previous sections, we focused on transport proteins that move molecules down their concentration gradients (facilitated transport). Here we focus our attention on the class of proteins — the ATP-powered pumps — that use the energy released by hydrolysis of the terminal phosphoanhydride bond of ATP to transport ions and various small molecules across membranes *against* their concentration gradients. All ATP-powered pumps are transmembrane proteins with one or more binding sites for ATP located on subunits or segments of the protein that face the cytosol. These proteins are ATPases, but they normally do not hydrolyze ATP into ADP and P_i unless ions or other molecules are simultaneously transported. Because of this tight *coupling* between ATP hydrolysis and transport, the energy stored in the phosphoanhydride bond is not dissipated as heat, but rather is used to move ions or other molecules uphill against an electrochemical gradient.

There Are Four Main Classes of ATP-Powered Pumps

The general structures of the four classes of ATP-powered pumps are depicted in Figure 11-9, with specific examples in each class listed below the figure. Note that the members of three of the classes

(P, F, and V) transport only ions, as do some members of the fourth class, the ABC superfamily. Most members of the ABC superfamily, however, transport small molecules such as amino acids, sugars, peptides, lipids, and many types of drugs.

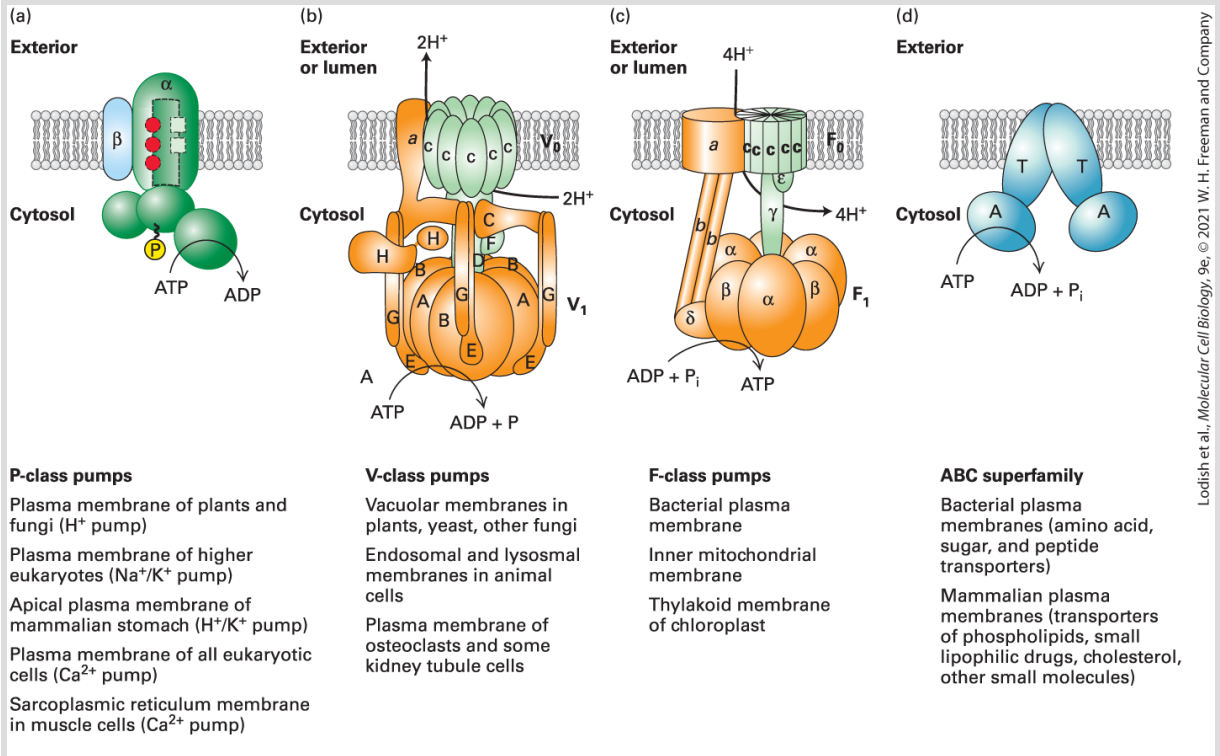


FIGURE 11-9 The four classes of ATP-powered transport proteins. The locations of specific examples are indicated below each class. (a) P-class pumps are composed of two catalytic α subunits, which become phosphorylated as part of the transport cycle. Two β subunits, present in some of these pumps, may regulate transport. Only one α and one β subunit are depicted here, from Na^+/K^+ ATPase, with the three red circles representing Na^+ ion binding sites and the two light green squares representing K^+ ion binding sites. (b, c) V-class and F-class pumps are not phosphorylated, and almost all transport only protons. *V-class pumps* couple ATP hydrolysis to transport protons against a concentration gradient, whereas *F-class pumps* normally operate in the reverse direction and use the energy in a proton concentration or voltage gradient to synthesize ATP. (d) All members of the large ABC superfamily of proteins contain two transmembrane (T) domains and two cytosolic ATP-binding (A) domains, which couple ATP hydrolysis to solute movement. These core domains

are present as separate subunits in some ABC proteins (as depicted here) but are fused into a single polypeptide in other ABC proteins.

[Information from S. Sharma et al., 2019, *J. Biol. Chem.* **294**:6439; C. Toyoshima et al., 2000, *Nature* **405**:647; D. McIntosh, 2000, *Nat. Struct. Biol.* **7**:532; and H. Okamoto and M. Futai, 2013, *Encyclopedia of Biophysics*, https://doi.org/10.1007/978-3-642-16712-6_203.]

All **P-class pumps** possess two identical catalytic α subunits, each of which contains an ATP-binding site. Most also have two smaller β subunits, which usually have regulatory functions. During transport, at least one of the α subunits becomes phosphorylated (hence the name "P" class), and the transported ions move through the phosphorylated subunit. The amino acid sequences around the phosphorylated residues are homologous in different pumps. This class includes the Na^+/K^+ ATPase in the plasma membrane, which generates the low cytosolic Na^+ and high cytosolic K^+ concentrations typical of animal cells (see Figure 11-3). Certain Ca^{2+} ATPases pump Ca^{2+} ions out of the cytosol into the external medium; others pump Ca^{2+} from the cytosol into the endoplasmic reticulum or into the specialized ER called the *sarcoplasmic reticulum (SR)* that is found in muscle cells. Another member of the P class, the gastric proton pump, is found in acid-secreting cells of the mammalian stomach; the gastric proton pump transports protons (H^+ ions) out of and K^+ ions into the cell. The recently solved crystal structure of this pump yielded important insights into the binding sites of drugs that are used to treat diseases of damage caused by gastric acid such as gastric ulcers and reflux esophagitis.

The structures of *V-class* and *F-class pumps* are similar to one another but are unrelated to, and more complicated than, those of P-class pumps. V- and F-class pumps contain several different transmembrane and cytosolic subunits. Virtually all known V and F pumps transport only protons and do so in a process that does not require phosphorylation. V-class pumps generally function to generate the low pH of plant vacuoles and of lysosomes and other acidic vesicles in animal cells by pumping protons from the cytosolic to the exoplasmic face of the membrane against a proton electrochemical gradient. In contrast, the proton pumps that generate and maintain the electric potential across the plasma membrane in plant, fungal, and many bacterial cells belong to the P-class pumps.

F-class pumps are found in bacterial plasma membranes and in mitochondria and chloroplasts. In contrast to V-class pumps, they generally function as reverse proton pumps, in which protons move from the exoplasmic to the cytosolic face of the membrane *down* the proton electrochemical gradient, in the energetically favored direction. The energy released is used to power the energetically unfavorable synthesis of ATP from ADP and P_i. Because of their importance in ATP synthesis in chloroplasts and mitochondria, F-class proton pumps, commonly called *ATP synthases,* are treated separately in Chapter 12 (see Figure 12-27).

The final class of ATP-powered pumps is a large family with multiple members that are more diverse in function than those of the other classes. Referred to as the **ABC** (*A*TP-*b*inding *c*assette) **superfamily**, this class includes several hundred different transport proteins found in organisms ranging from bacteria to humans. As detailed below, some of these transport proteins were first identified as multidrug-resistance proteins that, when overexpressed in cancer cells, export anticancer drugs and render tumors resistant to their action. Each ABC protein is specific for a single substrate or group of related substrates, which may be ions, sugars, amino acids, phospholipids, cholesterol, peptides, polysaccharides, or even proteins. All ABC membrane transport proteins share a structural organization consisting of four core domains: two transmembrane (T) domains, which form the passageway through which transported molecules cross the membrane, and two cytosolic ATP-binding (A) domains. In some ABC proteins, mostly those in bacteria, the core domains are present as four separate polypeptides; in others, the core domains are fused into one or two multidomain polypeptides. ATP binding and hydrolysis drives the transport process in most ABC membrane transporters. However, the energy from ATP binding and hydrolysis can also be used to regulate the opening and closing of a continuous channel, as described below for the cystic fibrosis transmembrane conductance regulator protein (CFTR).

ATP-Powered Ion Pumps Generate and Maintain Ionic Gradients Across Cellular Membranes

The specific ionic composition of the cytosol usually differs greatly from that of the surrounding extracellular fluid. In virtually all cells — including microbial, plant, and animal cells — the cytosolic pH is kept near 7.2 *regardless* of the extracellular pH. In the most extreme case, there is a 1-million-fold difference in H^+ concentration between the cytosol of the epithelial cells lining the stomach and the stomach contents after a meal. Furthermore, the cytosolic concentration of K^+ is much higher than that of Na^+. In both invertebrates and vertebrates, the concentration of K^+ is 20–40 times higher in the cytosol than in the blood, while the concentration of Na^+ is 8–12 times lower in the cytosol than in the blood (Table 11-2).

TABLE 11-2 • Typical Intracellular and Extracellular Ion Concentrations

Ion	Cell (mM)	Seawater (mM)
Squid Giant Axon (marine invertebrate)[i]		
K^+	400	10
Na^+	50	440

Cl^-	40–150	560
Ca^{2+}	0.0003	10
X^-[ii]	300–400	5–10
Mammalian Cell (vertebrate)	**Cell (mM)**	**Blood (mM)**
K^+	139	4
Na^+	12	145
Cl^-	4	116
HCO_3^-	12	29
X^-	138	9
Mg^{2+}	0.8	1.5
Ca^{2+}	<0.0002	1.8

[i]The large nerve axon of the squid has been widely used in studies of the mechanism of conduction of electric impulses.

[ii]X^- represents proteins, which have a net negative charge at the neutral pH of blood and cells.

Some Ca^{2+} in the cytosol is bound to the negatively charged groups in ATP and in proteins and other molecules, but it is the concentration of unbound (or "free") Ca^{2+} that is critical to its functions in signaling pathways and muscle contraction. The concentration of free Ca^{2+} in the cytosol is generally less than 0.2 micromolar ($2 \times 10^{-7}\mu M$), a thousand or more times lower than

that in the blood. Plant cells and many microorganisms maintain similarly high cytosolic concentrations of K^+ and low concentrations of Ca^{2+} and Na^+, even if the cells are cultured in very dilute salt solutions.

The ion pumps discussed in this section are largely responsible for establishing and maintaining the typical ionic gradients observed across the plasma and intracellular membranes. In carrying out this task, cells expend considerable energy. For example, up to 25 percent of the ATP produced by nerve and kidney cells is used for ion transport, and human erythrocytes consume up to 50 percent of their available ATP for this purpose; in both cases, most of this ATP is used to power the Na^+/K^+ pump (see Figure 11-3). In neurons, the pump maintains Na^+ and K^+ gradients that are essential for conducting electrical signals rapidly and efficiently, as we detail in Chapter 23. Certain enzymes required for protein synthesis in all cells require a high K^+ concentration and are inhibited by high concentrations of Na^+; these enzymes would cease to function without the operation of the Na^+/K^+ pump. In cells treated with poisons that inhibit the production of ATP (e.g., 2,4-dinitrophenol in aerobic cells), the pumping stops, and the ion concentrations inside the cell gradually approach those of the exterior environment as ions spontaneously move through channels in the plasma membrane down their electrochemical gradients. Eventually the treated cells die, partly because protein synthesis requires a high concentration of K^+ ions and partly because, in the absence of a Na^+

gradient across the plasma membrane, a cell cannot import certain nutrients such as amino acids (see Figure 11-3). Studies on the effects of such poisons provided early evidence for the existence and significance of ion pumps.

Muscle Relaxation Depends on Ca^{2+} ATPases That Pump Ca^{2+} from the Cytosol into the Sarcoplasmic Reticulum

In skeletal muscle cells, Ca^{2+} ions are concentrated and stored in the sarcoplasmic reticulum (SR), a specialized type of endoplasmic reticulum (ER). The release (via ion channels) of stored Ca^{2+} ions from the SR lumen into the cytosol causes muscle contraction, as discussed in Chapter 17. A P-class Ca^{2+} ATPase located in the SR membrane pumps Ca^{2+} from the cytosol back into the lumen of the SR, thereby inducing muscle relaxation.

The activity of the muscle Ca^{2+} ATPase increases as the free Ca^{2+} concentration in the cytosol rises. In the cytosol of muscle cells, the free Ca^{2+} concentration ranges from 0.1 μM (resting cells) to more than 1 μM (contracting cells), whereas the *total* Ca^{2+} concentration in the SR lumen can be as high as 10 mM. The lumen of the SR contains two abundant proteins, calsequestrin and the so-called

high-affinity Ca^{2+}-binding protein, each of which binds multiple Ca^{2+} ions at high affinity. By binding much of the Ca^{2+} in the SR lumen, these proteins reduce the concentration of "free" Ca^{2+} ions inside the SR. This reduction, in turn, reduces the Ca^{2+} concentration gradient between the cytosol and the SR lumen and consequently reduces the energy needed to pump Ca^{2+} ions into the SR from the cytosol.

In skeletal muscle cells, the calcium pump in the SR membrane works in concert with a similar Ca^{2+} pump located in the plasma membrane to ensure that the cytosolic concentration of free Ca^{2+} in resting muscle remains at or below 0.1 μM. As the Ca^{2+} pump in the SR membrane moves Ca^{2+} into the SR lumen, the Ca^{2+} pump in the plasma membrane moves Ca^{2+} out of the cell. The activity of plasma-membrane Ca^{2+} ATPases is regulated by *calmodulin*, a cytosolic Ca^{2+}-binding protein (see Figure 3-34). A rise in cytosolic Ca^{2+} induces the binding of Ca^{2+} ions to calmodulin, which triggers activation of the Ca^{2+} ATPase. In this way, free Ca^{2+} ion concentration is coordinately regulated in the muscle cell cytoplasm and the SR lumen.

The Mechanism of Action of the Ca^{2+} Pump Is Known in Detail

Because the calcium pump constitutes more than 80 percent of the integral membrane protein in muscle SR membranes, it is easily purified and has been studied extensively. The three-dimensional structure of this protein has been determined in several conformational states, representing different steps in the pumping process. These structures have revealed much about its mechanism of action, which serves as a model for understanding many P-class ATPases.

The current model for the mechanism of action of the Ca^{2+} ATPase in the SR membrane proposes that the ATPase passes through multiple conformational states. The Ca^{2+} ATPase has two binding sites for Ca^{2+}, located in the center of the membrane-spanning domain. For simplicity, we group the multiple conformation states into E1 states, in which the two binding sites face the cytosol, and E2 states, in which these binding sites face the exoplasmic face of the membrane, pointing into the lumen of the SR. The coupling of ATP hydrolysis with ion pumping requires several conformational changes in the protein that must occur in a defined order, as shown in Figure 11-10. When the protein is in the E1 conformation, two Ca^{2+} ions bind to two high-affinity binding sites accessible from the cytosolic side; even though the cytosolic Ca^{2+} concentration is low (see Table 11-2), calcium ions still fill these sites.

Next an ATP binds to a site on the cytosolic surface (step **1**). The bound ATP is hydrolyzed to ADP in a reaction that requires Mg^{2+},

and the liberated phosphate is transferred to a specific aspartate residue in the protein, forming the high-energy acyl phosphate bond denoted by E1 $\sim$ P (step **2**). The protein then undergoes a conformational change that generates E2, in which the affinity of the two Ca^{2+}-binding sites is reduced (shown in detail in the next figure) and in which these sites are now accessible to the SR lumen (step **3**). The free energy of hydrolysis of the aspartyl-phosphate bond in E1 $\sim$ P is greater than that in E2–P, and this reduction in the free energy of the aspartyl-phosphate bond can be said to power the E1 $\rightarrow$ E2 conformational change.

The Ca^{2+} ions spontaneously dissociate from the binding sites to enter the SR lumen because even though the Ca^{2+} concentration there is higher than in the cytosol, it is lower than the K_d for Ca^{2+} binding in the low-affinity state (step **4**). Finally, the aspartyl-phosphate bond is hydrolyzed (step **5**). This dephosphorylation, coupled with subsequent binding of cytosolic Ca^{2+} to the high-affinity E1 Ca^{2+}-binding sites, stabilizes the E1 conformational state relative to E2; furthermore, it can be said to power the E2 $\rightarrow$ E1 conformational change (step **6**). Now E1 is ready to transport two more Ca^{2+} ions. Thus the cycle is complete, and hydrolysis of one phosphoanhydride bond in ATP has powered the pumping of two Ca^{2+} ions against a concentration gradient into the SR lumen.

Much structural and biophysical evidence supports the model depicted in Figure 11-10. For instance, the muscle calcium pump has

been isolated with phosphate linked to the key aspartate residue, and spectroscopic studies have detected slight alterations in protein conformation during the E1 → E2 conversion. The two phosphorylated states can also be distinguished biochemically; addition of ADP to phosphorylated E1 results in synthesis of ATP, the reverse of step **2**, whereas addition of ADP to phosphorylated E2 does not. Each principal conformational state of the reaction cycle can also be characterized by a different susceptibility to various proteolytic enzymes such as trypsin.

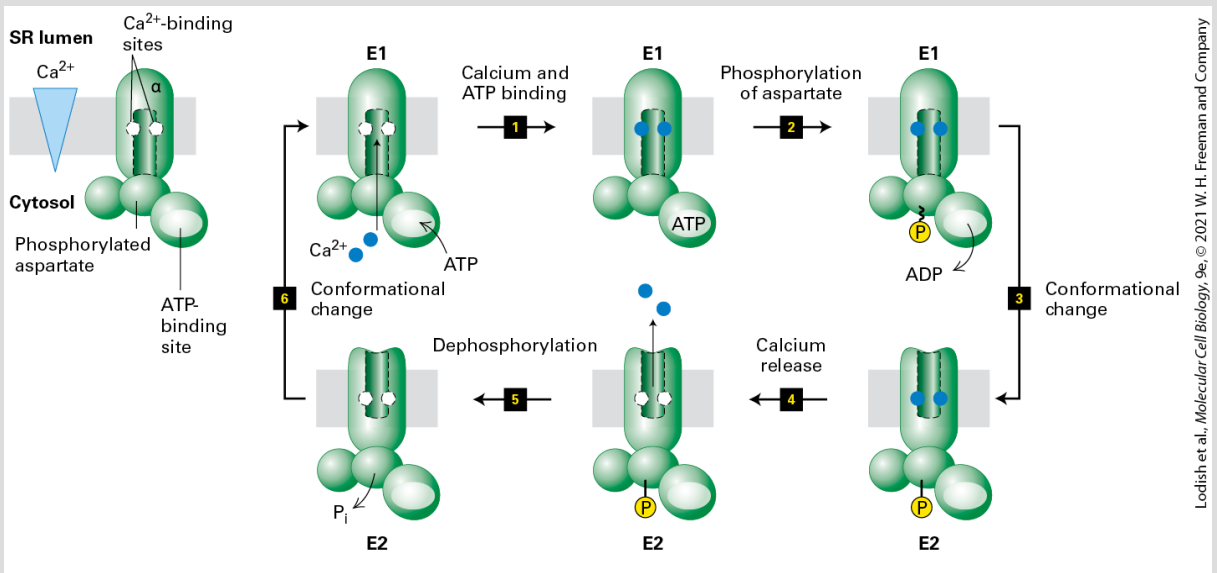


FIGURE 11-10 Operational model of the Ca^{2+} ATPase in the SR membrane of skeletal muscle cells. Only one of the two catalytic α subunits of this P-class pump is depicted. E1 and E2 are alternate conformations of the protein in which the Ca^{2+}-binding sites are accessible from the cytosolic and exoplasmic (SR lumen) faces of the membrane, respectively. An ordered sequence of steps, as diagrammed here, is essential for coupling ATP hydrolysis with the transport of Ca^{2+} ions across the membrane. In the figure, ~P indicates a high-energy aspartyl-phosphate bond; –P indicates a low-energy bond. Because the affinity of Ca^{2+} for the cytosolic-facing binding sites in E1 is 1000-fold greater than its affinity for the exoplasmic-facing sites in E2, this pump transports Ca^{2+} unidirectionally

from the cytosol to the SR lumen. See the text and Figure 11-11 for more detail. See C. Toyoshima and G. Inesi, 2004, *Annu. Rev. Biochem.* **73**:269–292.

Figure 11-11b shows the three-dimensional structure of the Ca^{2+} pump in the E1 state. As can be seen in Figure 11-11c, the 10 membrane-spanning α helices in the catalytic subunit form the passageway through which Ca^{2+} ions move. Amino acids in four of these helices form the two high-affinity E1 Ca^{2+}-binding sites (Figure 11-11a, *left*). One site is formed from negatively charged oxygen atoms from the carboxyl groups (COO^-) of glutamate and aspartate side chains, as well as from bound water molecules. The other site is formed from side- and main-chain oxygen atoms. Thus as Ca^{2+} ions bind to the Ca^{2+} pump, they lose the water molecules that normally surround a Ca^{2+} ion in aqueous solution (see Figure 2-7), but these water molecules are replaced by oxygen atoms with a similar geometry that are part of the transport protein. In contrast, in the E2 state (Figure 11-11a, *right*), several of these binding side chains have moved fractions of a nanometer and are unable to interact with bound Ca^{2+} ions, accounting for the low affinity of the E2 state for Ca^{2+} ions.

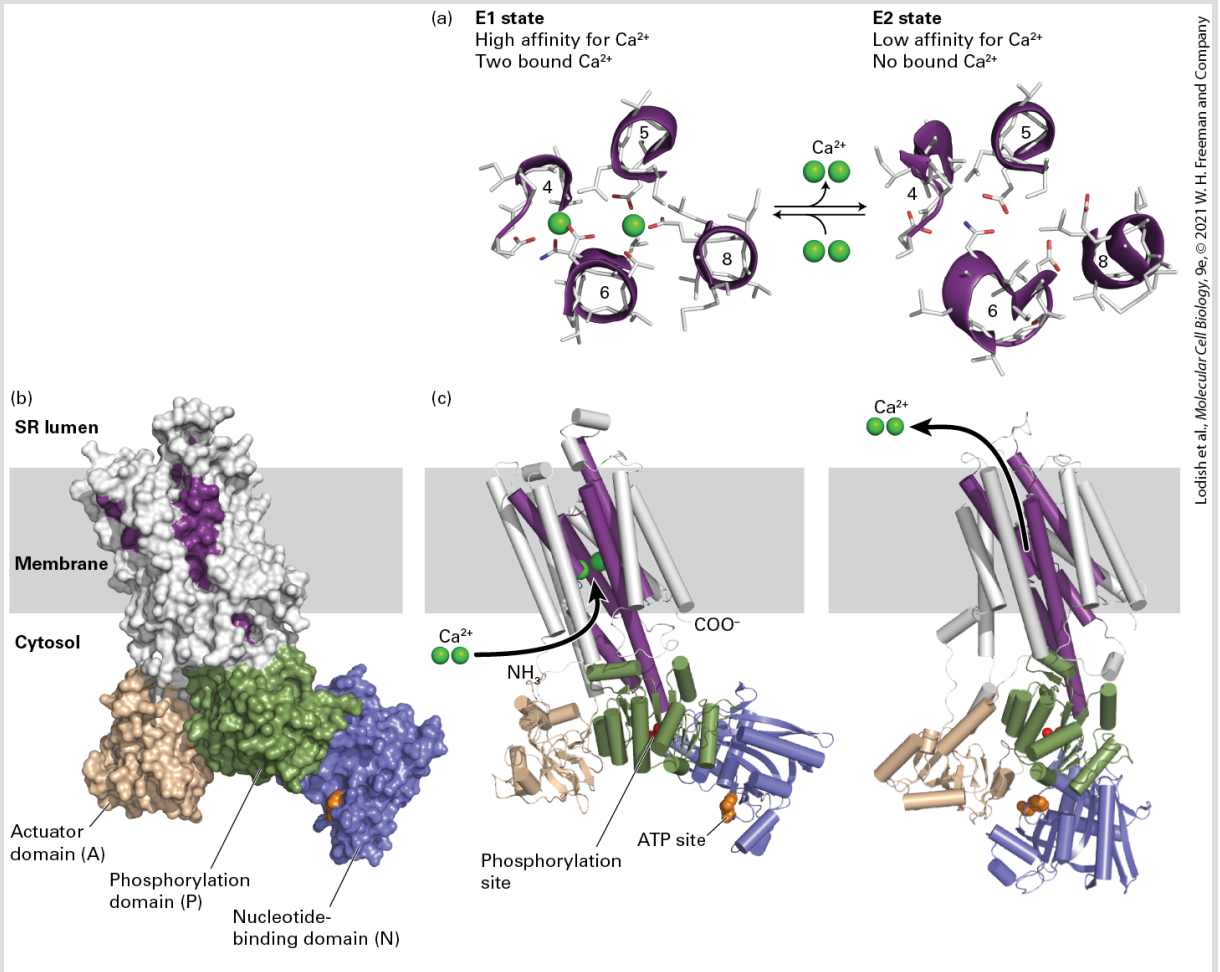


FIGURE 11-11 Structure of the catalytic α subunit of the muscle Ca^{2+} ATPase. (a) Ca^{2+}-binding sites in the E1 state (*left*), with two bound calcium ions, and the low-affinity E2 state (*right*), without bound ions. Side chains of key amino acids are white, and the oxygen atoms on the glutamate and aspartate side chains are red. In the high-affinity E1 conformation, Ca^{2+} ions bind at two sites between helices 4, 5, 6, and 8 inside the membrane. One site is formed out of negatively charged oxygen atoms from glutamate and aspartate side chains and from water molecules (not shown), and the other is formed out of side- and main-chain oxygen atoms. Seven oxygen atoms surround the Ca^{2+} ion in both sites. (b) Three-dimensional model of the protein in the E1 state based on the structure determined by x-ray crystallography. There are 10 transmembrane α helices, four of which (purple) contain residues that participate in Ca^{2+} binding. The cytosolic segment forms three domains: the nucleotide-binding domain (N, blue), the phosphorylation domain (P, green), and the actuator domain (A, beige), which connects two of the membrane-spanning helices. (c) Models of the pump in the E1 state (*left*) and in the E2 state (*right*). Note the differences between the E1 and E2 states in the conformations of the N and A domains. Movements of

these domains power the conformational changes of the membrane-spanning α helices (purple) that constitute the Ca^{2+}-binding sites, converting them from a conformation in which the Ca^{2+}-binding sites are accessible from the cytosolic face (E1 state) to one in which the now loosely bound Ca^{2+} ions gain access to the exoplasmic face (E2 state).

[Data from C. Toyoshima and G. Inesi, 2004, *Annu. Rev. Biochem.* **73**:269–292, PDB ID 1su4; and K. Obara et al., 2005, *Proc. Nat'l. Acad. Sci. USA* **102**:14489–14496, PDB ID 1agv.]

The binding of Ca^{2+} ions to the Ca^{2+} pump illustrates a general principle of ion binding to membrane transport proteins that we will encounter repeatedly in this chapter: as ions bind, they lose most of their waters of hydration, but interact with oxygen atoms in the transport protein that have a geometry similar to that of the water oxygen atoms that are bound to them in aqueous solution. This reduces the thermodynamic barrier for ion binding to the protein and allows the ion to bind tightly even from solutions of relatively low concentrations.

The cytosolic region of the Ca^{2+} pump consists of three domains that are well separated from one another in the E1 state (see Figure 11-11b). Each of these domains is connected to the membrane-spanning helices by short segments of amino acids. Movements of these cytosolic domains during the pumping cycle cause movements of the connecting segments, which are transmitted into movements of the attached membrane-spanning α helices. For example, the phosphorylated residue, Asp 351, is located in the phosphorylation (P) domain. The adenosine moiety of ATP binds to the nucleotide-binding (N) domain, but the γ phosphate of ATP

binds to specific residues on the P domain, requiring movements of both the N and P domains. Thus following ATP and Ca^{2+} binding, the γ phosphate of the bound ATP sits adjacent to the aspartate on the P domain that is to receive the phosphate. Although the precise details of these and other protein conformational changes are not yet clear, the movements of the N and P domains are transmitted by lever-like motions of the connecting segments into rearrangements of several membrane-spanning α helices. These changes are especially apparent in the four helices that contain the two Ca^{2+}-binding sites: the changes prevent the bound Ca^{2+} ions from moving back into the cytosol when released, but enable them to dissociate into the exoplasmic space (lumen).

All P-class ATP-powered pumps, regardless of which ion they transport, are phosphorylated on a highly conserved aspartate residue during the transport process. As deduced from cDNA sequences, the catalytic α subunits of all the P-class pumps examined to date have similar amino acid sequences and thus are presumed to have similar arrangements of transmembrane α helices and cytosol-facing A (actuator), P, and N domains (see Figure 11-11). These findings strongly suggest that all such proteins evolved from a common precursor, although they now transport different ions. This suggestion is borne out by the finding that the three-dimensional structures of the membrane-spanning segments of the Na^+/K^+ ATPase and the Ca^{2+} pump are highly similar; the molecular structures of the three cytosolic domains are also very similar. Thus

the operational model in Figure 11-11 is generally applicable to all of the P-class ATP-powered pumps.

The Na^+/K^+ ATPase Maintains the Intracellular Na^+ and K^+ Concentrations in Animal Cells

An important P-class pump that is present in the plasma membranes of all animal cells is the **Na^+/K^+ ATPase**. This ion pump is a tetramer of subunit composition $\alpha_2\beta_2$ and shares structural homology with the Ca^{2+} pump. The small, glycosylated β-transmembrane polypeptide apparently is not involved directly in ion pumping. During its catalytic cycle, the Na^+/K^+ ATPase moves three Na^+ ions *out* of and two K^+ ions *into* the cell per ATP molecule hydrolyzed.

The mechanism of action of the Na^+/K^+ ATPase, outlined in Figure 11-12, is similar to that of the muscle SR calcium pump, except that ions are pumped in *both* directions across the membrane, with each ion moving *against* its concentration gradient. In its E1 conformation, the Na^+/K^+ ATPase has three high-affinity Na^+-binding sites and two low-affinity K^+-binding sites accessible from the cytosolic surface of the protein. The K_d for binding of Na^+ to these cytosolic sites is 0.6 mM, a value considerably lower than the

intracellular Na^+ concentration of $\sim$12 mM; as a result, Na^+ ions normally fully occupy these sites. Conversely, the affinity of the cytosolic K^+-binding sites is low enough that K^+ ions, transported inward through the protein, dissociate from E1 and enter the cytosol despite the high intracellular K^+ concentration. During the E1 $\rightarrow$ E2 transition, the three bound Na^+ ions gain access to the exoplasmic face, and simultaneously, the affinity of the three Na^+-binding sites drops. The three Na^+ ions, now bound to low-affinity Na^+ sites, dissociate one at a time and enter the extracellular medium despite the high extracellular Na^+ concentration. Transition to the E2 conformation also generates two high-affinity K^+ sites accessible from the exoplasmic face. Because the K_d for K^+ binding to these sites (0.2 mM) is lower than the extracellular K^+ concentration (4 mM), these sites will fill with K^+ ions as the Na^+ ions dissociate. Similarly, during the subsequent E2 $\rightarrow$ E1 transition, the two bound K^+ ions are transported inward and then released into the cytosol.

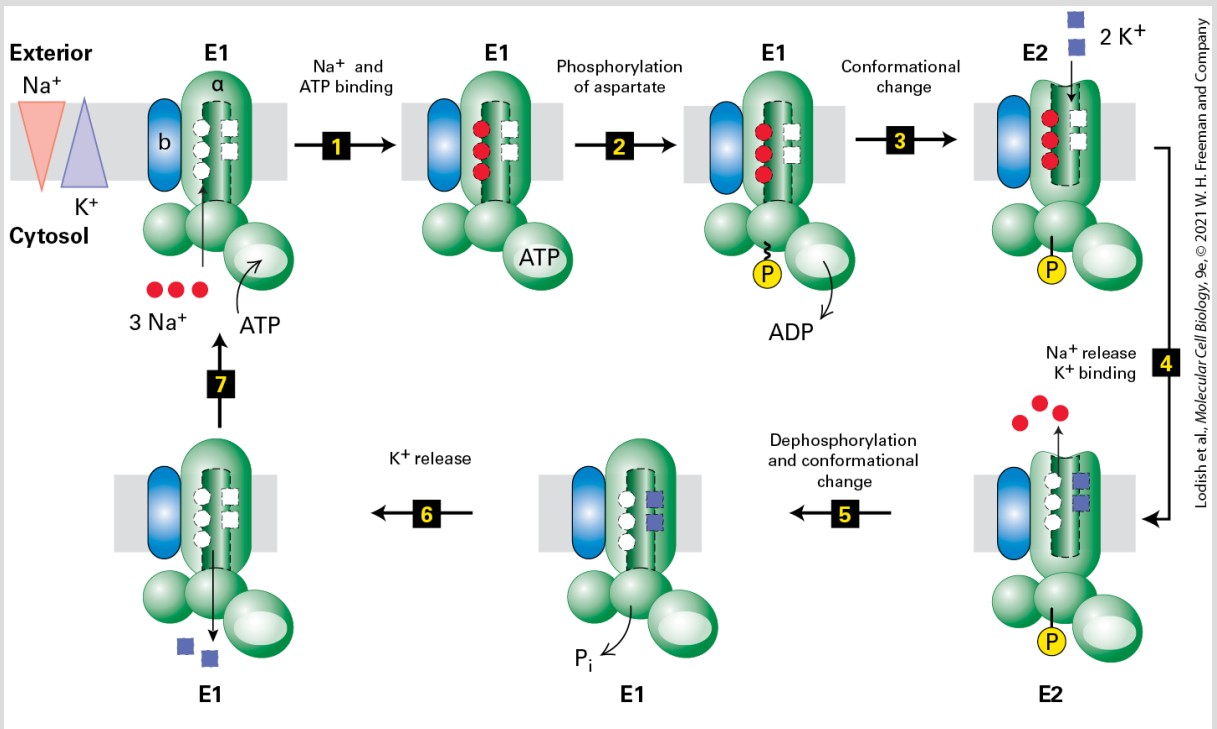


FIGURE 11-12 Operational model of the plasma-membrane Na^+/K^+ ATPase. Only one of the two catalytic α subunits of this P-class pump is depicted. It is not known whether just one or both subunits in a single ATPase molecule transport ions. Ion pumping by the Na^+/K^+ ATPase involves phosphorylation, dephosphorylation, and conformational changes similar to those in the muscle Ca^{2+} ATPase (see Figure 11-11). In this case, hydrolysis of the E2–P intermediate powers the $E2 \rightarrow E1$ conformational change and concomitant transport of two K^+ ions inward. Na^+ ions are indicated by red circles; K^+ ions, by purple squares; high-energy acyl phosphate bond, by $\sim P$; low-energy phosphoester bond, by –P.

Certain drugs (e.g., ouabain and digoxin) bind to the exoplasmic domain of the plasma-membrane Na^+/K^+ ATPase and specifically inhibit its ATPase activity. The resulting disruption in the Na^+/K^+ balance of cells is strong evidence for the critical role of this ion pump in maintaining the normal K^+ and Na^+ ion concentration

gradients. Classic Experiment 11-1 describes the discovery of this important pump, which is required for life.

V-Class H^+ ATPases Maintain the Acidity of Lysosomes and Vacuoles

All V-class ATPases transport only H^+ ions. These proton pumps, present in the membranes of lysosomes, endosomes, and plant vacuoles, function to acidify the lumina of these organelles. The pH of the lysosomal lumen can be measured precisely in live cells by use of particles labeled with a pH-sensitive fluorescent dye. When these particles are added to the extracellular fluid, the cells engulf and internalize them (phagocytosis; see Figure 1-19 and Chapter 17), ultimately transporting them into lysosomes. The lysosomal pH can be calculated from the spectrum of the fluorescence emitted. Alternatively, the DNA encoding a naturally fluorescent protein whose fluorescence depends on the pH can be modified (by adding DNA segments encoding "signal sequences," detailed in Chapters 13 and 14) such that the protein is targeted to the lysosome lumen; fluorescence measurements can then be used to determine the pH in the organelle lumen.

These measurements reveal that a hundredfold or more proton gradient is maintained between the lysosomal lumen (pH ~4.5–5.0) and the cytosol (pH ~7.0). The low lysosomal pH is necessary for

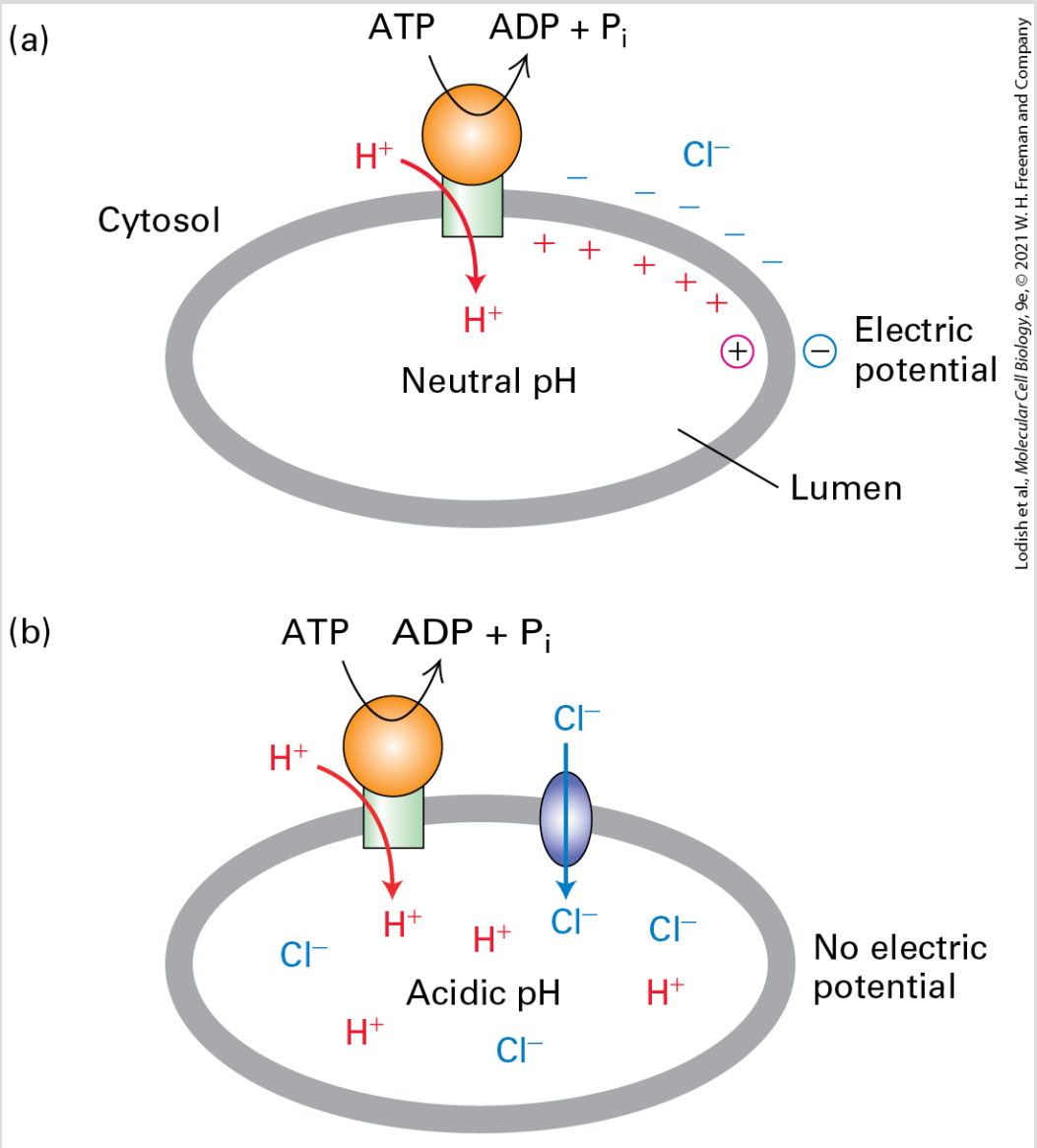
optimal function of the many proteases, nucleases, and other hydrolytic enzymes in the lumen; on the other hand, a cytosolic pH of 5 would disrupt the functions of many proteins optimized to act at pH 7 and lead to death of the cell. Maintenance of the proton gradient depends on a V-class ATPase and thus on ATP production by the cell.

Relatively few protons need to be pumped to acidify an intracellular vesicle. To understand why, recall that a solution of pH 4 has an H^+ ion concentration of 10^{-4} moles per liter, or 10^{-7} moles of H^+ ions per milliliter. There are 6.02×10^{23} atoms of H per mole (Avogadro's number), so a milliliter of a pH 4 solution contains 6.02×10^{16} H^+ ions. Thus at pH 4, a primary spherical lysosome with a volume of 4.18×10^{-15} ml (diameter of $0.2\,\mu m$) would contain just 252 protons. At pH 7, the same organelle would have an average of only 0.2 protons in its lumen, and thus pumping of only about 250 protons would be necessary for lysosome acidification.

By themselves, V-class proton pumps cannot acidify the lumen of an organelle (or the extracellular space) because these pumps are *electrogenic*; that is, a net movement of electric charge occurs during transport. Pumping of just a few protons causes a buildup of positively charged H^+ ions on the exoplasmic (inside) face of the organelle membrane. For each H^+ pumped across, a negative ion (e.g., OH^- or Cl^-) will be left behind on the cytosolic face, causing a buildup of negatively charged ions there. These oppositely charged

ions attract each other on opposite faces of the membrane, generating a charge separation, or electric potential, across the membrane. The lysosome membrane thus functions as a capacitor in an electric circuit, storing opposing charges (anions and cations) on opposite sides of a barrier impermeable to the movement of charged particles.

As more and more protons are pumped and build up excess positive charge on the exoplasmic face, the energy required to move additional protons against this rising electric potential gradient increases dramatically and prevents pumping of additional protons long before a significant transmembrane H^+ concentration gradient is established (Figure 11-13a). In fact, this is the way that P-class proton pumps generate a cytosol-negative potential across plant and yeast plasma membranes.

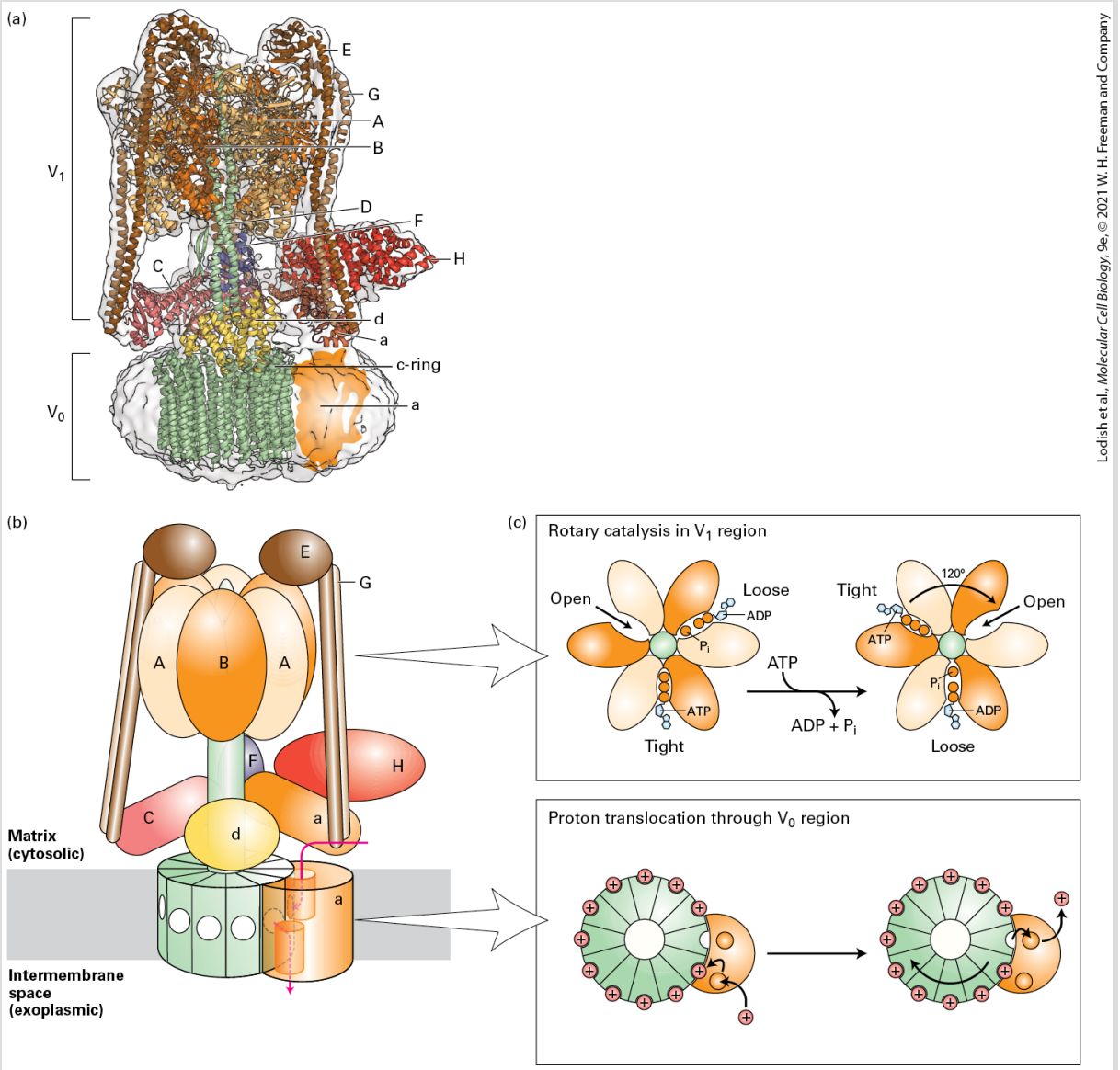


Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 11-13 Effect of V-class proton pumps on H^+ concentration gradients and electric potential gradients across cellular membranes. (a) If an intracellular organelle contains only V-class pumps, proton pumping generates an electric potential across the membrane (the cytosolic face becomes negative and the luminal face positive) but no significant change in the intraluminal pH. (b) If the organelle membrane also contains Cl^- channels, anions passively follow the pumped protons, resulting in an accumulation of H^+ and Cl^- ions in the lumen (low luminal pH) but no electric potential across the membrane.

In order for an organelle lumen or an extracellular space (e.g., the lumen of the stomach) to become acidic, movement of protons must be accompanied either by (1) movement of an equal number of anions (e.g., Cl^-) in the same direction or by (2) movement of equal numbers of a different cation in the opposite direction. The first process occurs in lysosomes and plant vacuoles, whose membranes contain V-class H^+ ATPases and anion channels through which accompanying Cl^- ions move (Figure 11-13b). The second process occurs in the lining of the stomach, which contains a P-class H^+/K^+ ATPase that is not electrogenic and pumps one H^+ outward and one K^+ inward. Operation of this pump is discussed later in the chapter.

The V-class proton pumps in lysosomal and vacuolar membranes have been solubilized, purified, and incorporated into liposomes for further study. As shown in Figure 11-14, these pumps consist of two discrete domains: a soluble hydrophilic domain (V_1) on the cytosolic side, and a ringlike transmembrane domain (V_0). V_1 and V_0 are linked by three peripheral stalk structures. The V_1 and V_0 domains each function as rotary motors. ATP hydrolysis by the V_1 domain is coupled to proton translocation through the membrane by means of the rotation of both domains. More specifically, ATP hydrolysis triggers conformational changes that drive the rotation of V_1 by 120° per molecule of ATP hydrolyzed. The rotation of V_1 in turn generates a torque that drives the rotation of the c-subunits against the a subunits in V_0, resulting in proton translocation across the membrane.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 11-14 Eukaryotic V-ATPase: rotation of V_0 and V_1 domains drive proton transport across the membrane. (a) The structure of yeast vacuolar V-ATPase as determined by cryo-electron microscopy reveals separate cytosolic V_0 and transmembrane V_1 domains. (b) Cartoon of the pump. (c) *Top:* Cross-sectional view of V_1. The three nucleotide binding sites in V_1 exist in three distinct conformations: tight, with ATP bound; loose, with ADP bound; and open, without any nucleotide bound. Each hydrolysis of ATP triggers a 120° rotation of V_1 relative to V_0. *Bottom:* Rotation of the c subunits in V_0 against the a subunit drives proton translocation from the cytosol to the lumen via a series of protonation and deprotonation of glutamate and aspartate residues.

[Data from J. Zhao et al., 2015, *Nature* **521**:241.]

Recent cryo-electron microscopy structures of the V_0 subunit of a yeast vacuolar V-ATPase has yielded molecular insights into how protons are translocated through the membrane by V_0. These studies revealed the presence of two negatively charged cavities: one gives access to the cytosolic side of the membrane and the other to the luminal side. Positively charged amino acid residues form a barrier in the middle of the membrane that separates the two cavities. Protons enter the cavity on the cytosolic side, and conserved glutamates or aspartates undergo a cycle of protonation and deprotonation as the protons translocate through the pore into the lumen of the vacuole (Figure 11-14b). The rotation of the subunits closes access to the cytoplasm and traps the proton within the pump, and further rotation then releases the protons to the lumen of the vacuole.

ABC Proteins Export a Wide Variety of Drugs and Toxins from the Cell

About 50 different mammalian ABC proteins are now recognized, and mutations in these proteins have been linked to many human diseases (Table 11-3). In eukaryotic cells, ABC proteins localize not only to the plasma membrane but also to the membranes of many intracellular organelles. Several are expressed in abundance in the liver, intestines, and kidneys — sites where natural toxic and waste

products are removed from the body. Substrates for these ABC proteins include sugars, amino acids, cholesterol, bile acids, phospholipids, peptides, proteins, toxins, and foreign substances. The normal function of ABCB1 most likely is to transport various natural and metabolic toxins into the bile or intestinal lumen for excretion or into the urine being formed in the kidney. During the course of its evolution, ABCB1 appears to have acquired the ability to transport drugs whose structures are similar to those of these endogenous toxins. Tumors derived from MDR-expressing cell types, such as hepatomas (liver cancers), are frequently resistant to virtually all chemotherapeutic agents and are thus difficult to treat, presumably because the tumors exhibit increased expression of ABCB1 or a related ABC protein.

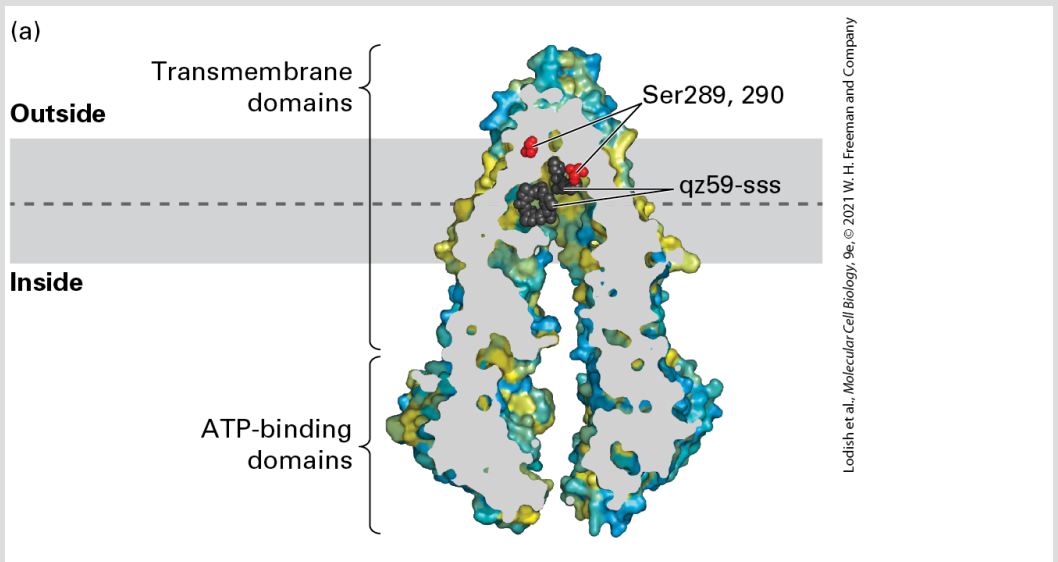
TABLE 11-3 • Selected Human ABC Proteins

Protein	Tissue Expression	Function	Disease Caused by Defective Protein
ABCB1 (MDR1)	Many epithelia, blood-brain barrier	Exports glucosylceramides and lipophilic drugs	Inflammatory bowel disease
ABCB4 (MDR2)	Liver	Exports phosphatidylcholine into bile	
ABCB11	Liver	Exports bile salts into bile	Progressive familial intrahepatic cholestasis
CFTR	Exocrine	Transports Cl ions	Cystic fibrosis

	tissue		
ABCDI	Ubiquitous in peroxisomal enzyme	Influences activity of peroxisomal enzyme that oxidizes very long chain fatty acids	Adrenoleukodystrophy (ADL)
ABCG5/8	Liver, intestine	Exports cholesterol and other sterols	β-Sitosterolemia
ABCA1	Ubiquitous	Exports cholesterol and phospholipid for uptake into high-density lipoprotein (HDL)	Tangier's disease
ABCA4	Retina	Transports N-retinyl-phosphatidylethanolamine	Stargardt's disease in photoreceptor cells (juvenile macular degeneration)
ABCA2	Brain, kidney, heart, lung	Exports cholesterol	Alzheimer's disease
ABCA3	Lung	Exports surfactant lipids	Surfactant metabolism dysfunction 3
ABCA12	Lung, skin	Exports lipids	Harlequin ichthyosis
ABCB4	Liver	Exports long-chain phosphatidylcholines	Progressive familial intrahepatic cholestasis 3
ABCC2	Liver, kidney, intestines	Exports bilirubin, bile salts	Dubin-Johnson

ABCD1	Ubiquitous	Exports very long chain fatty acids	Adrenoleukodystrophy
ABCG5	Liver, intestine	Exports cholesterol, plant sterols	Sitosterolemia
ABCG8	Liver, intestine	Exports cholesterol, plant sterols	Sitosterolemia, gallstones

As noted earlier, all members of the very large and diverse ABC superfamily of membrane transport proteins contain two transmembrane (TM) domains and two cytosolic ATP-binding (A) domains (see Figure 11-9). The TM domains, each built of 10 membrane-spanning α helices, form the pathway through which the transported substance (substrate) crosses the membrane (Figure 11-15a) and determine the substrate specificity of each ABC protein. The sequences of the A domains are approximately 30–40 percent homologous in all members of this superfamily, indicating a common evolutionary origin.



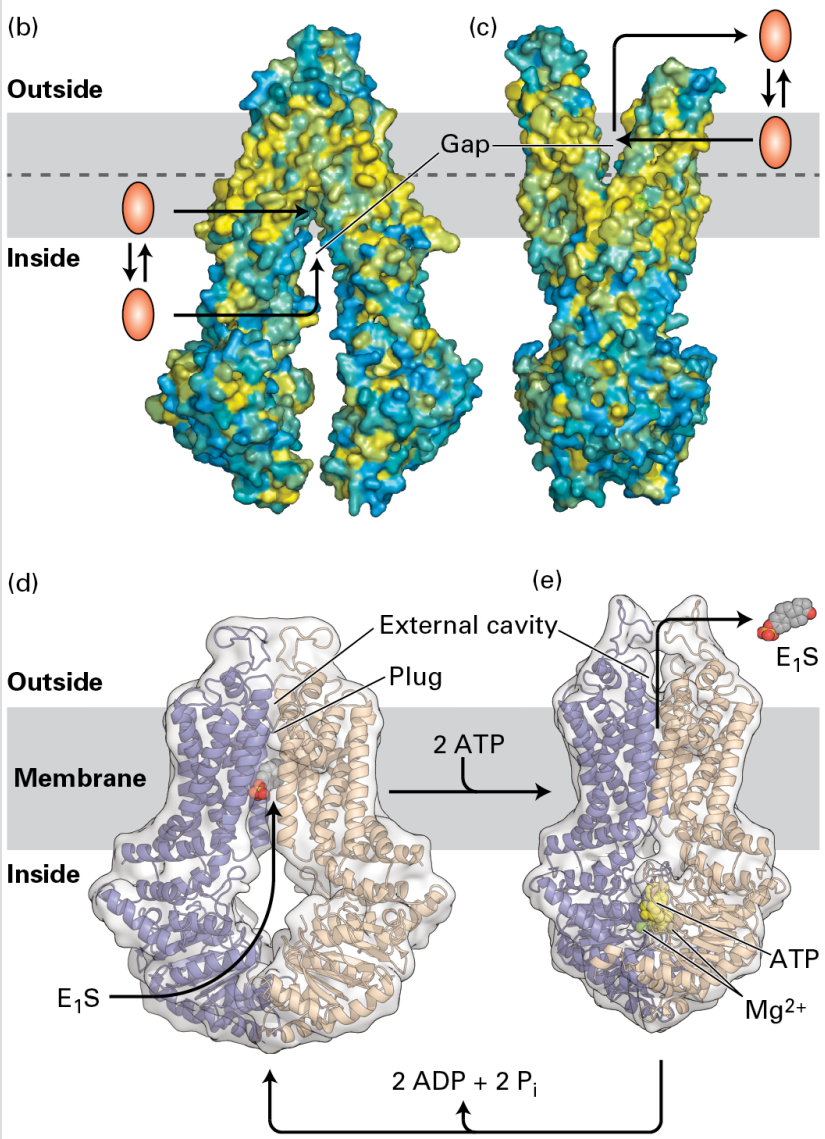


FIGURE 11-15 The multidrug transporter ABCB1 (MDR1): Structure and model of ligand export. (a) Cross-sectional view through the center of an ABCB1 protein bound to two molecules of a drug analog, qz59-sss (black), reveals the central location of the ligand-binding site in relation to the phospholipid bilayer: the central ligand-binding cavity is close to the leaflet-leaflet interface of the membrane. During transport, this binding cavity is alternately exposed to the exoplasmic and the cytosolic surface of the membrane. Serines 289 and 290 affect the ligand specificity of the transporter; they are shown as red spheres to highlight their juxtaposition to the bound ligand. Surface residues are colored yellow to denote hydrophobic and blue to denote hydrophilic amino acids. (b)

Three-dimensional structure of ABCB1 with its ligand-binding site facing inward toward the cytosol. In this conformation, a hydrophilic ligand can bind directly from the cytosol. A more hydrophobic ligand can enter into the inner leaflet of the plasma membrane bilayer and then enter the ligand-binding site through a gap in the protein that is accessible directly from the hydrophobic core of the inner leaflet. (c) Model for the structure of ABCB1 with its ligand-binding site facing outward, based on the structures of homologous bacterial ABC proteins. When the protein assumes this conformation, the ligand can either diffuse into the exoplasmic leaflet or directly into the aqueous extracellular medium. (d) Cryo-electron microscopy structure of human ABCG2 in the ADP-bound, inward-facing state, revealing a large cavity facing the cytosol into which the substrate can enter. (e) Mg^{2+} ions, which are important for ATP binding and activity, are shown as light green spheres. ATP (yellow spheres) binding triggers rotation of the α helices, compression of the large cavity, and extrusion of the substrate outside of the cell.

[Data from S. G. Aller et al., 2009, *Science* **323**:1718–1722, PDB ID 3g61; and I. Manolaridis et al., 2018, *Nature* **563**:426–430, PDB ID 6FEQ.]

The first eukaryotic ABC protein to be recognized was discovered during studies on tumor cells and cultured cells that exhibited resistance to several drugs with unrelated chemical structures. Such cells were eventually shown to express elevated levels of a *multidrug-resistance (MDR) transport protein* originally called *MDR1* and now known as *ABCB1*. This protein uses the energy derived from ATP hydrolysis to *export* a large variety of drugs from the cytosol to the extracellular medium. The *Mdr1* gene is frequently amplified in multidrug-resistant cells, resulting in a large overproduction of the ABCB1 protein. In contrast to bacterial ABC proteins, which are built of four discrete subunits, all four domains of mammalian ABCB1 are fused into a single 170-kDa protein.

The substrates of mammalian ABCB1 are primarily planar, lipid-soluble molecules with one or more positive charges; they all compete with one another for transport, which suggests that they bind to the same or overlapping sites on the protein. Many drugs transported out of the cell by ABCB1 have earlier diffused into the cell cytosol from the extracellular medium across the plasma membrane, unaided by transport proteins. Once in the cell, they block various cellular functions. Two such drugs, used to treat certain cancers, are colchicine and vinblastine, which block assembly of microtubules (see Chapter 18). ATP-powered export of such drugs by MDR1 reduces their concentration in the cytosol. As a result, a much higher extracellular drug concentration is required to kill cells that express ABCB1 than those that do not. That ABCB1 is an ATP-powered small-molecule pump has been demonstrated with liposomes containing the purified protein. Different drugs enhance the ATPase activity of these liposomes in a dose-dependent manner corresponding to their ability to be transported by ABCB1.

The three-dimensional structure of ABCB1, together with those of homologous bacterial and eukaryotic ABC proteins, revealed the protein's mechanism of transport as well as the key to its ability to bind and transport a wide array of hydrophilic and hydrophobic substrates (see Figure 11-15). The two TM domains form a binding site in the center of the membrane that alternates between an inward-facing (Figure 11-15b) and an outward-facing (Figure 11-15c) orientation, conforming to the alternating access model. The alternation between these two conformational states is powered by

ATP binding to the two A subunits and its subsequent hydrolysis to ADP and P_i.

The substrate-binding cavity formed by ABCB1 is large. Some of the amino acids that line the cavity — mainly tyrosine and phenylalanine — have aromatic side chains, allowing ABCB1 to bind multiple types of hydrophobic ligands. Other segments of the cavity are lined with hydrophilic residues, allowing hydrophilic or amphipathic molecules to bind. In the inward-facing conformation, the binding site is open directly to the surrounding aqueous solution, allowing hydrophilic molecules to enter the binding site directly from the cytosol. In addition, a gap in the protein is accessible directly from the hydrophobic core of the inner leaflet of the membrane bilayer; through this gap hydrophobic molecules can enter the binding site directly from the inner leaflet (see Figure 11-15b). After the ATP-powered change to the outward-facing conformation, molecules can exit the binding site into the outer membrane leaflet or directly into the extracellular medium (see Figure 11-15c).

The human ABCG2 structure has recently been determined in substrate- and ATP-bound states. This structure provided additional insights into the mechanisms underlying ATP-driven substrate translocation (Figure 11-15d, e). ABCG2 is also known as the Breast Cancer Resistance Protein and its expression levels are correlated with a poor outcome in breast as well as other cancers. A clever

experimental strategy made use of a mutant form of ABCG2 that exhibits decreased rates of transport and of ATP hydrolysis. High-resolution cryo-electron microscopy was used to capture the transporter bound to the substrate in the internal facing, pre-translocation state and bound to ATP in the post-translocation state. These studies revealed a large substrate-binding cavity in the internal-facing state, with a plug creating a smaller cavity toward the external-facing part of the protein. ATP binding triggered rotation of the α helices of the A domains, compressing the substrate-binding cavity and extruding the substrate through the plug to the external cavity, with release of the substrate to the outside. While not clear, ATP hydrolysis may be required to return the protein transporter to the inward-facing state.

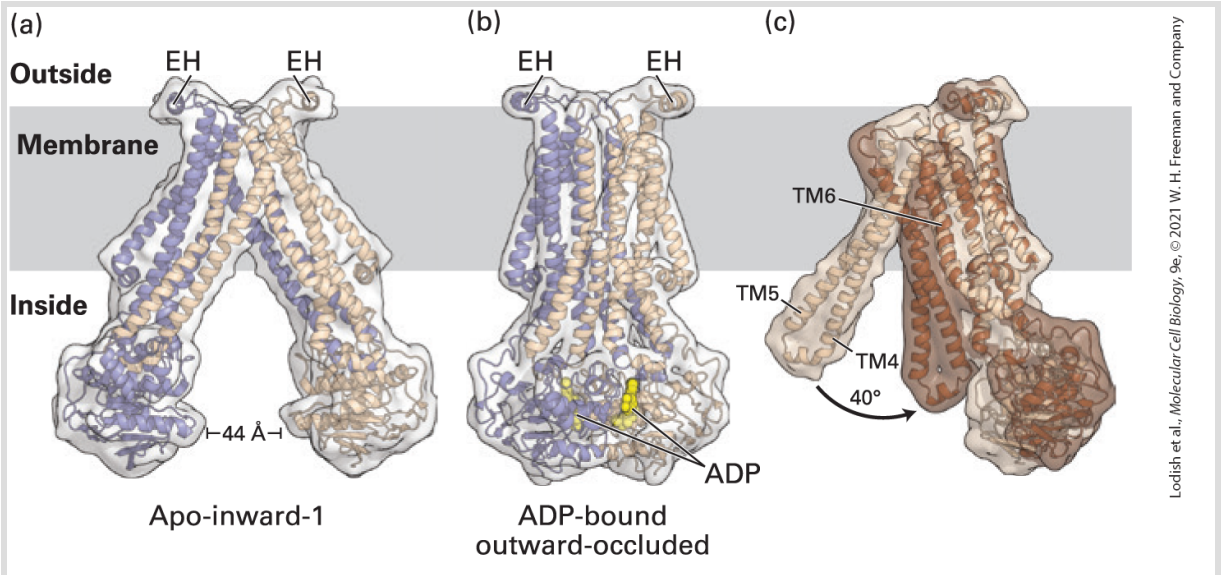
Certain ABC Proteins "Flip" Phospholipids and Other Lipid-Soluble Substrates from One Membrane Leaflet to the Other

As shown in Figure 11-15, ABCB1 and ABCG2 can move, or "flip," a hydrophobic or amphipathic substrate molecule from the inner leaflet of the membrane to the outer leaflet. This otherwise energetically unfavorable reaction is powered by the ATPase activity of the protein, and is required for maintaining the asymmetric distribution of phospholipids across the cytoplasmic and exoplasmic

membrane leaflets (Chapter 10) and for transporting lipids across cellular membrane. Support for this so-called **flippase** model of transport by ABCB1 comes from experiments on ABCB4 (originally called *MDR2*), a protein homologous to ABCB1 that is present in liver cells in the region of the plasma membrane that faces the thin tubes (canaliculi) that collect bile. ABCB4 moves phosphatidylcholine from the cytosolic to the exoplasmic leaflet of the plasma membrane for subsequent release into the bile in combination with cholesterol and bile acids, which themselves are transported by other ABC superfamily members. Still other ABC superfamily members are also flippases that export various lipids from cells, presumably by mechanisms similar to that of ABCB1 (see Table 11-3).

The flipping of lipids across membranes is required for the generation of glycoproteins within the endoplasmic reticulum in eukaryotic cells or the periplasm of prokaryotes (see Chapter 13). This process provides a way to get the hydrophilic oligosaccharides needed to form glycoproteins across the hydrophobic membrane. In this process, oligosaccharides are added to the cytosolic side of long polyisoprenol membrane lipids to form lipid-linked oligosaccharides (LLOs). These LLOs are then flipped across the membrane into the endoplasmic reticulum (or into the periplasm in prokaryotes) so that the oligosaccharides can be transferred to membrane and secreted proteins to generate glycoproteins. Recent x-ray crystal structures of a dimeric ABC transporter from

Campylobacter jejuni (a bacteria that commonly causes food poisoning) that mediates flipping of LLOs across the membrane captured the transporter in inward- and outward-facing states. These structures revealed a novel lipid-flipping mechanism (Figure 11-16d). Initially, the lipid tail of a lipid-linked oligosaccharide (LLO) binds the extracellular-facing helix of one of the flippase subunits (step **1**). Its binding promotes exchange of ATP for ADP, triggering the dimer to convert to an outward-open state (step **2**) that allows the hydrophilic oligosaccharide head of the LLO to enter the positively charged pore of the transporter (step **3**). ATP hydrolysis drives the return of the dimer to an inward-open conformation, squeezing the LLO head to the outward side of the membrane (step **4**). The LLO lipid tail then dissociates from the flippase and extends into the cytosolic side of the membrane. In this mechanism, the transporter protects the hydrophilic oligosaccharide head of the LLO from the hydrophobic milieu of the membrane, allowing the lipid tails to flip from one side of the membrane to the other.



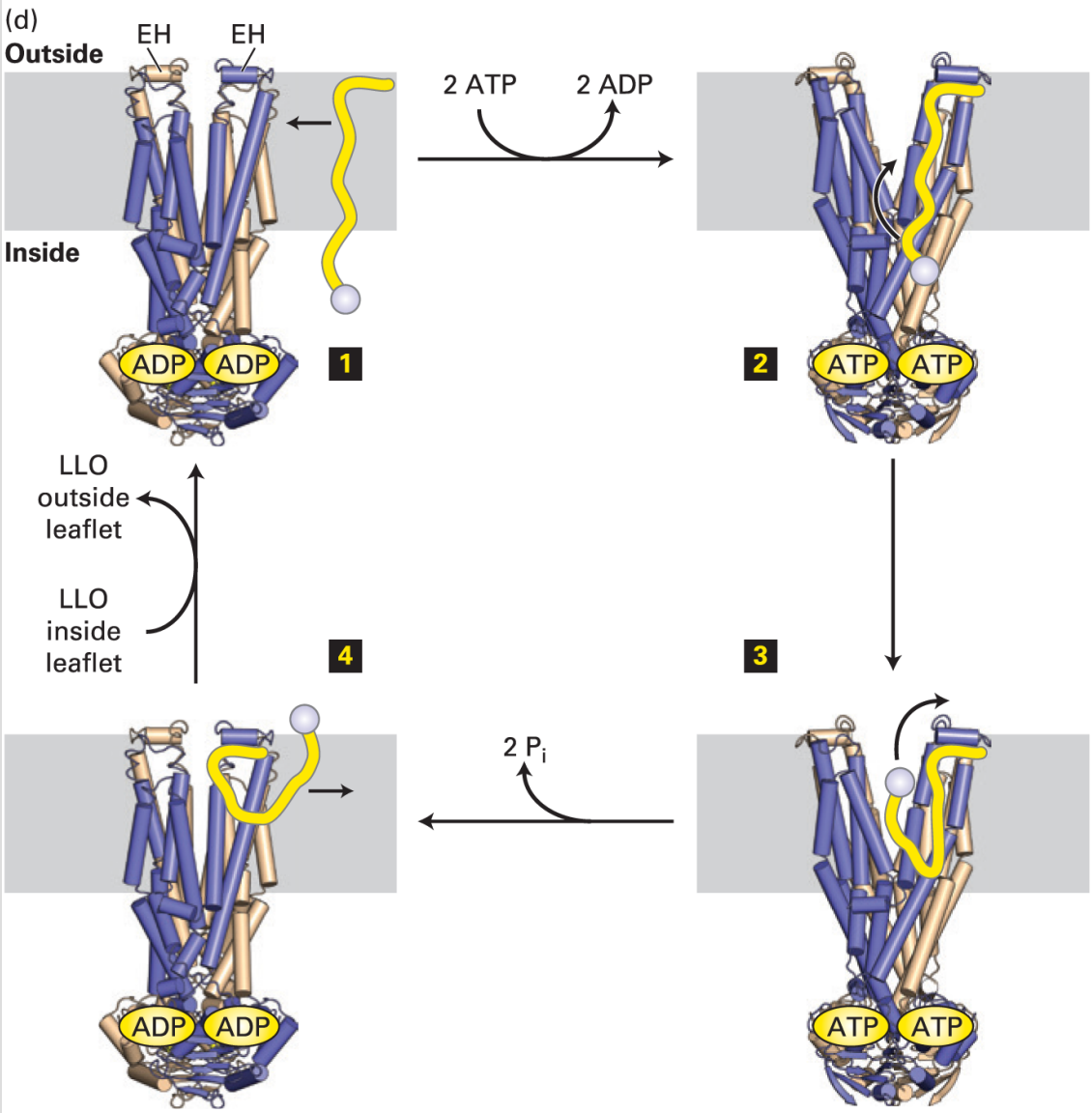


FIGURE 11-16 An ABC transporter from *Campylobacter jejuni* flips lipid-linked oligosaccharides across the membrane. (a) The inward-facing structure of the PglK ABC transporter, with a larger cavity facing the cytosol. Subunits are shown in peach and purple, with the extracellular-facing helix (EH) labeled. (b) The ADP-bound, outward-facing structure in which the cytosolic cavity is occluded. (c) The conversion of inward- to outward-facing states is associated with conformational changes. (d) In the proposed model for the flippase mechanism, the PglK subunits are shown as cartoon-like structures in peach and purple, with the EH labeled. The lipid tail of the LLO is yellow and the oligosaccharide head light purple. The lipid tail interacts with the EH (step 1 – 2); ATP replaces ADP, which opens the cytosolic cavity (step 2), into which the large oligosaccharide head of the LLO enters (step

3); ATP hydrolysis triggers conformational changes that achieve occlusion of the cytosolic cavity, dislodging and ejecting the oligosaccharide head into the periplasm (step 4) and flipping the direction of the lipid tail within the membrane.

[Information from C. Perez et al., 2015, *Nature* **524**:433–438.]

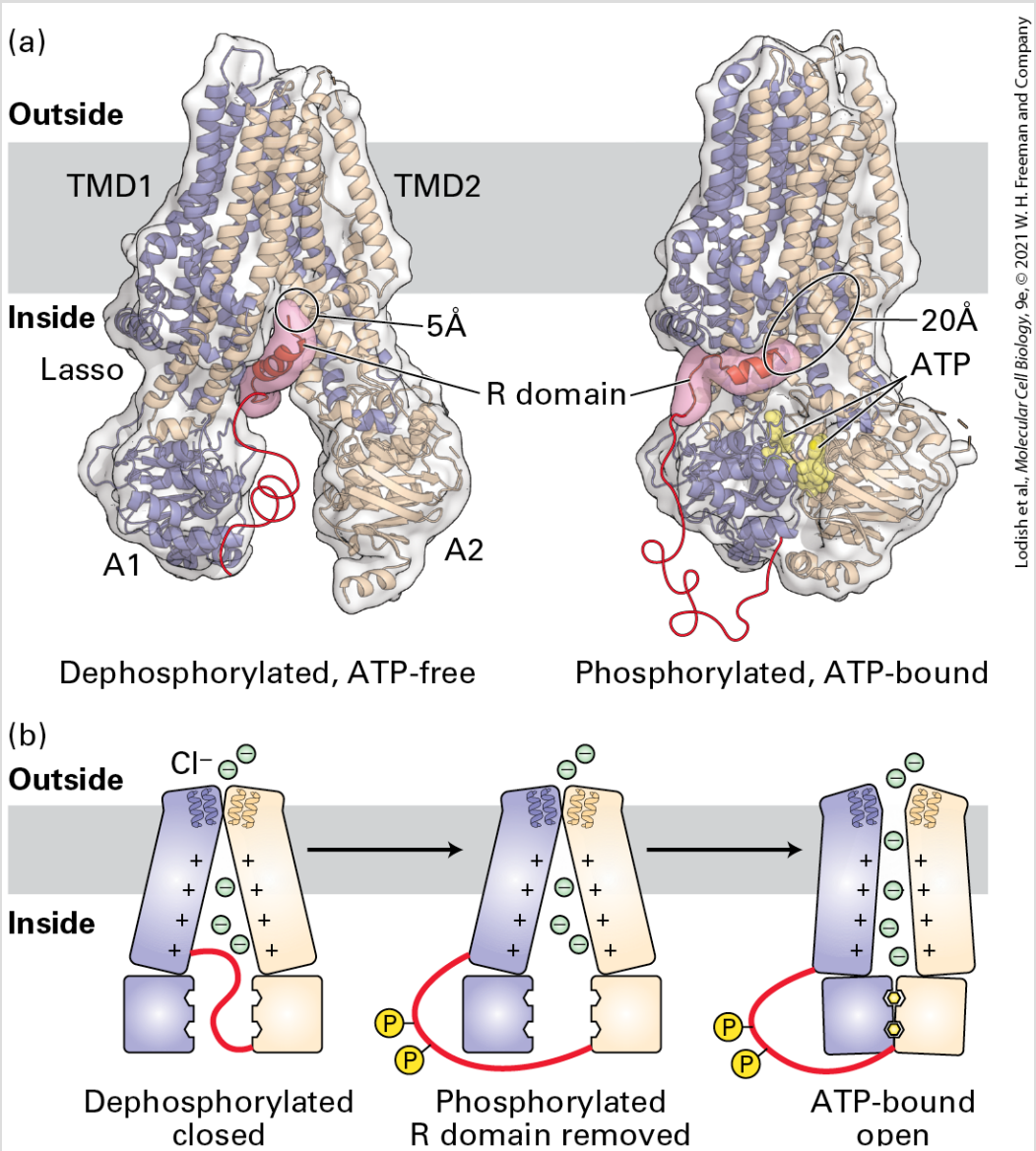
The ABC Cystic Fibrosis Transmembrane Regulator Is a Chloride Channel, Not a Pump

Several human genetic diseases are associated with defective ABC proteins (see Table 11-3). The most common (1 in 2500 births) is cystic fibrosis (CF), caused by a mutation in the gene encoding the *cystic fibrosis transmembrane regulator* (*CFTR*, also called *ABCC7*). Like other ABC proteins, CFTR has two transmembrane TM domains and two cytosolic ATP-binding domains (A domains). CFTR contains an additional R (regulatory) domain on the cytosolic face; R links the two homologous halves of the protein, creating an overall domain organization of TM1–A1–R–TM2–A2. CFTR is a Cl^- channel that opens upon phosphorylation and ATP binding to allow Cl^- anions to flow down their electrochemical gradient. It is expressed in the apical plasma membranes of epithelial cells in the lungs, sweat glands, pancreas, and other tissues, and mutations in the protein decrease the capacity of these membranes to transport Cl^-. For instance, CFTR protein is important for reuptake into the cells of sweat glands of Cl^- lost by sweating; babies with cystic fibrosis, if

licked, often taste salty because this reuptake is inhibited. Over time, as a result of disrupted salt and water homeostasis, patients with cystic fibrosis accumulate dehydrated mucus in their lungs and other tissues, leading to chronic respiratory infections and digestive system disorders, which can ultimately result in organ failure.

Until recently, our understanding of CFTR channel structure and function was derived from biochemical and biophysical studies, along with x-ray crystal structures of fragments of the CFTR protein. In 2017, however, the high-resolution structures of the full-length zebrafish and human CFTR channels were solved using cryo-electron microscopy (Figure 11-17). While it was known that the channel was gated by phosphorylation and ATP binding, comparison of the structure of open and closed CFTR channels revealed the major conformational changes produced by phosphorylation and ATP binding to open the channel. These structures show that channel gating is regulated by two processes: phosphorylation of the cytosolic R domain by cAMP-dependent protein kinase (PKA, discussed in Chapter 15), followed by binding of ATP to the cytosolic A domains. In the unphosphorylated state, the R domain is wedged between transmembrane α helices, keeping the A1 and A2 domains apart and blocking the anion channel. Phosphorylation by PKA triggers a conformational change that moves the R domain away from the A1 and A2 interface, and, following ATP binding, the two A domains dimerize, narrowing (but not closing) the opening of the channel on the cytosolic side and widening it on the extracellular side of the membrane, allowing the

anions to flow through the channel. The ion channel pathway is lined with positively charged amino acids, making the channel specific for conductance of anions.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 11-17 Structure and function of the cystic fibrosis transmembrane regulator (CFTR). (a) Ribbon structures of the human CFTR in the dephosphorylated, ATP-free state on the left and in the phosphorylated, ATP-bound state on the right. The unresolved regions of the structure are shown as thin red lines, and unstructured regions within the R domain are shown in pink.

Phosphorylation of the R domain moves it out of the ATP-binding region and allows ATP binding and dimerization of the two ATP-binding (A) domains. This narrows the opening on the cytosolic side, with a widening of the pore on the extracellular side of the membrane. (b) A cartoon of the CFTR, with each subunit of the dimer shown in purple and peach. In the unphosphorylated, ATP-free state, the R-domain (shown as a red line) is wedged between the two subunits. Phosphorylation of R moves it out of the region between the two subunits, ATP binds and the two ATP-binding (A) sites dimerize. Dimerization narrows the channel on the cytosolic face and opens the channel on the extracellular face so that the chloride ions can travel through the pore (lined with positively charged amino acid residues) down its electrochemical gradient.

[Data from Z. Zhang et al., 2017, *Cell* **170**:483; and Z. Zhang et al., 2018, *Proc. Nat'l Acad. Sci. USA* **115**:12757.]

About two-thirds of all CF cases can be attributed to a single mutation in CFTR: deletion of Phe 508 (F508del) in the ATP-binding A1 domain. At body temperature, the mutant protein fails to fold properly and to move to the cell surface, where it normally functions. Interestingly, if cells expressing the mutant protein are incubated at room temperature, the protein folds and accumulates normally on the plasma membrane, where it functions nearly as well as the wild-type CFTR channel. Recently, a series of small molecules have been chemically synthesized that bind to this mutant CFTR protein in CF patients and stabilize the folded form at 37 °C, allowing it to traffic normally to the cell surface and partially reverse the effects of the disease. These molecules are known as *CFTR correctors*. They only promote the cell surface trafficking of a fraction of CFTR channels and thus only partially reverse the effects

of the disease. Another CFTR mutation, Gly 551 to Asp (G551D), accounts for approximately 5 percent of CF cases and results in a channel that folds normally and accumulates normally on the plasma membrane but is defective in Cl^- transport because the mutation disrupts ATP binding. In 2012, a small molecule drug called *Ivacaftor*, which increases the flow of Cl^- ions through the mutant channel, was approved by the U.S. Food and Drug Administration (FDA) to treat CF patients whose disease is caused by this mutation. This drug, known as a *CFTR potentiator*, represents one of the first successful personalized therapies resulting from a molecular understanding of the disease-causing protein. Recently, the FDA has approved treatment of F508del cystic fibrosis patients with a combination of CFTR correctors and the CFTR potentiator Ivacaftor, which increases the flow of Cl^- ions through those channels that do reach the plasma membrane. This drug combination exemplifies the use of targeted combination drug therapies resulting from detailed molecular understanding of how specific mutations cause disease. ■

KEY CONCEPTS OF SECTION 11.3

ATP-Powered Pumps and the Intracellular Ionic Environment

- Four classes of transmembrane proteins couple the energy-releasing hydrolysis of ATP with the energy-requiring transport of substances against their concentration gradients: P-, V-, and F-class pumps and ABC proteins (see Figure 11-9).
- The combined action of P-class Na^+/K^+ ATPases in the plasma membrane and homologous Ca^{2+} ATPases in the plasma membrane or SR creates the usual ionic

milieu of animal cells: high K^+, low Ca^{2+}, and low Na^+ in the cytosol; low K^+, high Ca^{2+}, and high Na^+ in the extracellular fluid.

- In P-class pumps, phosphorylation of the α (catalytic) subunit and changes in conformational states are essential for coupling ATP hydrolysis to transport of H^+, Na^+, K^+, or Ca^{2+} ions (see Figures 11-10 through 11-13).
- V- and F-class ATPases, which transport protons exclusively, are large, multisubunit complexes with a proton-conducting channel in the transmembrane domain and ATP-binding sites in the cytosolic domain.
- V-class proton pumps in animal lysosomal and endosomal membranes and plant vacuolar membranes are responsible for maintaining a lower pH inside the organelles than in the surrounding cytosol (see Figure 11-14).
- V- and F-class ATPases transport protons across the membrane using a rotary motor mechanism.
- All members of the large and diverse ABC superfamily of membrane transport proteins contain four core domains: two transmembrane domains, which form a pathway for solute movement and determine substrate specificity, and two cytosolic ATP-binding domains (see Figure 11-15).
- The two T domains of the multidrug transporter ABCB1 form a ligand-binding site in the middle of the plane of the membrane; ligands can bind directly from the cytosol or from the inner-membrane leaflet through a gap in the protein.
- The ABC superfamily includes about 50 mammalian proteins (e.g., ABCB1, ABCA1) that transport a wide array of substrates, including toxins, drugs, phospholipids, peptides, and proteins, into or out of the cell.
- Biochemical experiments directly demonstrate that ABCB4 (MDR2) possesses phospholipid flippase activity (see Figure 11-16).
- High resolution cryo-electron microscopy of a flippase from *Campylobacter jejuni* reveals a novel mechanism by which an ABC transporter flips a lipid-linked oligosaccharide across the membrane.
- CFTR, an ABC protein, is a Cl^- channel, not a pump. Channel opening is triggered by protein phosphorylation and by binding of ATP to the two A domains (see Figure 11-17).
- Mutations in CFTR cause cystic fibrosis; small molecules targeting these mutations, classified as CFTR correctors and CFTR potentiators, are used to treat cystic fibrosis patients.

11.4 Nongated Ion Channels and the Resting Membrane Potential

We have just seen that ATP-powered ion pumps transport ions *against* their concentration gradients. In addition, the plasma membrane contains channel proteins that allow the principal cellular ions (Na^+, K^+, Ca^{2+}, and Cl^-) to move through them *down* their concentration gradients. Ion pumps and channels work together to generate a difference in voltage, or electric potential, across the plasma membrane. ATP-powered ion pumps generate differences in ion concentrations across the plasma membrane, and ions move selectively through ion channels down these concentration gradients (see Figure 11-3). All channel proteins have transmembrane domains that create a hydrophilic passage or pore for ions to travel through in order to cross the membrane. The size of the pore and the composition of the amino acids that line the pore determine which ions are transported through the channel. By allowing the passage of selective ions across the membrane, channels function to generate a tightly controlled electric potential across the membrane.

In all cells, the magnitude of this electric potential is generally between 60 and 90 millivolts (mV), with the *inside* cytosolic face of the plasma membrane always *negative* with respect to the outside exoplasmic face. This value does not seem like much until we

consider that the thickness of the plasma membrane is only $\sim 3.5\,\mathrm{nm}$. Thus the voltage gradient across the plasma membrane is 0.07 V per 3.5×10^{-7} cm, or 200,000 volts per centimeter! (To appreciate what this means, consider that high-voltage transmission lines for electricity use gradients of about 200,000 volts per kilometer, 10^5-fold less!)

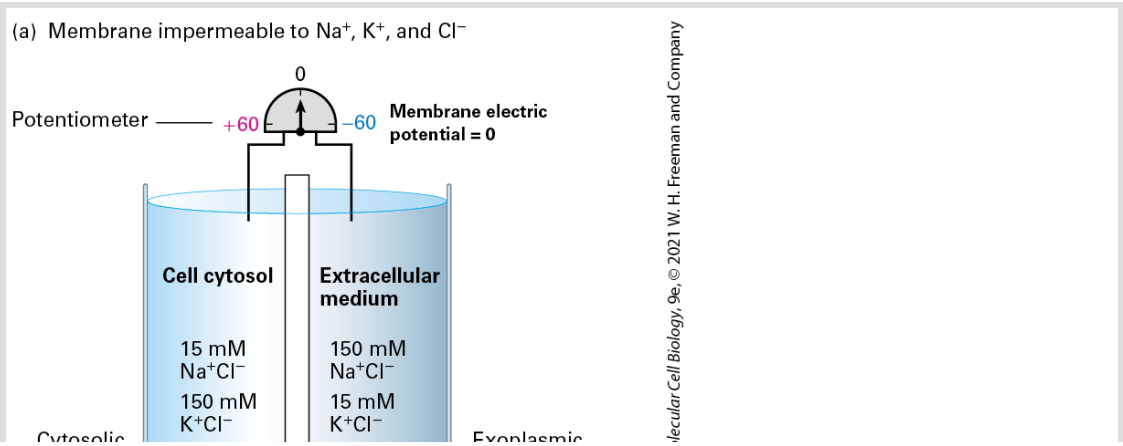
The ionic gradients and electric potential across the plasma membrane play crucial roles in many biological processes. As noted previously, and as described more fully in Chapter 15, a rise in the cytosolic Ca^{2+} concentration is an important regulatory signal, initiating contraction in muscle cells and triggering in many cells secretion of proteins, such as digestive enzymes from pancreatic cells. In many animal cells, the combined force of the Na^+ concentration gradient and the membrane electric potential drives the uptake of amino acids and other molecules against their concentration gradients by symporters and antiporters (see Figure 11-3 and Section 11.5). Furthermore, electrical signaling by neurons depends on the opening and closing of ion channels in response to changes in the membrane electric potential (see Chapter 23).

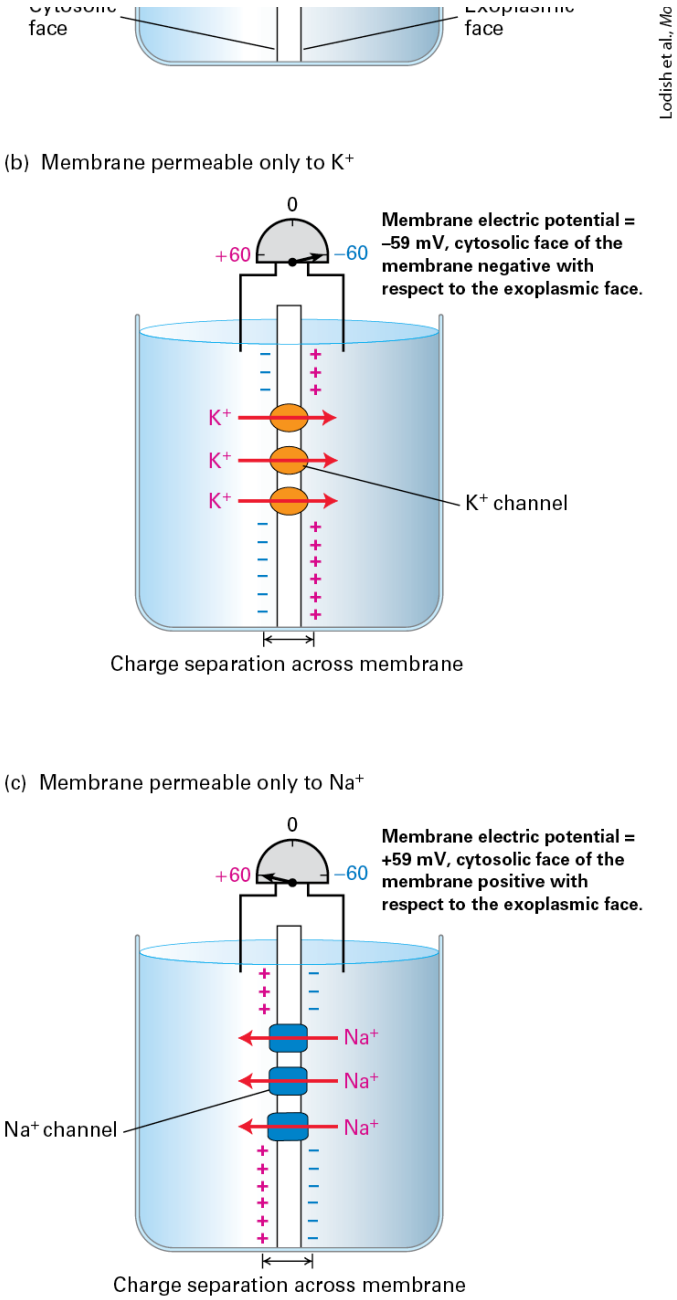
Here we discuss the origin of the membrane electric potential in resting non-neuronal cells (often called the cell's **resting membrane potential**). We also examine how ion channels mediate the selective movement of ions across a membrane and consider useful experimental techniques for characterizing the functional

properties of channel proteins. We address the role of ion channels in electrical signaling in neuronal cells in Chapter 23.

Selective Movement of Ions Creates a Transmembrane Electric Gradient

To help explain how an electric potential across the plasma membrane can arise, we first consider a set of simplified experimental systems in which a membrane separates a 15 mM NaCl/150 mM KCl solution in the cytosol (Figure 11-18a, *left*) from a 150 mM NaCl/15 mM KCl solution in the extracellular fluid (Figure 11-18a, *right*). A potentiometer (voltmeter) is connected to both solutions to measure any electric potential across the membrane. Initially, both solutions contain an equal number of positive and negative ions. If the membrane is impermeable to all ions, no ions will flow across it. Furthermore, there will be no difference in voltage, or electric potential gradient, across the membrane, as shown in Figure 11-18a.





EXPERIMENTAL FIGURE 11-18 Generation of a transmembrane electric potential (voltage) depends on the selective movement of ions across a semipermeable membrane. In this experimental system, a membrane separates a 15 mM NaCl/150 mM KCl solution in the cytosol (*left*) from a 150 mM NaCl/15 mM KCl solution in the extracellular fluid (*right*); these ion concentrations are similar to those in cytosol and blood, respectively. If the membrane separating the two solutions is impermeable to all ions (a), no ions can move across the membrane, and no electric potential is registered on the potentiometer connecting the two solutions. If the membrane is

selectively permeable only to K^+ (b) or only to Na^+ (c), then diffusion of these ions through their respective channels leads to a separation of charge across the membrane. At equilibrium, the membrane potential caused by the charge separation becomes equal to the Nernst potential E_K or E_{Na} registered on the potentiometer. In (b) and (c), the K^+ and Na^+ cations are denoted by the pink plus signs and the Cl^- anions are denoted by the light blue minus signs. See the text for further explanation.

Now suppose that the membrane contains K^+ channels that accommodate K^+ ions but exclude Na^+ and Cl^- ions (Figure 11-18b). K^+ ions then tend to move down their concentration gradient from the cytosol (*left*) to the extracellular fluid (*right*), leaving an excess of negative Cl^- ions (light blue) compared with K^+ ions (pink) on the cytosolic (*left*) side and generating an excess of positive K^+ ions compared with Cl^- ions on the extracellular (*right*) side. The excess K^+ in the extracellular (*right*) side and Cl^- on the cytosolic (*left*) side remain near the respective surfaces of the membrane because the excess positive charges on one side of the membrane are attracted to the excess negative charges on the other side. The resulting separation of charge across the membrane constitutes an electric potential, or voltage, and the cytosolic (*left*) side of the membrane has excess positive charge with respect to the extracellular (*right*) side.

As more and more K^+ ions move through channels across the membrane, the magnitude of this charge difference (i.e., voltage) increases. However, continued cytosol (*left*) to extracellular (*right*)

movement of the K^+ ions is eventually inhibited by the mutual repulsion between the excess positive (K^+) charges accumulated on the extracellular (*right*) side of the membrane and by the attraction of K^+ ions to the excess negative charges built up on the cytosolic (*left*) side. The system soon reaches an equilibrium point at which the two opposing factors that determine the movement of K^+ ions — the membrane electric potential and the ion concentration gradient — balance each other out. At equilibrium, there is no net movement of K^+ ions across the membrane. Thus this membrane, like all cellular membranes, acts as a *capacitor* — a device consisting of a thin sheet of nonconducting material (the hydrophobic interior) surrounded on both sides by electrically conducting material (the polar phospholipid head groups and the ions in the surrounding aqueous solution) that can store positive charges on one side and negative charges on the other.

If a membrane is permeable only to K^+ ions, then at equilibrium, the measured electric potential across the membrane equals the potassium equilibrium potential in volts, E_{K}. The magnitude of E_{K} is given by the *Nernst equation,* which is derived from basic principles of physical chemistry:

$$E_{\mathrm{Na}} = \frac{RT}{ZF} \ln \frac{[\mathrm{K}_{\mathrm{right}}]}{[\mathrm{K}_{\mathrm{left}}]} \tag{11-2}$$

where R (the gas constant) = 8.28 joules/(degree · mol); T (the absolute temperature in degrees Kelvin) = 293 °K at 20 °C; Z (the charge, also called the *valence*) is here equal to $+1$; F (the Faraday constant) = 96,500 coulombs/(mol · V); and $[\mathrm{K}_{\mathrm{left}}]$ and $[\mathrm{K}_{\mathrm{right}}]$ are the Ka^{+} concentrations on the left and right sides, respectively, at equilibrium. By convention, the potential is expressed as the *cytosolic* face of the membrane relative to the *exoplasmic* face, and the equation is written with the ion concentration of the extracellular solution (here the right side of the membrane) placed in the numerator and that of the cytosol in the denominator.

At 20 °C (room temperature), Equation 11-2 reduces to

$$E_{\mathrm{K}} = 0.059 \log_{10} \frac{[\mathrm{K}_{\mathrm{right}}]}{[\mathrm{K}_{\mathrm{left}}]} \tag{11-3}$$

If $[\mathrm{K}_{\mathrm{right}}]/[\mathrm{K}_{\mathrm{left}}] = 1/10$, a tenfold ratio of concentrations as in Figure 11-18b, then $E_{\mathrm{K}} = -0.059\,\mathrm{V}$ (or $-59\,\mathrm{mV}$), with the left, cytosolic side negative with respect to the right, exoplasmic side.

If the membrane is permeable only to Na^{+} ions and not to K^{+} or Cl^{-} ions, then a similar equation describes the sodium equilibrium potential E_{Na}:

$$E_{\mathrm{Na}} = 0.059 \log_{10} \frac{[\mathrm{Na}_{\mathrm{right}}]}{[\mathrm{Na}_{\mathrm{left}}]}$$

(11-4)

The *magnitude* of the membrane electric potential is the same (59 mV, for a tenfold difference in ion concentrations), but the cytosolic (left) side is now *positive* with respect to the extracellular (right) side (Figure 11-18c), so the polarity is opposite to that obtained across a membrane selectively permeable to K^+ ions.

The Resting Membrane Potential in Animal Cells Depends Largely on the Outward Flow of K^+ Ions Through Open K^+ Channels

The plasma membranes of animal cells contain many open K^+ channels but few open Na^+, Cl^-, or Ca^{2+} channels. As a result, the major ionic movement across the plasma membrane is the movement of K^+ from the *inside outward*, powered by the K^+ concentration gradient. This movement leaves an excess of *negative* charge on the cytosolic face of the plasma membrane and creates an excess of *positive* charge on the exoplasmic face, as in the experimental system shown in Figure 11-18b. This outward flow of K^+ ions is the major determinant of the inside-negative membrane potential. The channels through which the K^+ ions flow, called **resting K^+ channels**, alternate, like all channels, between an open

and a closed state (see Figure 11-2), but since their opening and closing is not affected by the membrane potential or by small signaling molecules, these channels are referred to as *nongated*. In contrast, the various gated channels in neurons and other excitable cells (see Chapter 23) open only in response to specific ligands or to changes in membrane potential.

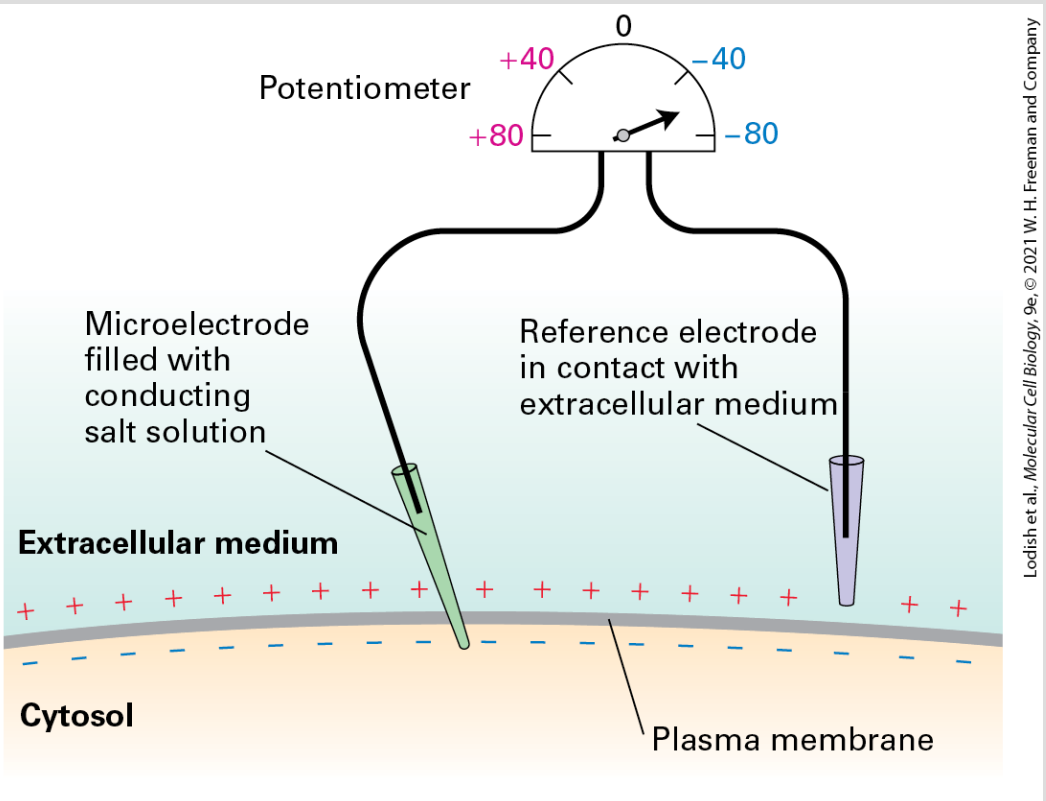
Quantitatively, the usual resting membrane potential of approximately –60 mV is relatively close to the potassium equilibrium potential, calculated from the Nernst equation and the K^+ concentrations in cells and surrounding media (see Table 11-2; also note that the resting membrane potential can vary in different cell types from –60 to –90 mV). Usually the resting membrane potential is slightly lower (less negative) than that calculated from the Nernst equation because of the presence of a few open Na^+ channels. These channels allow the net *inward* flow of Na^+ ions, making the cytosolic face of the plasma membrane more positive — that is, less negative — than predicted by the Nernst equation for K^+.

The K^+ ion would not flow inward at all if it were not for the presence of the K^+ concentration gradient. The concentration gradient that drives the flow of ions through resting K^+ channels is generated by the Na^+/K^+ ATPase described previously (see Figures 11-3 and 11-12). In the absence of this pump, or when it is inhibited,

the K^+ concentration gradient cannot be maintained, the membrane potential falls to zero, and the cell eventually dies.

Although resting K^+ channels play the dominant role in generating the electric potential across the plasma membranes of animal cells, this is not the case in bacterial, plant, and fungal cells. The inside-negative membrane potential in plant and fungal cells is generated by transport of positively charged protons (H^+) out of the cell by ATP-powered proton pumps, a process similar to what occurs in lysosomal membranes lacking Cl^- channels (see Figure 11-13a): each H^+ pumped out of the cell leaves behind a Cl^- ion, generating an inside-negative electric potential across the membrane. In aerobic bacterial cells, an inside-negative potential is generated by outward pumping of protons during electron transport, a process similar to proton pumping in mitochondrial inner membranes that will be discussed in detail in Chapter 12 (see Figure 12-19).

The electric potential across the plasma membrane of a cell can be measured with a microelectrode inserted into the cell and a reference electrode placed in the extracellular fluid. The two electrodes are connected to a potentiometer capable of measuring small potential differences (Figure 11-19). The potential across the plasma membrane of most animal cells generally does not vary with time. In contrast, neurons and muscle cells — the principal types of electrically active cells — undergo controlled changes in their membrane potential, as we discuss in Chapter 23.



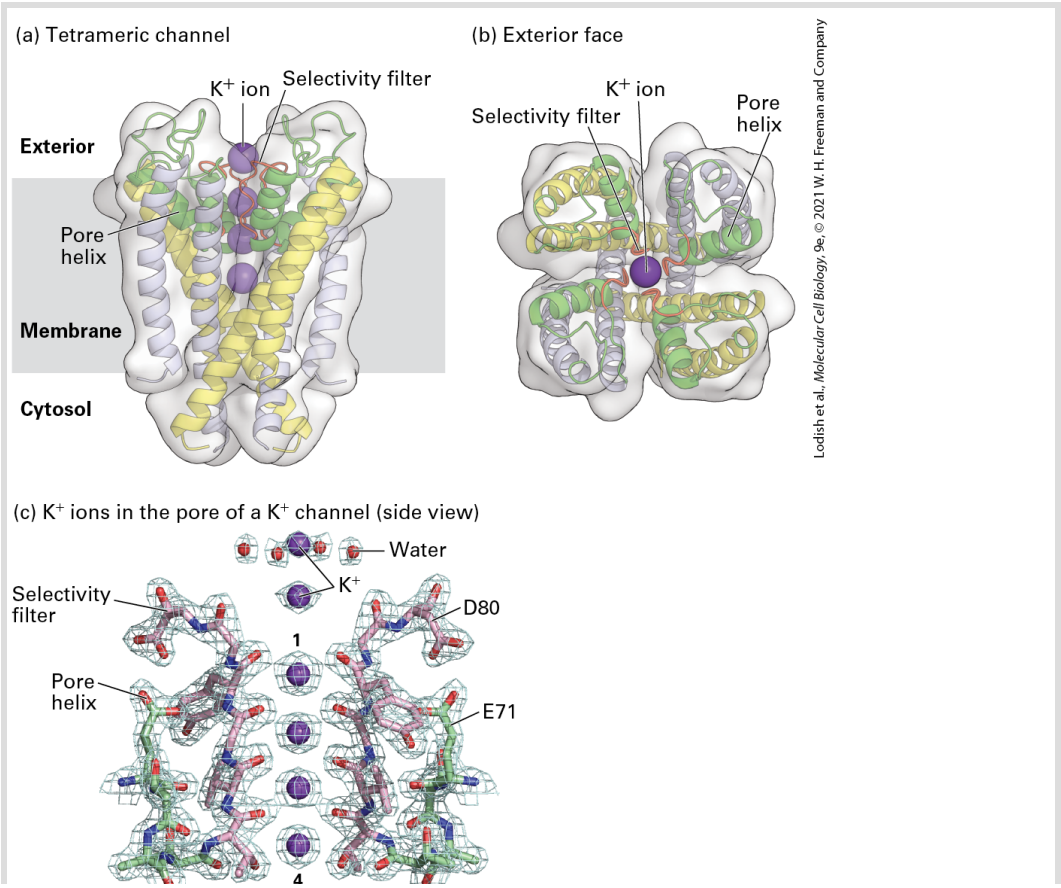
EXPERIMENTAL FIGURE 11-19 The electric potential across the plasma membrane of a live cell can be measured. A microelectrode, constructed by filling a glass tube of extremely small diameter with a conducting fluid such as a KCl solution, is inserted into a cell in such a way that the plasma membrane seals itself around the tip of the electrode. A reference electrode is placed in the extracellular medium. A potentiometer connecting the two electrodes registers the potential — in this case, –60 mV, with the cytosolic face *negative* with respect to the exoplasmic face of the membrane. A potential difference is registered only when the microelectrode is inserted into the cell; no potential is registered if the microelectrode is in the extracellular fluid.

Ion Channels Are Selective for Certain Ions by Virtue of a Molecular Selectivity Filter

All ion channels exhibit specificity for particular ions: K^+ channels allow K^+ ions, but not closely related Na^+ ions, to enter them, whereas Na^+ channels admit Na^+, but not K^+. K^+ channels can be divided into four classes on the basis of their structure and function: the tandem-pore domain K^+ channels (constitutively open, e.g., the resting K^+ channel); ligand-activated K^+ channels; inwardly rectifying K^+ channels (transport positive K^+ ions more easily in the inward direction, into the cell), and voltage-gated K^+ channels (which we discuss in Chapter 23 in the context of their role in electrical signaling in neurons). All K^+ channels possess transmembrane helices that span the lipid bilayer, and all K^+ channels can be divided into a pore-forming domain and a regulatory domain. K^+ ions are transported through the pore-forming domain, whereas the regulatory domain opens or closes the pore in response to distinct stimuli. The regulatory domain differs between classes, but the pore-forming domain is similar in all K^+ channels.

Determination of the three-dimensional structure of a bacterial K^+ channel first revealed how the exquisite ion selectivity of K^+ channels is achieved, using a mechanism that has been shown to be conserved in the many bacterial and eukaryotic K^+ channel structures that have been solved in recent years. The similarity of the structural features that are used to generate ion selectivity suggest that all K^+ channels derived from a single type of channel protein.

Like all other K^+ channels, bacterial K^+ channels are built of four identical transmembrane subunits symmetrically arranged around a central pore (Figure 11-20a, b). Each subunit contains two membrane-spanning α helices and a short P (pore) segment that partly penetrates the membrane bilayer from the exoplasmic surface. In the tetrameric K^+ channel, the eight transmembrane α helices (two from each subunit) form an inverted cone, generating a water-filled cavity called the *vestibule* in the central portion of the channel that extends halfway through the membrane toward the cytosolic side. Four extended loops that are part of the four P segments form the actual *selectivity filter* in the narrow part of the pore near the exoplasmic surface, above the vestibule.



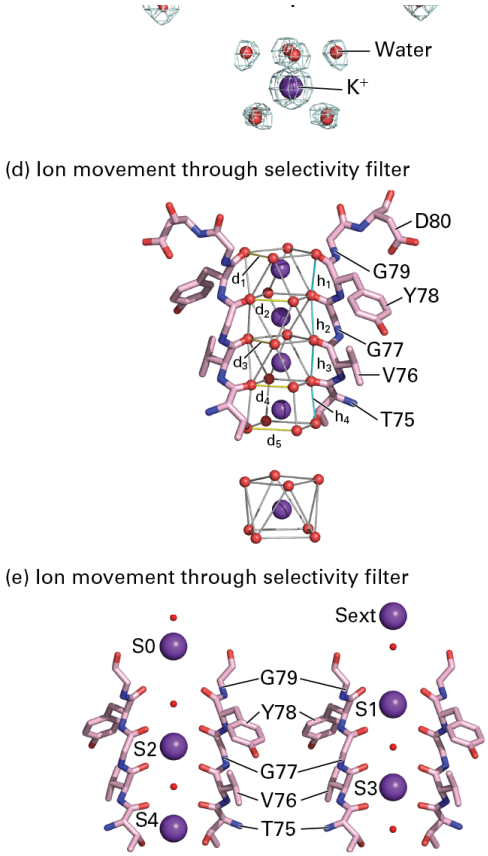


FIGURE 11-20 Structure of a resting K^+ channel from the bacterium *Streptomyces lividans.* All K^+ channels are tetramers comprising four identical subunits, each containing two conserved membrane-spanning α helices, and a shorter P, or pore, segment. The gated K^+ channels have additional α helices and are described in Chapter 23. (a) Side view of the KcsA K^+ channel, with each monomer consisting of two transmembrane α helices and a pore that is selective for K^+ ions (purple spheres). (b) Top view of the channel, revealing its fourfold symmetry, with a K^+ ion in the pore. (c) Electronic density map of two diagonal subunits, showing the orientation of the carbonyl oxygen atoms (red) in coordinating the K^+ ions. (d) Four conserved Thr-Val-Gly-Tyr-Gly residues form the selectivity filter, with dehydrated K^+ ions stabilized by the negatively charged carbonyl groups so that they can translocate through the pore in single file. (e) The K^+ ions are in two conformations, either in S2 and S4, or in S1 and S3 during translocation, with water molecules (red spheres) in the vacant positions.

[Data from Y. Zhou et al., 2001, *Nature* **414**:43, PDB ID 1k4c; D. Naranjo, 2016, *J. Gen. Physiol.* **148**:277; and Q. Kuang et al., 2015, *Cell. Mol. Life Sci.* **72**:3677.]

Several related pieces of evidence support the role of P segments in ion selection. First, the amino acid sequences of the P segments in all known K^+ channels are highly homologous and are different from those in other ion channels. Second, mutations of certain amino acids in this segment alter the ability of a K^+ channel to distinguish Na^+ from K^+. Finally, replacing the P segment of a bacterial K^+ channel with the homologous segment from a mammalian K^+ channel yields a chimeric protein that exhibits normal selectivity for K^+ over other ions. Thus all K^+ channels are thought to use the same mechanism to distinguish K^+ from other ions.

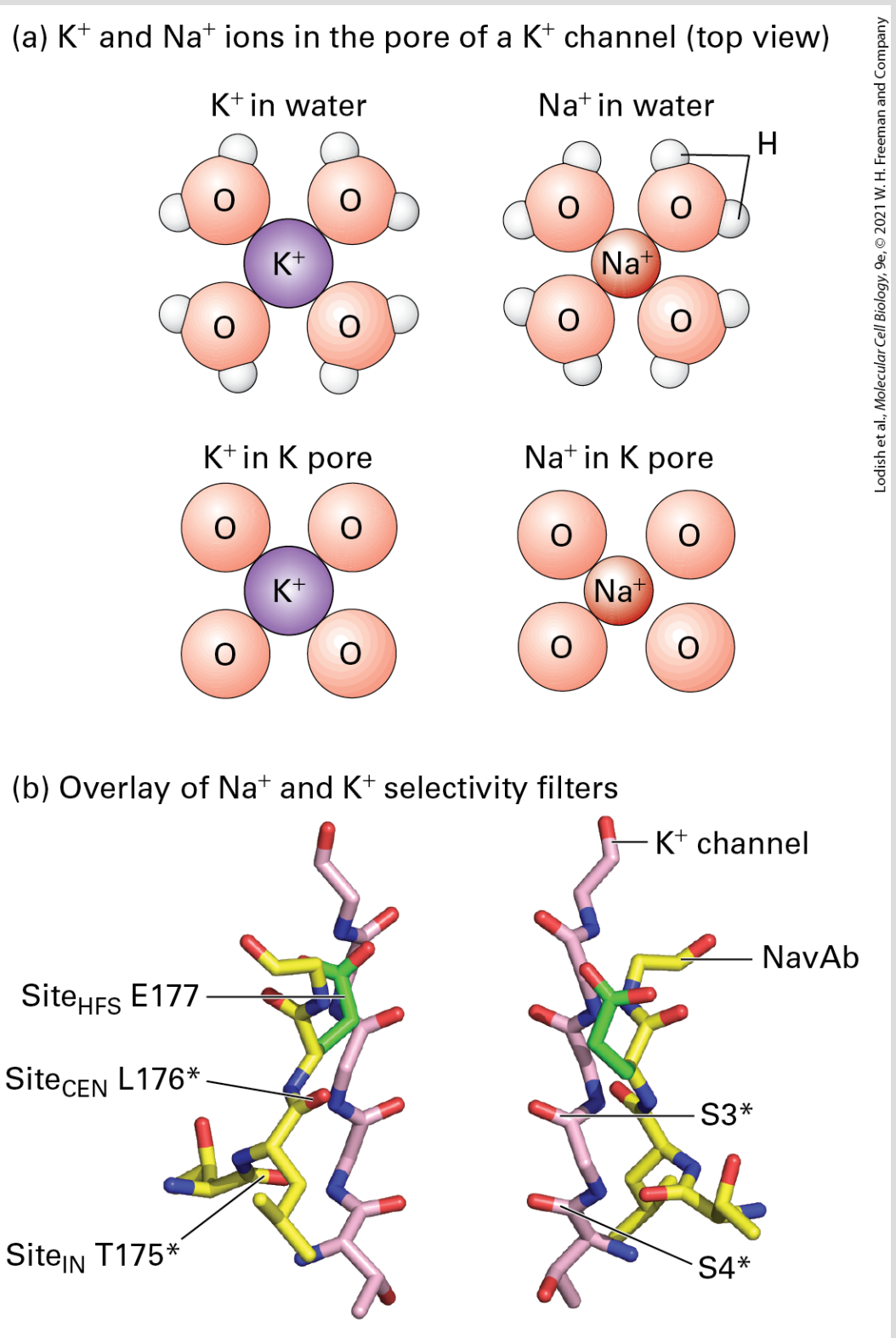
The selectivity filter in K^+ channels is able to select K^+ over Na^+ due mainly to the backbone carbonyl oxygen atoms on residues located in a Thr-Val-Gly-Tyr-Gly sequence that is found in an analogous position in the P segment in every known K^+ channel (note that in some channels, the Tyr residue is a Phe). As a K^+ ion enters the selectivity filter, it passes into the narrow space between the filter sequences of the four P-segments contributed by the four adjacent subunits. There the K^+ ion loses its eight waters of hydration, but is bound in the same geometry to eight backbone carbonyl oxygen atoms of the Thr-Val-Gly-Tyr-Gly sequence, two from the extended loop in each of the four P segments lining the channel (Figure 11-20c, d). These carbonyl oxygen atoms form four

evenly spaced K^+ ion binding sites which can bind to four K^+ ions. These protein oxygen atoms mimic the oxygen in water, and as such, the activation energy required for passage of K^+ ions into the channel from an aqueous solution is low.

The ease of passage into the channel in turn raises the question of how K^+ ions are released from their binding sites in the P-segment to allow rapid transport. Analysis of ion occupancy revealed that only two K^+ ions are present in the P segment at a time, either in the S1 and S3 conformation, or in the S2 and S4 conformation (Figure 11-20e), with two water molecules occupying the other, vacant sites. The finding that K^+ ions move through the channel at near the diffusion limit suggests that the binding of K^+ to the sites in P segment is weak, with repulsion of the two K^+ ions in the P segment promoting their transport through the pore.

Na^+ ions are smaller than K^+ ions. How, then, can a channel protein exclude smaller Na^+, yet allow passage of larger K^+? The answer to this question is simple and elegant. A dehydrated Na^+ ion is too small to bind to all eight carbonyl oxygen atoms that line the selectivity filter with the same geometry as a Na^+ ion surrounded by its normal eight water molecules in aqueous solution. As a result, Na^+ ions "prefer" to remain in water rather than enter the selectivity filter, and thus the change in free energy for entry of Na^+ ions into the channel is relatively high (Figure 11-21a, *right*). This difference in free energies favors passage of K^+ ions through the

channel over Na^+ ions by a factor of 1000. Like Na^+, the dehydrated Ca^{2+} ion is smaller than the dehydrated K^+ ion and cannot interact properly with the oxygen atoms in the selectivity filter. Furthermore, because a Ca^{2+} ion has two positive charges and binds water oxygen atoms more tightly than does a single positive Na^+ or K^+ ion, more energy is required to strip the waters of hydration from Ca^{2+} than from K^+ or Na^+.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 11-21 Mechanism of ion selectivity and transport in resting K^+ channels. (a) Schematic diagrams of K^+ and Na^+ ions hydrated in solution

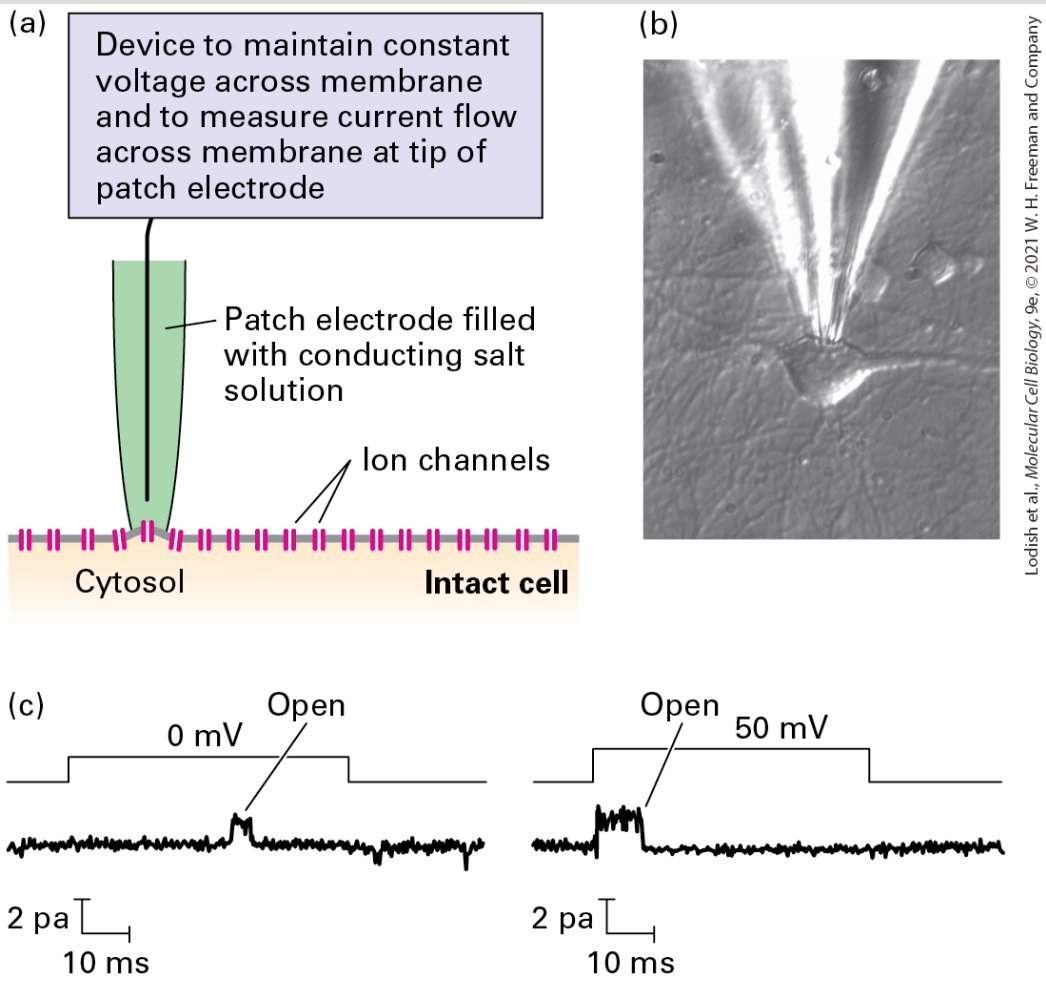
and in the pore of a K^+ channel. As K^+ ions pass through the selectivity filter, they lose their bound water molecules and become bound instead to eight backbone carbonyl oxygen atoms (four of which are shown) that are part of the conserved amino acid sequence in the channel-lining selectivity filter loop of each P segment. The smaller Na^+ ions, with their tighter shell of water molecules, cannot perfectly bind to the channel oxygen atoms and therefore pass through the channel only rarely. (b) Overlay of the selectivity filters in the KcsA K^+ channel (pink) and the NavAb Na^+ channel (yellow). The selectivity pore in the NavAb Na^+ channel is wider and shorter than the selectivity filter in the KcsA K^+ channel. See C. Armstrong, 1998, *Science* **280**:56.

[Data from R. F. Stephens et al., 2015, *Front. Physiol.* **6**:153.]

The first three-dimensional crystal structure of a bacterial voltage-gated Na^+ channel was determined in late 2011. This structure showed that the channel contains an ion selectivity filter that is slightly larger than the filter in K^+ channels, and that it transports hydrated rather than dehydrated Na^+ ions (Figure 11-21b). The structures of multiple eukaryotic voltage-gated Na^+ channels have been solved in recent years. We will examine these structures and their function in electrical signaling in neurons in more detail in Chapter 23.

Patch Clamps Permit Measurement of Ion Movements Through Single Channels

Once it was realized that there are only one or a few ion channels per square micrometer of plasma membrane in most cells, it became possible to record ion movements through single ion channels, using a technique known as **patch clamping**. This technique measures the rates at which these channels open and close and conduct specific ions. As illustrated in Figure 11-22a and b, a tiny glass pipette is tightly sealed, using suction, to the surface of a cell; the segment of the plasma membrane within the tip of the pipette will contain only one or a few ion channels. The only current that crosses through the patch of membrane must pass through these channels as ions flow. An electrical recording device detects this ion flow, measured as electric current, through the channels; this ion flow usually occurs in small bursts when a channel is open (Figure 11-22c). The electrical device also "clamps" or locks the electric potential across the membrane at a predetermined value (hence the term *patch clamping*). The inward or outward movement of ions across a patch of membrane is quantified from the amount of electric current needed to maintain the membrane potential at a particular clamped value (see Figure 11-22a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 11-22 Current flow through individual ion channels can be measured by patch clamping. (a) Basic experimental arrangement for measuring current flow through individual ion channels in the plasma membrane of a live cell. The patch electrode, filled with a current-conducting saline solution, is applied, with a slight suction, to the plasma membrane. The 0.5-μm-diameter tip covers a region that contains only one or a few ion channels. A recording device measures current flow only through the channel or channels in that region. (b) Photomicrograph of the cell body of a cultured neuron and the tip of a patch pipette touching the plasma membrane. (c) Current traces from patch-clamp experiments measuring activity of a single voltage-gated K^+ channel. Patches of membrane containing a voltage-gated K^+ channel were clamped at a negative membrane potential and voltage steps were applied while measuring the changes in current. The top row shows the voltage applied to the membrane; first the membrane is held at a negative

potential and then voltage steps are applied to hold ("clamp") the membrane potential at 0 mV (A) or $+50$ mV (B). The voltage traces below show K^+ channel opening by an increase in the current trace and the channel closing by the rapid decrease in the current trace. The amplitude of the current trace is greater and the channel opening is quicker when the membrane is more depolarized $(+50\ \mathrm{mV})$ compared to when it is clamped at 0 mV.

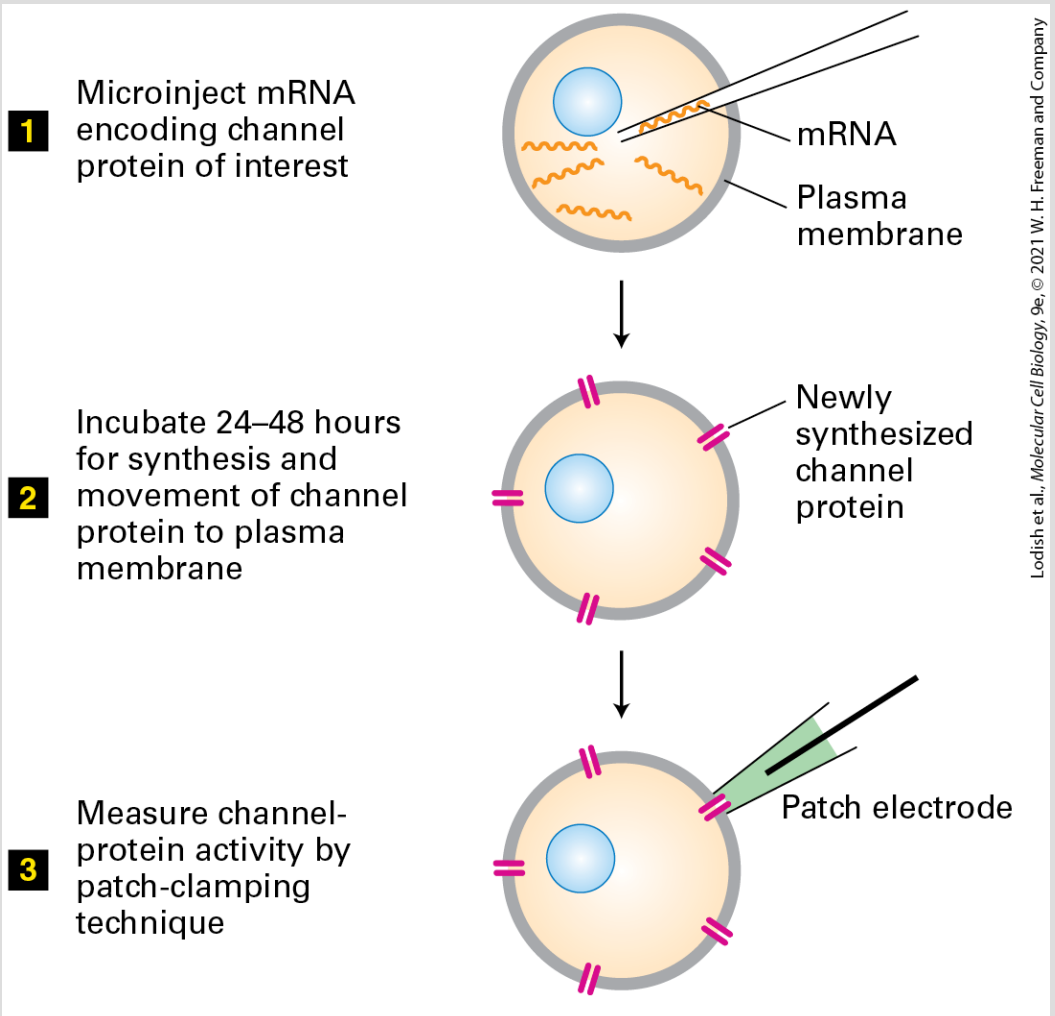
[Part (c) Data from W. N. Zagotta, M. S. Brainard, and R. W. Aldrich, 1988, *J. Neurosci.* **8**:4765–4779.]

Novel Ion Channels Can Be Characterized by a Combination of Oocyte Expression and Patch Clamping

The current traces in Figure 11-22c illustrate the use of the single-channel patch clamp technique to study the properties of voltage-gated K^+ channels. These types of K^+ channels are critical for multiple physiological functions including regulating neuron excitability and cardiac muscle activity. Here a portion of membrane containing a single potassium channel is clamped at a negative membrane potential (a voltage at which these K^+ channels are closed), and then a voltage that depolarizes the membrane — that is, reverse the charge across the membrane. This voltage is applied in two steps; the first raises the membrane potential from the initial potential to 0 mV and the second raises it to 50 mV. Each depolarization results in the channel opening for $\sim$7 ms to allow K^+

ions to flow out of the cell. The efflux of K^+ ions through a single channel can be visualized by the transient rise of the current traces in Figure 11-22c. This experiment illustrates the voltage-dependence of these channels; the channel opens more rapidly and shows greater K^+ current efflux when the depolarization is greater (step to +50 mV) compared to when the depolarization is weaker (step to 0 mV).

Cloning of human-disease-causing genes and sequencing of the human genome have identified many genes encoding putative channel proteins, including 79 putative K^+ channels and 23 putative Na^+ channels. One way of characterizing the function of these proteins is to transcribe a cloned cDNA in a cell-free system to produce the corresponding mRNA. Injecting this mRNA into frog oocytes and taking patch-clamp measurements of the newly synthesized channel protein can often reveal its function (Figure 11-23). This experimental approach is especially useful because frog oocytes normally do not express any channel proteins on their plasma membranes, so only the channel under study is present in the membrane. In addition, because of the large size of frog oocytes, patch-clamping studies are technically easier to perform on them than on smaller cells.



EXPERIMENTAL FIGURE 11-23 The oocyte expression assay is useful in comparing the function of normal and mutant forms of a channel protein. An oocyte from the follicle of a frog is first treated with collagenase to remove the surrounding follicle cells, leaving a denuded oocyte, which is then microinjected with mRNA encoding the channel protein under study.

[Data from T. P. Smith, 1988, *Trends Neurosci*. **11**:250.]

KEY CONCEPTS OF SECTION 11.4

Nongated Ion Channels and the Resting Membrane Potential

- An inside-negative electric potential (voltage) of between 60 and 90 mV exists across the plasma membrane of all cells.
- The electric potential generated by the selective flow of ions across a membrane can be calculated using the Nernst equation (see Equation 11-2).
- The resting membrane potential in animal cells is the result of the combined action of the ATP-powered Na^+/K^+ pump, which establishes Na^+ and K^+ concentration gradients across the membrane, and resting K^+ channels, which permit selective movement of K^+ ions back down their concentration gradient to the external medium (see Figure 11-3).
- Unlike the more common gated ion channels, which open only in response to various signals, the nongated resting K^+ channels are usually open.
- In plants and fungi, the membrane potential is maintained by the ATP-driven pumping of protons from the cytosol to the exterior of the cell.
- K^+ channels are assembled from four identical subunits, each of which has at least two conserved membrane-spanning α helices and a nonhelical P segment that lines the ion pore and forms the selectivity filter (see Figure 11-20).
- The ion specificity of K^+ channels is due mainly to binding of the dehydrated K^+ ion with eight carbonyl oxygen atoms of specific amino acids in the P segments, which lowers the activation energy for the passage of K^+ compared with Na^+ or other ions (see Figure 11-21).
- The pore in Na^+ channels is wider than in K^+ channels and transports partially hydrated Na^+ ions.
- Patch-clamping techniques, which permit measurement of ion movements through single channels, are used to determine the ion conductivity of a channel and the effect of various reagents on its activity (see Figure 11-22).
- Recombinant DNA techniques and patch clamping allow the expression and functional characterization of channel proteins in frog oocytes (see Figure 11-23).

11.5 Cotransport by Symporters and Antiporters

In previous sections, we saw how ATP-powered pumps generate ion concentration gradients across cellular membranes and how K^+ channels use the K^+ concentration gradient to establish an electric potential across the plasma membrane. In this section, we see how cotransporters use the energy stored in electric potentials and concentration gradients of Na^+ or H^+ ions to power the uphill movement of another substance, which may be a small organic molecule such as glucose, or an amino acid, or a different ion. An important feature of such **cotransport** is that neither substance can move alone; movement of both substances together is obligatory, or *coupled*.

Cotransporters share features with uniporters such as the GLUT proteins. The two types of transporters exhibit certain structural similarities, operate at equivalent rates, and undergo cyclical conformational changes during transport of their substrates. They differ in that uniporters can only accelerate thermodynamically favorable transport down a concentration gradient, whereas cotransporters can harness the energy released when one substance moves down its concentration gradient to drive the movement of another substance against its concentration gradient.

When the transported molecule and cotransported ion move in the same direction, the process is called *symport*; when they move in opposite directions, the process is called *antiport* (see Figure 11-2). Some cotransporters transport only positive ions (cations), while others transport only negative ions (anions). Yet other cotransporters mediate movement of both cations and anions together. Cotransporters are present in all organisms, including bacteria, plants, and animals. In this section, we describe the operation and function of several physiologically important symporters and antiporters.

Na^+ Entry into Mammalian Cells Is Thermodynamically Favored

Mammalian cells express many types of Na^+-linked symporters. The human genome encodes literally hundreds of different types of transporters that use the energy stored in the Na^+ concentration gradient and in the inside-negative electric potential across the plasma membrane to transport a wide variety of molecules into cells against their concentration gradients. To see how such transporters allow cells to accumulate substrates against a considerable concentration gradient, we first need to calculate the change in free energy (ΔG) that occurs during Na^+ entry. As mentioned earlier, two forces govern the movement of ions across a selectively permeable membrane: the voltage across the membrane and the ion

concentration gradient across the membrane. The sum of these forces constitutes the electrochemical gradient. To calculate the free-energy change, ΔG, corresponding to the transport of any ion across a membrane, we need to consider the independent contributions from each of these forces to the electrochemical gradient.

For example, when Na^+ moves from the outside to the inside of a cell, the free-energy change generated from the Na^+ concentration gradient is given by

$$\Delta G_c = RT \ln \frac{[Na_{in}]}{[Na_{out}]} \tag{11-5}$$

At the concentrations of Na_{in} and Na_{out} shown in Figure 11-24, which are typical for many mammalian cells, ΔG_c, the change in free energy due to the concentration gradient, is -6.067 Joules for transport of 1 mole of Na^+ ions from the outside to the inside of a cell, assuming there is no membrane electric potential. Note that the free energy is negative, indicating spontaneous movement of Na^+ into the cell down its concentration gradient.

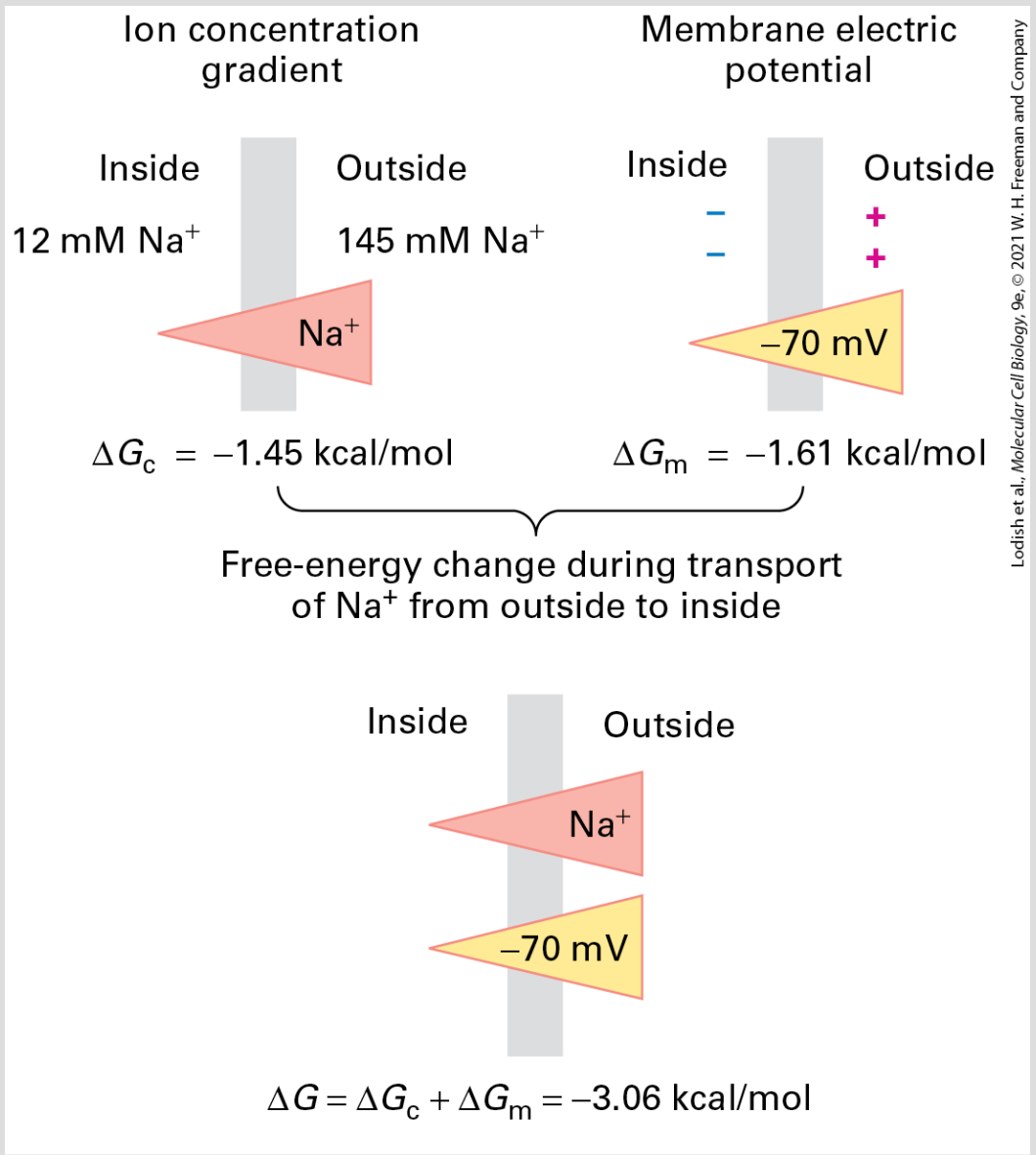


FIGURE 11-24 Transmembrane forces acting on Na^+ ions. As with all ions, the movement of Na^+ ions across the plasma membrane is governed by the sum of two separate forces: the ion concentration gradient and the membrane electric potential. At the internal and external Na^+ concentrations typical of mammalian cells, these forces usually act in the same direction, making the inward movement of Na^+ ions energetically favorable.

The free-energy change generated from the membrane electric potential is given by

$$\Delta G_{\mathrm{m}} = FE$$

(11-6)

where F is the Faraday constant [$= 96{,}500\,\mathrm{Joules/(mol \cdot V)}$]; and E is the membrane electric potential. If $E = -70\,\mathrm{mV}$, then ΔG_{m}, the free-energy change due to the membrane potential, is –6755 Joules for transport of 1 mole of $\mathrm{Na^+}$ ions from the outside to the inside of a cell, assuming there is no $\mathrm{Na^+}$ concentration gradient. Since both forces do in fact act on $\mathrm{Na^+}$ ions, the total ΔG is the sum of the two partial values:

$$\Delta G = \Delta G_{\mathrm{c}} + \Delta G_{\mathrm{m}} = (-6{,}067) + (-6{,}755) = -12{,}822\ \mathrm{Joules/mol}$$

In this example, the $\mathrm{Na^+}$ concentration gradient and the membrane electric potential contribute almost equally to the total ΔG for transport of $\mathrm{Na^+}$ ions. Since ΔG is less than 0, the inward movement of $\mathrm{Na^+}$ ions is thermodynamically favored. As we will see next, the inward movement of $\mathrm{Na^+}$ is used to power the uphill movement of other ions and several types of small molecules into or out of animal cells. The rapid, energetically favorable movement of $\mathrm{Na^+}$ ions through gated $\mathrm{Na^+}$ channels is also critical in generating action potentials in neurons and muscle cells, as discussed in Chapter 23.

Na^+-Linked Symporters Enable Animal Cells to Import Sugars Including Glucose and Galactose as Well as Amino Acids Against High Concentration Gradients

Most body cells import glucose from the blood *down* a concentration gradient of glucose, using GLUT proteins to facilitate this transport. However, certain cells need to import glucose from extracellular fluids against a very large concentration gradient. For example, the cells lining the small intestine and the kidney tubules need to import glucose from digestive products or urine against a very large concentration gradient because glucose concentration is higher inside the cell. Such cells use a *two-Na^+/one-glucose symporter*, a protein that couples the import of one glucose molecule to the import of two Na^+ ions:

$$2\,\mathrm{Na^+}_{\mathrm{out}} + \mathrm{glucose}_{\mathrm{out}} \rightleftharpoons 2\,\mathrm{Na^+}_{\mathrm{in}} + \mathrm{glucose}_{\mathrm{in}}$$

Quantitatively, the free-energy change for the symport of two Na^+ ions and one glucose molecule can be written

$$\Delta G = RT\,\ln\frac{[\mathrm{glucose}_{\mathrm{in}}]}{[\mathrm{glucose}_{\mathrm{out}}]} + 2RT\ln\frac{[\mathrm{Na^+}_{\mathrm{in}}]}{[\mathrm{Na^+}_{\mathrm{out}}]} + 2FE \tag{11-7}$$

Thus the ΔG for the overall reaction is the sum of the free-energy changes generated by the glucose concentration gradient (1 molecule transported), the Na^+ concentration gradient (2 Na^+ ions transported), and the membrane potential (2 Na^+ ions transported). As illustrated in Figure 11-25, the movement of Na^+ ions into a mammalian cell down their electrochemical gradient has a free-energy change, ΔG, of about $-12{,}552$ Joules per mole of Na^+ transported. Thus the ΔG for the transport of two moles of Na^+ inward would be twice this amount, or about $-25{,}104$ Joules. This negative free-energy change for sodium import is coupled to the uphill transport of glucose, a process with a positive ΔG. We can calculate the glucose concentration gradient (inside greater than outside) that can be established by the action of this Na^+-powered symporter by realizing that at equilibrium for Na^+-linked glucose import, $\Delta G = 0$. By substituting the values for sodium import into Equation 11-7 and setting $\Delta G = 0$, we see that

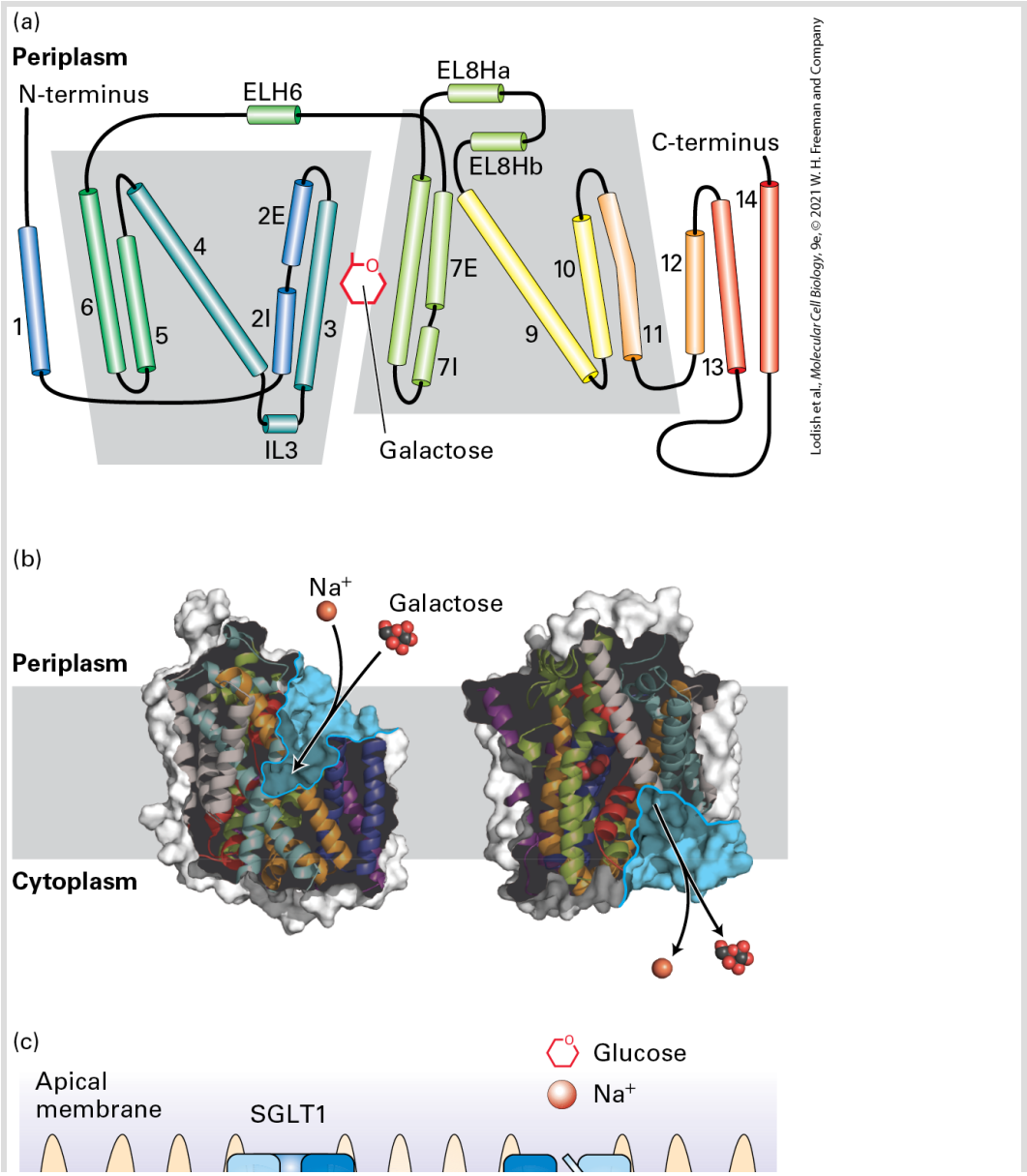
$$0 = RT \ln \frac{[\text{glucose}_{\text{in}}]}{[\text{glucose}_{\text{out}}]} - 25{,}104 \text{ Joules}$$

and we can calculate that at equilibrium, the ratio of $\text{glucose}_{\text{in}}/\text{glucose}_{\text{out}} = \sim 30{,}000$. Thus the inward flow of two moles of Na^+ can generate an intracellular glucose concentration that is $\sim 30{,}000$ times greater than the exterior concentration. If only one Na^+ ion were imported (ΔG of approximately $-12{,}552$ Joules/mol)

per glucose molecule, then the available energy could generate a glucose concentration gradient (inside/outside) of only about 170-fold. Thus by coupling the transport of two Na^+ ions to the transport of one glucose molecule, the two-Na^+/one-glucose symporter permits cells to accumulate a very high concentration of glucose relative to the external concentration. This means that glucose that is present even at very low concentrations in the lumen of the intestine or in the kidney tubules can be efficiently transported into the cells of the lining and not lost from the body.

The high-resolution crystal structure of a bacterial Na^+ galactose symporter was solved in 2008; it provides a model for the function of Na^+ sugar symporters. In this particular symporter (and in select other symporters), the stoichiometry of Na^+ ions to sugar molecules is 1:1 rather than 2:1. As shown in Figure 11-25, the Na^+ galactose symporter contains 14 transmembrane (TM) α helices, with TM domains 2-6 and 7-11 as inverted repeat structures within the membrane. Both N- and C-termini extend into the periplasm, the gel-like matrix in the space between the inner cytosolic membrane and the bacterial outer membrane. The galactose and Na^+ ions interact with specific side chains of the TM α helices that are present in their binding sites. Substrates enter from the periplasm through an extracellular cavity in the transporter. Upon substrate binding, the extracellular cavity closes and a cavity on the cytosolic side opens, releasing Na^+ and galactose into the cytoplasm. All substrates must bind to their sites on the extracellular domain

before the protein undergoes the change that converts the substrate-binding sites from the outward- to the inward-facing conformation; this ensures that inward transport of galactose and Na^+ ions are coupled. Figure 11-25c shows a model comparing the alternating access model of the Na^+ glucose symporter with a glucose uniporter, which does not require a cotransported ion (compare with Figure 11-5).



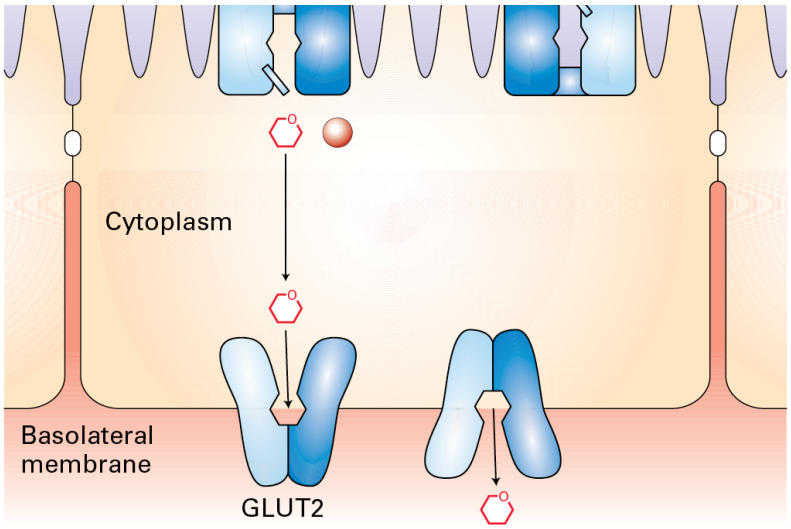


FIGURE 11-25 Structure and mechanism of transport by the Na^+ galactose symporter vSGLT from *Vibrio parahaemolyticus*. (a) Topology of vSGLT showing TM domains in rainbow colors from N-terminus (blue) to C-terminus (red). TMs 2-6 and TMs 7-11 form inverted repeats, which are highlighted by the gray trapezoids. (b) Outward-facing model through the membrane plane revealing the extracellular cavity in blue highlighting, to which Na^+ (red sphere) and galactose (red and black spheres) bind, triggering a conformational change to an inward-facing structure on the right, with the intracellular cavity (blue highlighting), releasing Na^+ and galactose into the cytoplasm. The conformational change involves structural rearrangements of TM helices. (c) A model of the pore of a Na^+-glucose symporter (SGLT1) and a glucose uniporter (GLUT2), which function together to transport glucose across the intestinal epithelium (see Figure 11-30). In the Na^+-glucose symporter, the rotation of TM helices underlies an alternating access mechanism in which the pore alternates between being closed (gated) on the extracellular and intracellular sides of the membrane; in the glucose uniporter, rotation of the two domains of the transporter against each other alternately exposes the substrate-binding site to the intracellular and extracellular sides of the membrane (see Figure 11-5).

[Data from S. Faham, 2008, *Science* **321**:810–814; and N. K. Karpowich and D. N. Weng, 2008, *Science* **321**:781–782.]

There are two human Na^+/glucose symporters. SGLT1 is found in the absorptive cells lining the small intestine as well as in the epithelial cells lining part of the kidney tubules. SGLT2 is found only in kidney tubules, where, together with SGLT1, it resorbs glucose into the blood from the forming urine. Inhibition of SGLT2 leads to excretion of glucose in the urine and a reduction in blood glucose levels; therefore, SGLT2 inhibitors have potential use in the treatment of type II diabetes. Indeed, several drug candidates that selectively inhibit SGLT2 and not SGLT1 have been developed or are currently undergoing clinical trials, including several that have been approved by the FDA for use in the United States and Canada.

Cells use Na^+-powered symporters to transport substances other than glucose into the cell against high concentration gradients. For example, several types of Na^+/amino acid symporters allow cells to import many amino acids. As another example, Na^+/neurotransmitter symporters couple the import of Na^+ to the reuptake and recycling of neurotransmitters, and they are the targets of many therapeutic drugs, including many antidepressants. They are also the targets of several drugs of abuse, including cocaine and amphetamines. ■

A Bacterial Na^+/Amino Acid Symporter Reveals How Symport

Works

Until recently, no three-dimensional structure had been determined for any mammalian Na^+-linked symporter, but the structures of several homologous bacterial Na^+/substrate symporters, such as the bacterial Na^+-galactose symporter, have provided considerable insights into the mechanisms of symporter function. The structure of the bacterial two-Na^+/one-leucine symporter has served as a particularly useful model for understanding how antidepressants work, since the bacterial symporter is closely related to the human neurotransmitter transporters that are the targets of many antidepressant drugs.

As shown in Figure 11-26a, the bacterial two-Na^+/one-leucine symporter consists of 12 membrane-spanning α helices. Two of the helices (numbers 1 and 6) have nonhelical segments in the middle of the membrane that form part of the leucine-binding site. The amino acid residues involved in binding the leucine and the two Na^+ ions are located in the middle of the membrane-spanning segment and are close together in three-dimensional space. This proximity suggests that the coupling of amino acid and ion transport in these transporters is the consequence of direct or nearly direct physical interactions of the substrates. Indeed, one of the Na^+ ions is bound to the carboxyl group of the transported leucine (Figure 11-26b). Thus neither substance can bind to the transporter without the other. Each of the two Na^+ ions is bound to six oxygen atoms in the

transporter. One of the two Na^+ ions, for example, is bound to carbonyl oxygens of several transporter amino acids as well as to carbonyl oxygens and the hydroxyl oxygen of one threonine. Equally importantly, there are no water molecules surrounding either of the bound Na^+ atoms, as is the case for K^+ ions in potassium channels (see Figure 11-20). Thus as the Na^+ ions lose their water of hydration in binding to the transporter, they bind to six oxygen atoms with a similar geometry. This reduces the energy change required for the binding of Na^+ ions and prevents other ions, such as K^+, from binding in place of Na^+.

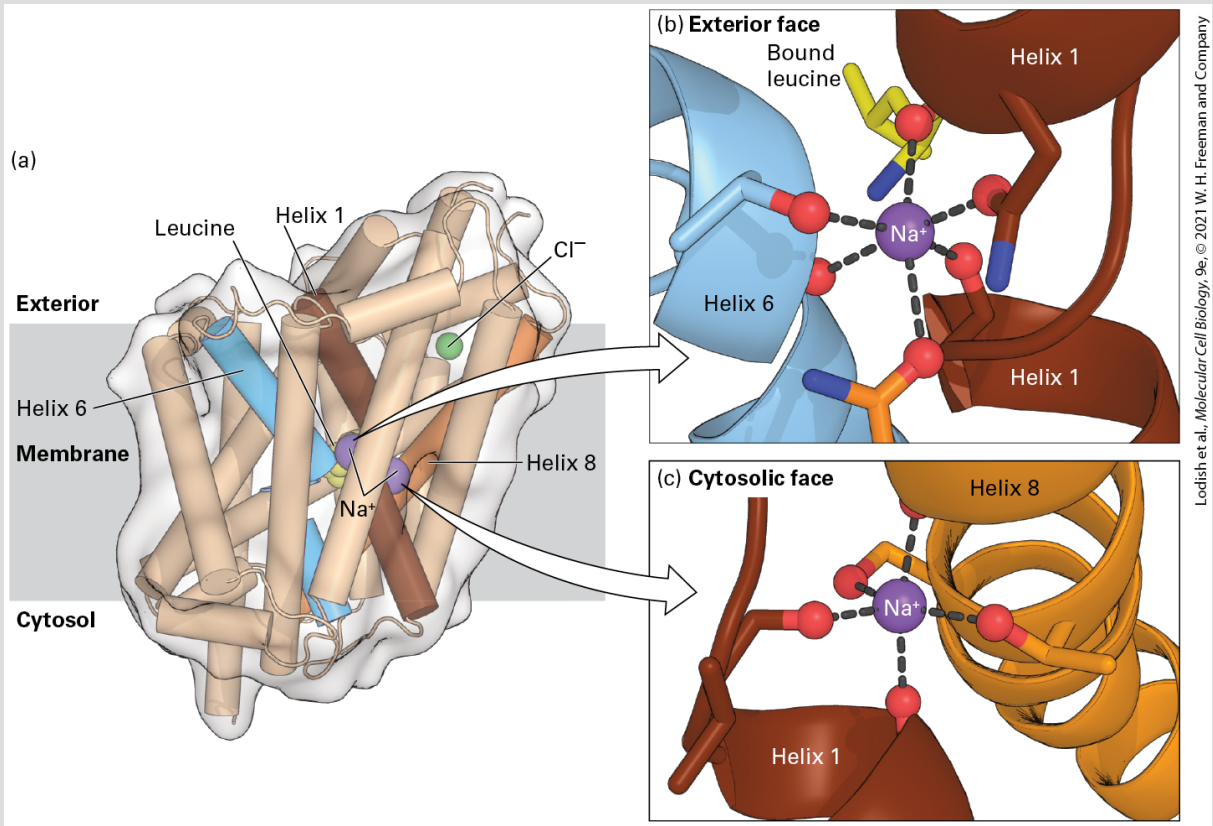


FIGURE 11-26 Three-dimensional structure of the two-Na^+/one-leucine symporter from the bacterium *Aquifex aeolicus*. (a) The bound L-leucine, two Na^+ ions, and a Cl^- ion are

shown in yellow, purple, and green, respectively. The three membrane-spanning α helices that bind the Na^+ or the leucine are colored brown, blue, and orange. (b, c) Binding of the two Na^+ ions to carbonyl main-chain or carboxyl side-chain oxygen atoms (red) that are part of helices 1 (brown), 6 (blue), or 8 (orange). It is important that one of the Na^+ ions is also bound to the carboxyl group of the transported leucine (b).

[Data from A. Yamashita et al., 2005, *Nature* **437**:215, PDB ID 2a65.]

One striking feature of the structure depicted in Figure 11-26 is that the bound Na^+ ions and leucine are *occluded* — that is, they cannot diffuse out of the protein to either the surrounding extracellular or cytoplasmic media. This structure represents an intermediate in the transport process in which the transporter appears to be changing from a conformation with an exoplasmic-facing to one with a cytosolic-facing binding site.

A Na^+-Linked Ca^{2+} Antiporter Regulates the Strength of Cardiac Muscle Contraction

In all muscle cells, a rise in the cytosolic Ca^{2+} concentration triggers contraction. Thus when muscle is resting, the cytosolic Ca^{2+} concentration must be low. In cardiac muscle cells, a *three-Na^+/one-Ca^{2+} antiporter*, rather than the plasma-membrane Ca^{2+} ATPase discussed earlier, plays the principal role in maintaining a low

concentration of Ca^{2+} in the cytosol. The transport reaction mediated by this *cation antiporter* can be written as

$$3\,Na^{+}{}_{out} + Ca^{2+}{}_{in} \rightleftharpoons 3\,Na^{+}{}_{in} + Ca^{2+}_{out}$$

Note that the inward movement of three Na^{+} ions is required to power the export of one Ca^{2+} ion from the cytosol, which has a [Ca^{2+}] of $\sim 2 \times 10^{-7}$ M, to the extracellular medium, which has a [Ca^{2+}] of $\sim 2 \times 10^{-3}$ M — a concentration gradient of some 10,000-fold (higher on the outside). By lowering cytosolic Ca^{2+}, operation of the Na^{+}/Ca^{2+} antiporter reduces the strength of heart muscle contraction.

The Na^{+}/K^{+} ATPase in the plasma membrane of cardiac muscle cells, as in other body cells, creates the Na^{+} concentration gradient necessary for export of Ca^{2+} by the Na^{+}-linked Ca^{2+} antiporter. As mentioned earlier, inhibition of the Na^{+}/K^{+} ATPase by the drugs ouabain and digoxin lowers the cytosolic K^{+} concentration and, more importantly here, simultaneously increases cytosolic Na^{+}. The resulting reduced Na^{+} electrochemical gradient across the membrane causes the Na^{+}-linked Ca^{2+} antiporter to function less efficiently. As a result, fewer Ca^{2+} ions are exported, and the cytosolic Ca^{2+} concentration increases, causing the muscle to contract more strongly. Because of their ability to increase the force of heart muscle contractions, drugs such

as ouabain and digoxin that inhibit the Na^+/K^+ ATPase are widely used in the treatment of congestive heart failure. ■

Several Cotransporters Regulate Cytosolic pH

The anaerobic metabolism of glucose yields lactic acid, and aerobic metabolism yields CO_2, which combines with water to form carbonic acid (H_2CO_3). These weak acids dissociate, yielding H^+ ions (protons); if these excess protons were not removed from cells, the cytosolic pH would drop precipitously, endangering cellular functions. Two types of cotransporters help remove some of the excess protons generated during metabolism in animal cells. One is a $Na^+HCO_3{}^-/Cl^-$ *antiporter*, which imports one Na^+ ion, together with one $HCO_3{}^-$, in exchange for export of one Cl^- ion. The cytosolic enzyme *carbonic anhydrase* catalyzes the hydration of carbon dioxide and the dehydration of bicarbonate to convert water and carbon dioxide to carbonic acid, which dissociates to create H^+ and bicarbonate. Carbonic anhydrase is a zinc metalloenzyme, and zinc binding causes deprotonation of water to create a zinc-hydroxide intermediate that attacks carbon dioxide to generate bicarbonate.

$$HCO_3{}^- + H^+ \overset{\text{Carbonic anhydrase}}{\rightleftharpoons} CO_2 + H_2O$$

Thus the overall action of this enzyme is to *consume* cytosolic H^+ ions, thereby *raising* the cytosolic pH. Also important in raising cytosolic pH is a Na^+/H^+ *antiporter*, which couples the movement of one Na^+ ion into the cell down its concentration gradient to the export of one H^+ ion.

Under certain circumstances, the cytosolic pH can rise beyond the normal range of 7.2–7.5. To cope with the excess OH^- ions associated with elevated pH, many animal cells use an *anion antiporter* that catalyzes a one-for-one exchange of HCO_3^- for Cl^- across the plasma membrane. At high pH, this Cl^-/HCO_3^- *antiporter* exports one molecule of HCO_3^- (which can be viewed as a "complex" of OH^- and CO_2) in exchange for the import of one molecule of Cl^-, thus lowering the cytosolic pH. The movement of Cl^- down its concentration gradient ($Cl^-_{exterior} > Cl^-_{cytosol}$; see Table 11-2) powers the export of HCO_3^-.

The activity of all three of these antiporters is regulated by the cytosolic pH, providing cells with a finely tuned mechanism for controlling cytosolic pH. The two antiporters that operate to increase cytosolic pH are activated when the pH of the cytosol falls. Similarly, a rise in pH above 7.2 stimulates the Cl^-/HCO_3^- antiporter, leading to a more rapid export of HCO_3^- and a drop in the cytosolic pH. In this manner, the cytosolic pH of growing cells is maintained very close to pH 7.4.

An Anion Antiporter Is Essential for Transport of CO_2 by Erythrocytes

Transmembrane anion exchange is essential for an important function of erythrocytes: the transport of waste CO_2 from peripheral tissues to the lungs for exhalation. Waste CO_2 released from cells into the capillary blood freely diffuses across the erythrocyte membrane (Figure 11-27a). In its gaseous form, CO_2 dissolves poorly in aqueous solutions such as the cytosol or blood plasma, as is apparent to anyone who has opened a bottle of a carbonated beverage. However, the large amount of the potent enzyme carbonic anhydrase in the erythrocyte combines CO_2 with hydroxyl ions (OH^-) to form water-soluble bicarbonate ($HCO_3{}^-$) anions. This process occurs while erythrocytes are in systemic (tissue) capillaries and releasing oxygen into the blood plasma. The release of oxygen from hemoglobin induces a change in its conformation that enables a histidine side chain of a globin polypeptide to bind a proton. Thus when erythrocytes are in systemic capillaries, water is split into a proton that binds hemoglobin and an OH^- that reacts with CO_2 to form an $HCO_3{}^-$ anion.

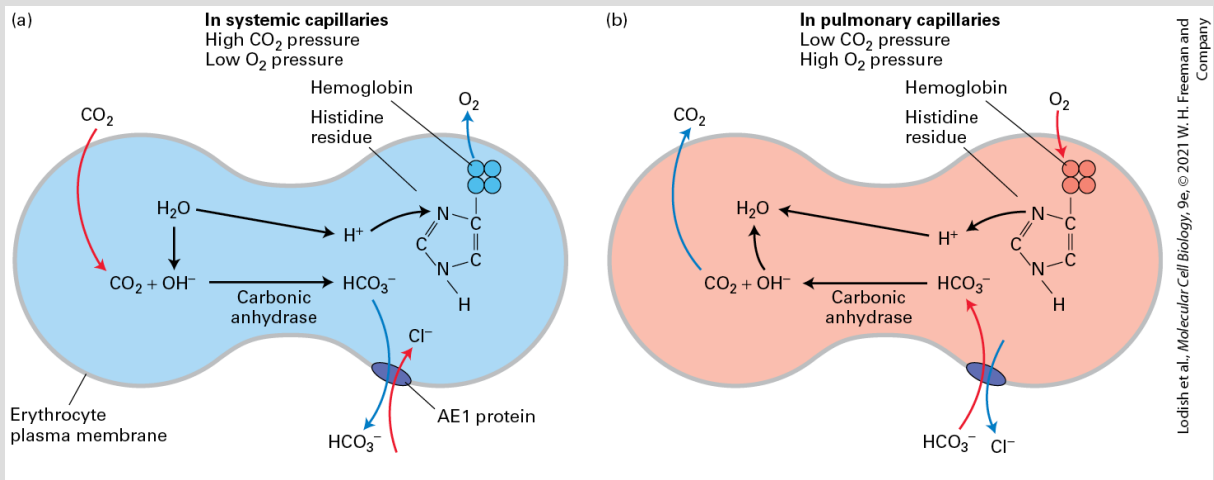


FIGURE 11-27 Carbon dioxide transport in blood requires a Cl^-/HCO_3^- antiporter. (a) In systemic capillaries, carbon dioxide gas diffuses across the erythrocyte plasma membrane and is converted into soluble HCO_3^- by the enzyme carbonic anhydrase; at the same time, oxygen leaves the cell and hemoglobin binds a proton. The anion antiporter AE1 (purple) catalyzes the reversible exchange of Cl^- and HCO_3^- ions across the membrane. The overall reaction causes HCO_3^- to be released from the cell, which is essential for maximal CO_2 transport from the tissues to the lungs and for maintaining pH neutrality in the erythrocyte. (b) In the lungs, where carbon dioxide is excreted, the overall reaction is reversed. See text for additional discussion.

In a reaction catalyzed by the antiporter *Anion Exchanger 1 (AE1)*, also known as *Band 3 anion transporter*, cytosolic HCO_3^- is transported out of the erythrocyte in exchange for an entering Cl^- anion:

$$HCO_3^-{}_{in} + Cl^-{}_{out} \rightleftharpoons HCO_3^-{}_{out} + Cl^-{}_{in}$$

(see Figure 11-27a). The entire anion-exchange process is completed within 50 milliseconds (ms). If anion exchange did not occur, then during periods such as exercise, when much CO_2 is generated,

HCO_3^- would accumulate inside the erythrocyte to toxic levels, as the cytosol would become alkaline. The exchange of HCO_3^- (equal to $OH^- + CO_2$) for Cl^- causes the cytosolic pH to remain nearly neutral. Normally, about 80 percent of the CO_2 in blood is transported as HCO_3^- generated inside erythrocytes; anion exchange allows about two-thirds of this HCO_3^- to be transported by blood plasma external to the cells, increasing the amount of CO_2 that can be transported from tissues to the lungs. In the lungs, where CO_2 leaves the body, the overall direction of this anion-exchange process is reversed (Figure 11-27b).

AE1 catalyzes the precise one-for-one sequential exchange of anions on opposite sides of the plasma membrane required to preserve electroneutrality in the cell; only once every 10,000 or so transport cycles does an anion move from one side of the membrane to the other without another anion moving across the membrane in the opposite direction. AE1 is the most abundant membrane protein in erythrocytes, with nearly one million molecules per cell. The recently solved crystal structure of human erythrocyte AE1 reveals that it has 14 transmembrane α helices and a cytosolic-facing domain that anchors certain cytoskeletal proteins, including ankyrin (see Figure 17-21), to the membrane. The 14 transmembrane α helices form two structurally distinct domains within the membrane, defined as the core and the gate (Figure 11-28a). The core and the gate are divided by a cleft, which, in the outward-facing conformation of the protein whose structure was

solved, is on the extracellular side of the protein. The anion binding site (for chloride) is formed by positively charged dipoles between two of the α helices within the core domain. Comparison of the AE1 structure to the structures of related transporters, such as the Na^+ leucine transporter shown in Figure 11-26, suggests that the antiporter uses an alternating access mechanism: the chloride ion binds the anion binding site in the outward-facing protein, triggering conformational changes such that the core domain rotates against the gate domain to form an inward-facing structure. The chloride ion diffuses out to be replaced by bicarbonate, followed by a reversal of this process.

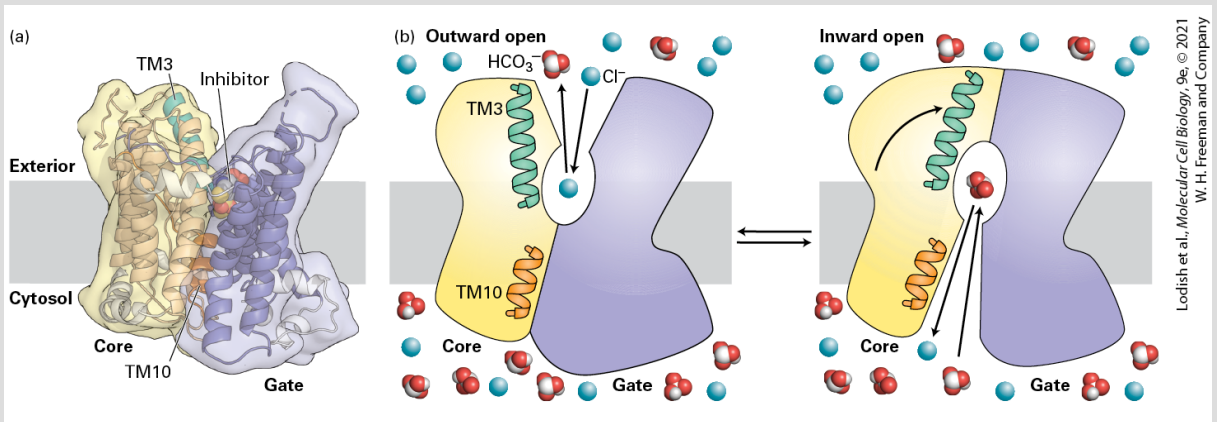


FIGURE 11-28 The antiporter Anion Exchanger 1 (AE1) transports HCO_3^- out of the cell and Cl^- into the cell. (a) Topology of Human AE1, with 14 transmembrane domains (TMs) drawn as cylinders viewed in the plane of the membrane, showing the core domain in yellow and the gate in blue. The putative binding site is located between TM domain 3 (cyan) and TM 10 (orange). (b) HCO_3^- and Cl^- bind in the anion binding site between TMs 3 and 10. The anions are transported across the membrane via an alternating access mechanism triggered by the relative movements of the core and gate domains.

Numerous Transport Proteins Enable Plant Vacuoles to Accumulate Metabolites and Ions

The lumen of a plant vacuole is much more acidic (pH 3–6) than is the cytosol (pH 7.5). The acidity of vacuoles is maintained by a V-class ATP-powered proton pump (see Figure 11-9) and by a pyrophosphate-powered proton pump that is unique to plants. Both of these pumps, located in the vacuolar membrane, import H^+ ions into the vacuolar lumen against a concentration gradient. The vacuolar membrane also contains Cl^- and NO_3^- channels that transport these anions from the cytosol into the vacuole. The proton pumps generate an inside-positive electric potential that drives the entry of these anions against their concentration gradients. The combined operation of these proton pumps and anion channels produces an inside-positive electric potential of about 20 mV across the vacuolar membrane as well as a substantial pH gradient (Figure 11-29).

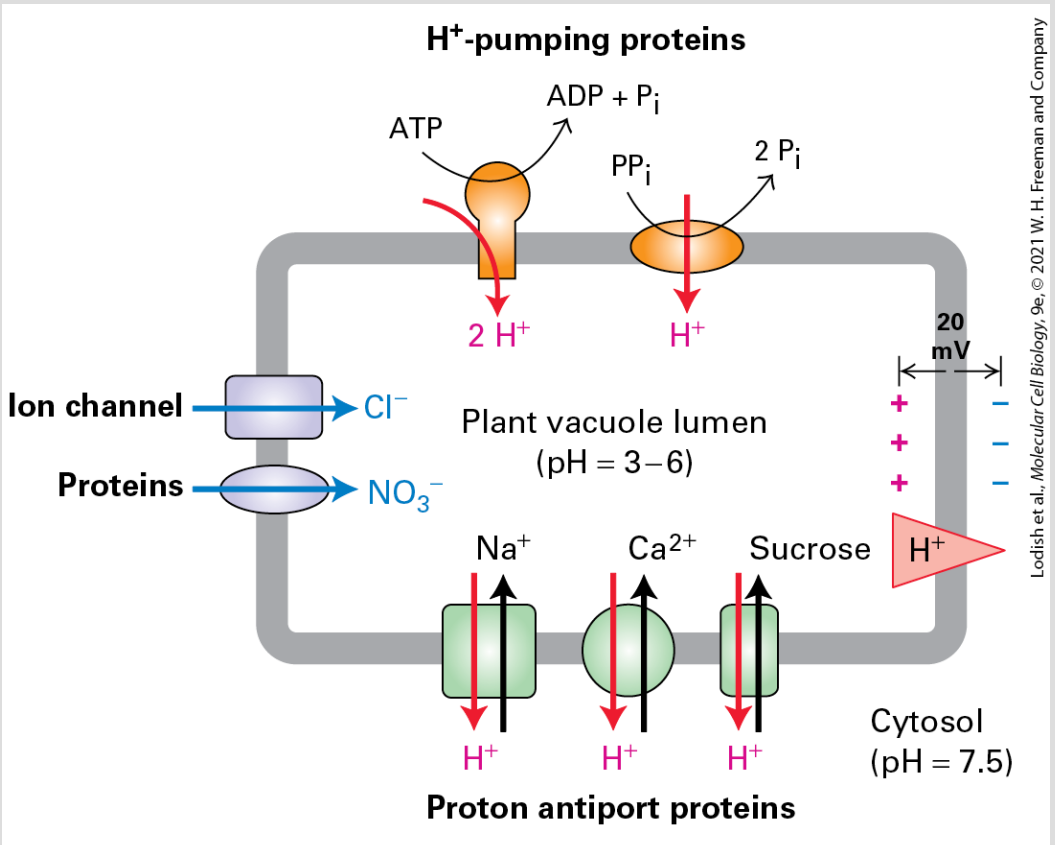


FIGURE 11-29 Concentration of ions and sucrose by the plant vacuole. The vacuolar membrane contains two types of proton pumps (orange): a V-class H^+ ATPase (*left*) and a pyrophosphate-hydrolyzing proton pump (*right*) that differs from all other ion pumps and is probably unique to plants. These pumps generate a low luminal pH as well as an inside-positive electric potential across the vacuolar membrane by their inward pumping of H^+ ions. The inside-positive potential powers the movement of Cl^- and NO_3^- from the cytosol through separate channel proteins (purple). Proton antiporters (green), powered by the H^+ gradient, accumulate Na^+, Ca^{2+}, and sucrose inside the vacuole. See B. J. Barkla and O. Pantoja, 1996, *Annu. Rev. Plant Phys.* **47**:159–184; and P. A. Rea et al., 1992, *Trends Biochem. Sci.* **17**:348.

The proton electrochemical gradient across the plant vacuole membrane is used in much the same way as the Na^+ electrochemical gradient across the animal-cell plasma membrane:

to power the selective uptake or extrusion of ions and small molecules by various antiporters. In the leaf, for example, excess sucrose generated by photosynthesis during the day is stored in the vacuole; during the night, the stored sucrose moves into the cytoplasm and is metabolized to CO_2 and H_2O with concomitant generation of ATP from ADP and P_i. A *proton/sucrose antiporter* in the vacuolar membrane operates to accumulate sucrose in plant vacuoles. The inward movement of sucrose is powered by the outward movement of H^+, which is favored by its concentration gradient (lumen > cytosol) and by the cytosolic-negative potential across the vacuolar membrane (see Figure 11-29). Uptake of Ca^{2+} and Na^+ into the vacuole from the cytosol against their concentration gradients is similarly mediated by proton antiporters.

Our understanding of the transporters in plant vacuolar membranes has the potential for increasing agricultural production in soils with a high salt (NaCl) concentration, which are found throughout the world. Because most agriculturally useful crops cannot grow in such saline soils, agricultural scientists have long sought to develop salt-tolerant plants by traditional breeding methods. With the availability of the cloned gene encoding the vacuolar Na^+/H^+ antiporter, researchers can now produce transgenic plants that overexpress this transporter, leading to increased sequestration of Na^+ in the vacuole. For instance, transgenic tomato plants that overexpress the vacuolar Na^+/H^+ antiporter can grow, flower, and produce fruit in the presence of soil

NaCl concentrations that kill wild-type plants. Interestingly, although the leaves of these transgenic tomato plants accumulate large amounts of salt, the fruit has a very low salt content. ■

KEY CONCEPTS OF SECTION 11.5

Cotransport by Symporters and Antiporters

- The electrochemical gradient across a semipermeable membrane determines the direction of ion movement through transmembrane proteins. The two forces constituting the electrochemical gradient — the membrane electric potential and the ion concentration gradient — may act in the same or opposite directions (see Figure 11-25).
- Cotransporters use the energy released by movement of an ion (usually H^+ or Na^+) down its electrochemical gradient to power the import or export of a small molecule or different ion against its concentration gradient.
- The Na^+ galactose symporter provides a model for Na^+ glucose and other sugar transporters. Binding of Na^+ and galactose causes specific TM helices to rotate, creating a gating mechanism that provides alternate access to the substrates on the extracellular and cytoplasmic sides of the membrane.
- The cells lining the small intestine and kidney tubules contain symporters that couple the energetically favorable entry of Na^+ to the import of glucose against its concentration gradient (see Figure 11-26). Amino acids also enter cells by means of Na^+-linked symporters.
- The molecular structure of a bacterial Na^+/amino acid symporter reveals how binding of Na^+ and leucine are coupled and provides a snapshot of an occluded transport intermediate in which the bound substrates cannot diffuse out of the protein (see Figure 11-27).
- In cardiac muscle cells, the export of Ca^{2+} is coupled to and powered by the import of Na^+ by a cation antiporter, which transports three Na^+ ions inward for each Ca^{2+} ion exported.
- Two cotransporters that are activated at low pH help maintain the cytosolic pH in animal cells very close to 7.4 despite metabolic production of carbonic and lactic acids. One, a Na^+/H^+ antiporter, exports excess protons. The other, a $Na^+HCO_3{}^-/Cl^-$ cotransporter, imports $HCO_3{}^-$, which dissociates in the cytosol to yield pH-raising OH^- ions.

- A $Cl^-/HCO_3{}^-$ antiporter that is activated when the cytosolic pH rises above normal decreases pH by exporting $HCO_3{}^-$.
- AE1, a $Cl^-/HCO_3{}^-$ antiporter in the erythrocyte membrane, increases the ability of blood to transport CO_2 from tissues to the lungs (see Figure 11-28).
- The structure of human erythrocyte AE1 has been solved, revealing that anions are transported across the membrane by means of an alternating access mechanism triggered by the relative movements of the core and gate domains.
- Uptake of sucrose, Na^+, Ca^{2+}, and other substances into plant vacuoles is carried out by proton antiporters in the vacuolar membrane. Ion channels and proton pumps in the membrane are critical in generating a large enough proton concentration gradient to power these proton antiporters (see Figure 11-29).

11.6 Transcellular Transport

The previous sections have illustrated how several types of transporters function together to carry out important cellular functions. Here we extend this concept by focusing on the transport of several types of molecules and ions across *polarized* cells, which are cells that are asymmetric (have different "sides") and thus have biochemically distinct regions of the plasma membrane. A particularly well-studied class of polarized cells includes many of the epithelial cells that form sheet-like layers (epithelia) covering most external and internal surfaces of body organs. (Epithelial cells are discussed in greater detail in Chapter 20.) Like many epithelial cells, an intestinal epithelial cell that absorbs nutrients from the gastrointestinal tract has a plasma membrane organized into two major discrete regions: the surface that faces the outside of the organism, called the **apical**, or top, surface, and the surface that faces the inside of the organism (or the bloodstream-facing side), called the **basolateral** surface, which is composed of the basal and lateral surfaces of the cell (see Figure 20-11).

Specialized regions of the epithelial-cell plasma membrane, called **tight junctions**, separate the apical and basolateral membranes and prevent many, but not all, water-soluble substances on one side from moving across to the other side through the extracellular space between cells. For this reason, many nutrients are absorbed from the intestinal lumen across the epithelial cell layer and eventually

into the blood by a two-stage process called **transcellular transport**: molecules are imported through the plasma membrane on the apical side of the cells and exported through the plasma membrane on the basolateral (blood-facing) side (Figure 11-30, also see Figure 11-25d). The apical portion of the plasma membrane, which faces the intestinal lumen, is specialized for absorption of sugars, amino acids, and other molecules that are produced from food by multiple digestive enzymes.

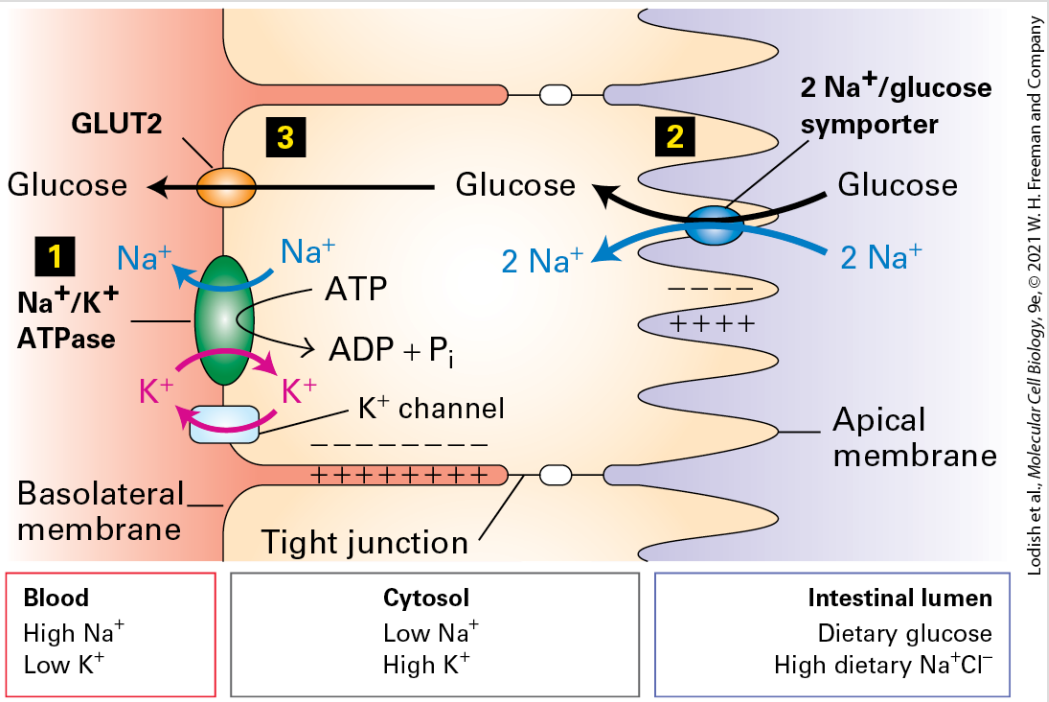


FIGURE 11-30 Transcellular transport of glucose from the intestinal lumen into the blood. The Na^+/K^+ ATPase in the basolateral surface membrane generates Na^+ and K^+ concentration gradients (step 1). The outward movement of K^+ ions through nongated K^+ channels generates an inside-negative membrane potential across the entire plasma membrane. Both the Na^+ concentration gradient and the membrane potential are used to drive the uptake of glucose from the intestinal lumen by the two-Na^+/one-glucose symporter located in the apical surface membrane (step 2). Glucose leaves

the cell via facilitated transport catalyzed by GLUT2, a glucose uniporter located in the basolateral membrane (step 3).

Multiple Transport Proteins Are Needed to Move Glucose and Amino Acids Across Epithelia

Figure 11-30 depicts the proteins that mediate absorption of glucose from the intestinal lumen into the blood. This figure illustrates the important concept that different types of proteins are localized to the apical and basolateral membranes of epithelial cells. The passage of K^+ and Na^+ ions through nongated K^+ channels and through the Na^+/K^+ ATPase generates K^+ and Na^+ gradients as well as an inside-negative membrane potential. The Na^+ concentration gradient and the membrane potential are used to drive the uptake of glucose, against its concentration gradient, from the intestinal lumen into the cell by the two-Na^+/one-glucose symporter located in the apical surface membrane. As noted above, this symporter couples the energetically unfavorable inward movement of one glucose molecule to the energetically favorable inward transport of two Na^+ ions. In the steady state, all the Na^+ ions transported from the intestinal lumen into the cell during Na^+/glucose symport, or the similar process of Na^+/amino acid symport that also takes place on the apical membrane, are pumped out across the basolateral membrane, which faces the blood. Thus the low intracellular Na^+

concentration is maintained. The Na^+/K^+ ATPase that accomplishes this pumping is found exclusively in the basolateral membrane of intestinal epithelial cells. The coordinated operation of these two transport proteins allows uphill movement of glucose and amino acids from the intestine into the cell. This first stage in transcellular transport is ultimately powered by ATP hydrolysis by the Na^+/K^+ ATPase.

In the second stage, the glucose and amino acids concentrated inside intestinal cells by apical symporters are exported down their concentration gradients into the blood via uniport proteins in the basolateral membrane. In the case of glucose, this movement is mediated by GLUT2 (see Figure 11-30 and Figure 11-25d). As noted earlier, this GLUT isoform has a relatively low affinity for glucose but increases its rate of transport substantially when the glucose gradient across the membrane rises (see Figure 11-4).

The net result of this two-stage process is movement of Na^+ ions, glucose, and amino acids from the intestinal lumen across the intestinal epithelium into the extracellular medium that surrounds the basolateral surface of intestinal epithelial cells, and eventually into the blood. Tight junctions between the epithelial cells prevent these molecules from diffusing back into the intestinal lumen. The transport of salt, glucose, and amino acids across the intestinal epithelium increases osmotic pressure and thereby draws water from the intestinal lumen, mainly through the tight junctions, into

the extracellular medium that surrounds the basolateral surface; aquaporins do not appear to play a major role. In a sense, salts, glucose, and amino acids "carry" the water along with them.

Simple Rehydration Therapy Depends on the Osmotic Gradient Created by Absorption of Glucose and Na^+

An understanding of osmosis and the intestinal absorption of salt and glucose forms the basis for a simple therapy that saves millions of lives each year, particularly in developing countries. In these countries, cholera and other intestinal pathogens are major causes of death for young children. A toxin released by these bacteria activates chloride secretion from the apical surfaces of intestinal epithelial cells into the lumen; water follows osmotically, and the resultant massive loss of water causes diarrhea, dehydration, and ultimately death. A cure demands not only killing the bacteria with antibiotics but also *rehydration*: replacement of the water that is lost from the blood and other tissues.

Simply drinking water does not help because it is excreted from the gastrointestinal tract almost as soon as it enters. However, as we have just learned, the coordinated transport of glucose and Na^+ across the intestinal epithelium creates a difference in osmotic pressure across the plasma membrane, forcing water to move from

the intestinal lumen across the epithelial cell layer and ultimately into the blood. Thus giving affected children a solution of sugar and salt to drink (but not sugar or salt alone) increases the transport of sodium and sugar across epithelial cells and, consequently, increases osmotic flow of water into the blood from the intestinal lumen, leading to rehydration. Similar sugar-salt solutions are the basis of popular drinks used by athletes to get sugar as well as water into the body quickly and efficiently. ■

Parietal Cells Acidify the Stomach Contents While Maintaining a Neutral Cytosolic pH

The mammalian stomach contains a 0.1 M solution of hydrochloric acid (HCl). This strongly acidic medium kills many ingested pathogens and denatures many ingested proteins so that they can be degraded by proteolytic enzymes (e.g., pepsin) that function at acidic pH. Hydrochloric acid is secreted into the stomach by specialized epithelial cells called *parietal cells* (also known as *oxyntic cells*) in the stomach lining. These cells contain a H^+/K^+ *ATPase* in the apical membrane (which faces the stomach lumen) that pumps H^+ ions into the stomach in exchange for K^+ ions. The action of this ATPase generates a 1-million-fold H^+ concentration gradient: pH $\sim$1.0 in the stomach lumen versus pH $\sim$7.2 in the cell cytosol. This P-class ATP-powered ion pump is similar in structure and function

to the plasma-membrane Na^+/K^+ ATPase discussed earlier. The numerous mitochondria in parietal cells produce abundant ATP for use by the H^+/K^+ ATPase.

If parietal cells simply exported H^+ ions in exchange for K^+ ions, the loss of protons would lead to a rise in the concentration of OH^- ions in the cytosol and thus a marked increase in cytosolic pH. (Recall that $[H^+] \times [OH^-]$ is always a constant, 10^{-14} M^2.) Parietal cells avoid this rise in cytosolic pH by using Cl^-/HCO_3^- antiporters in the basolateral membrane to export the excess OH^- ions from the cytosol to the blood. As noted earlier, these anion antiporters are activated at high cytosolic pH.

The overall process by which parietal cells acidify the stomach lumen is illustrated in Figure 11-31. In a reaction catalyzed by carbonic anhydrase, the excess cytosolic OH^- combines with CO_2 that diffuses in from the blood, forming HCO_3^-. This bicarbonate ion is exported across the basolateral membrane (and ultimately into the blood) by the Cl^-/HCO_3^- antiporter in exchange for a Cl^- ion. The Cl^- ions then exit through Cl^- channels in the apical membrane, entering the stomach lumen. To preserve electroneutrality, each Cl^- ion that moves into the stomach lumen across the apical membrane is accompanied by a K^+ ion that moves outward through a separate K^+ channel. In this way, the excess K^+ ions pumped inward by the H^+/K^+ ATPase are returned to the stomach lumen, thus maintaining the normal intracellular K^+

concentration. The net result is secretion of equal amounts of H^+ and Cl^- ions (i.e., HCl) into the stomach lumen, while the pH of the cytosol remains neutral and the excess OH^- ions, as HCO_3^-, are transported into the blood, where the change in pH is minimal.

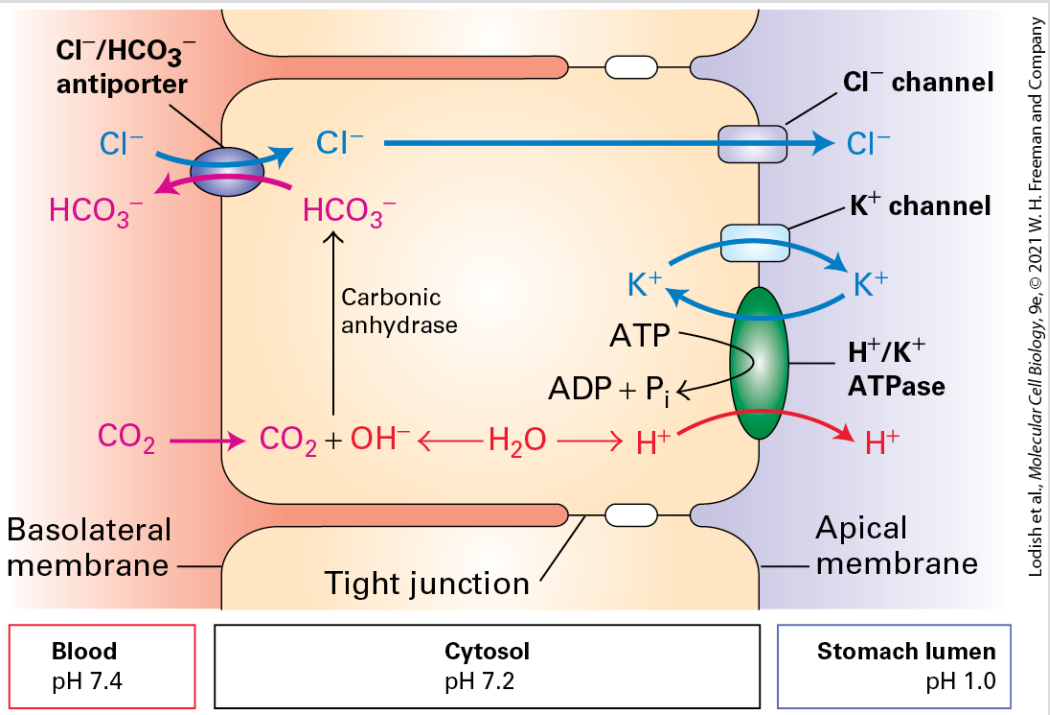


FIGURE 11-31 Acidification of the stomach lumen by parietal cells in the gastric lining. The apical membrane of parietal cells contains a H^+/K^+ ATPase (a P-class pump) as well as Cl^- and K^+ channels. Note the cyclic K^+ transport across the apical membrane: K^+ ions are pumped inward by the H^+/K^+ ATPase and exit via a K^+ channel. The basolateral membrane contains an anion antiporter that exchanges HCO_3^- and Cl^- ions. The combined operation of these four different transport proteins and carbonic anhydrase acidifies the stomach lumen while maintaining the neutral pH of the cytosol.

Bone Resorption Requires the Coordinated Function of a V-Class Proton Pump and a Specific Chloride Channel

Net bone growth in mammals subsides just after puberty, but a finely balanced, highly dynamic process of disassembly (resorption) and reassembly (bone formation) goes on throughout adulthood. Such continual bone *remodeling* permits the repair of damaged bones and can release calcium, phosphate, and other ions from mineralized bone into the blood for use elsewhere in the body.

Osteoclasts, the bone-dissolving cells, are macrophages, a type of cell best known for their role in protecting the body from infections. Osteoclasts are polarized cells that adhere to bone and form specialized, very tight seals between themselves and the bone, creating an enclosed extracellular space (Figure 11-32). An adhered osteoclast then secretes into this space a corrosive mixture of HCl and proteases that dissolves the inorganic components of the bone into Ca^{2+} and phosphate and digests its protein components. The mechanism of HCl secretion is similar to that used by the stomach to generate digestive juice (see Figure 11-31). As in gastric HCl secretion, carbonic anhydrase and an anion antiporter are important for osteoclast function. Osteoclasts employ a V-class

proton pump to export H^+ ions into the bone-facing space, rather than the P-class H^+/K^+ pump used by gastric epithelial cells.

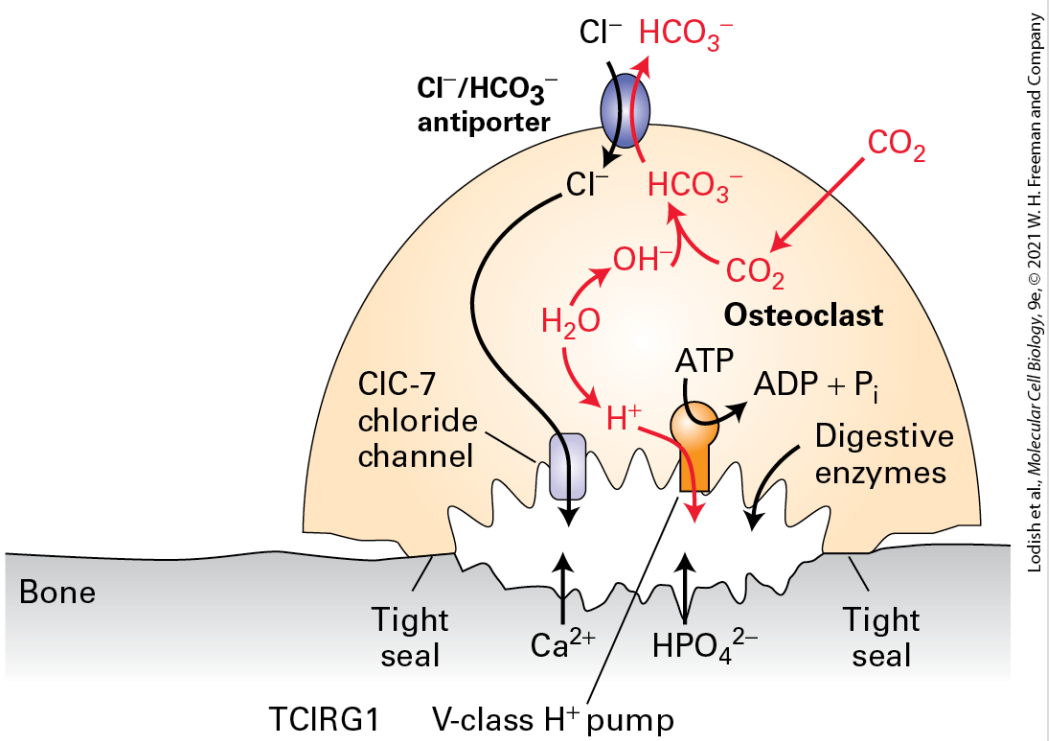


FIGURE 11-32 Dissolution of bone by polarized osteoclast cells requires a V-class proton pump and the ClC-7 chloride channel. The osteoclast plasma membrane is divided into two domains separated by the tight seal between a ring of membrane and the bone surface. The membrane domain facing the bone contains V-class proton pumps and ClC-7 Cl^- channels. The opposing membrane domain contains anion antiporters that exchange HCO_3^- and Cl^- ions. The combined operation of these three transport proteins and carbonic anhydrase acidifies the enclosed space and allows bone resorption while maintaining the neutral pH of the cytosol. See R. Planells-Cases and T. Jentsch, 2009, *Biochim. Biophys. Acta* **1792**:173 for discussion of ClC-7.

The rare hereditary disease *osteopetrosis*, marked by increased bone density, is due to abnormally low bone resorption. Many

patients have mutations in the gene encoding TCIRG1, a subunit of the osteoclast V-class proton pump, whose action is required to acidify the space between the osteoclast and the bone. Other patients have mutations in the gene encoding ClC-7, the chloride channel localized to the domain of the osteoclast plasma membrane that faces the space near the bone. As with lysosomes (see Figure 11-13), in the absence of a chloride channel, the proton pump cannot acidify the enclosed extracellular space, and thus bone resorption is defective. ■

KEY CONCEPTS OF SECTION 11.6

Transcellular Transport

- The apical and basolateral regions of epithelial cell plasma membranes contain different transport proteins and carry out quite different transport processes.
- In intestinal epithelial cells, the coordinated operation of Na^+-linked symporters in the apical membrane and Na^+/K^+ ATPases and uniporters in the basolateral membrane mediates transcellular transport of amino acids and glucose from the intestinal lumen to the blood (see Figure 11-30).
- The increased osmotic pressure created by transcellular transport of salt, glucose, and amino acids across the intestinal epithelium draws water from the intestinal lumen into the body, a phenomenon that serves as the basis for rehydration therapy using sugar-salt solutions.
- The combined action of carbonic anhydrase and four different transport proteins permits parietal cells in the stomach lining to secrete HCl into the lumen while maintaining their cytosolic pH near neutrality (see Figure 11-31).
- Bone resorption requires coordinated function in osteoclasts of a V-class proton pump and the ClC-7 chloride channel (see Figure 11-32).

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Classic Experiment 11-1: Stumbling upon Active Transport
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

ABC superfamily
active transport
antiporter
apical
aquaporins
ATP-powered pump
basolateral
channel
cotransport
cotransporter
electrochemical gradient

facilitated transport
flippase
membrane potential
membrane transport proteins
Na^+/K^+ ATPase
osmosis
patch clamping
P-class pump
resting K^+ channel
resting membrane potential
symporter
tight junction
transcellular transport
transporter
uniport

Review the Concepts

1. Nitric oxide (NO) is a gaseous molecule with lipid solubility similar to that of O_2 and CO_2. Endothelial cells lining arteries use NO to signal surrounding smooth muscle cells to relax, thereby increasing blood flow. What mechanism or mechanisms would transport NO from where it is produced in the cytoplasm of an endothelial cell into the cytoplasm of a smooth muscle cell, where it acts?
2. Acetic acid (a weak acid with a pK_a of 4.75) and ethanol (an alcohol) are each composed of two carbon atoms, hydrogen

and oxygen, and both enter cells by passive diffusion. At pH 7, one is much more able to permeate a cellular membrane than the other. Which is more membrane permeable, and why? Predict how the membrane permeability of each is altered when the extracellular pH is reduced to 1.0, a value typical of the stomach.

3. Uniporters and ion channels support facilitated transport across cellular membranes. Although both are examples of facilitated transport, the rates of ion movement via an ion channel are roughly 10^4- to 10^5-fold faster than the rates of molecule movement via a uniporter. What key mechanistic difference results in this large difference in transport rate? What contribution to free energy (ΔG) determines the direction of transport?
4. Name the three classes of membrane transport proteins. Explain which one or ones of these classes is able to move glucose and which can move bicarbonate (HCO_3^-) against an electrochemical gradient. In the case of bicarbonate, but not glucose, the ΔG of the transport process has two terms. What are these two terms, and why does the second not apply to glucose? Why are cotransporters often referred to as examples of secondary active transport?
5. An H^+ ion is smaller than an H_2O molecule, and a glycerol molecule, a three-carbon alcohol, is much larger. Both readily dissolve in H_2O. Why do aquaporins fail to transport H^+ whereas some can transport glycerol?
6. GLUT1, found in the plasma membrane of erythrocytes, is a classic example of a uniporter.

a. Design a set of experiments to prove that GLUT1 is indeed a glucose-specific uniporter rather than a galactose- or mannose-specific uniporter.
b. Glucose is a six-carbon sugar, and ribose is a five-carbon sugar. Despite its smaller size, ribose is not efficiently transported by GLUT1. How can this be explained?
c. A drop in blood sugar from 5 mM to 2.8 mM or below can cause confusion and fainting. Calculate the effect of this drop on glucose transport into cells expressing GLUT1.
d. How do liver and muscle cells maximize glucose uptake without changing V_{max}?
e. Tumor cells expressing GLUT1 often have a higher V_{max} for glucose transport than do normal cells of the same type. How could these cells increase the V_{max}?
f. Fat and muscle cells modulate the V_{max} for glucose uptake in response to insulin signaling. How?

7. Name the four classes of ATP-powered pumps that produce active transport of ions and molecules. Indicate which of these classes transport only ions and which transport primarily small organic molecules. The initial discovery of one class of these ATP-powered pumps came from studying the transport not of a natural substrate, but rather of artificial substrates used as cancer chemotherapy drugs. What do investigators now think are common examples of the natural substrates of this particular class of ATP-powered pumps?

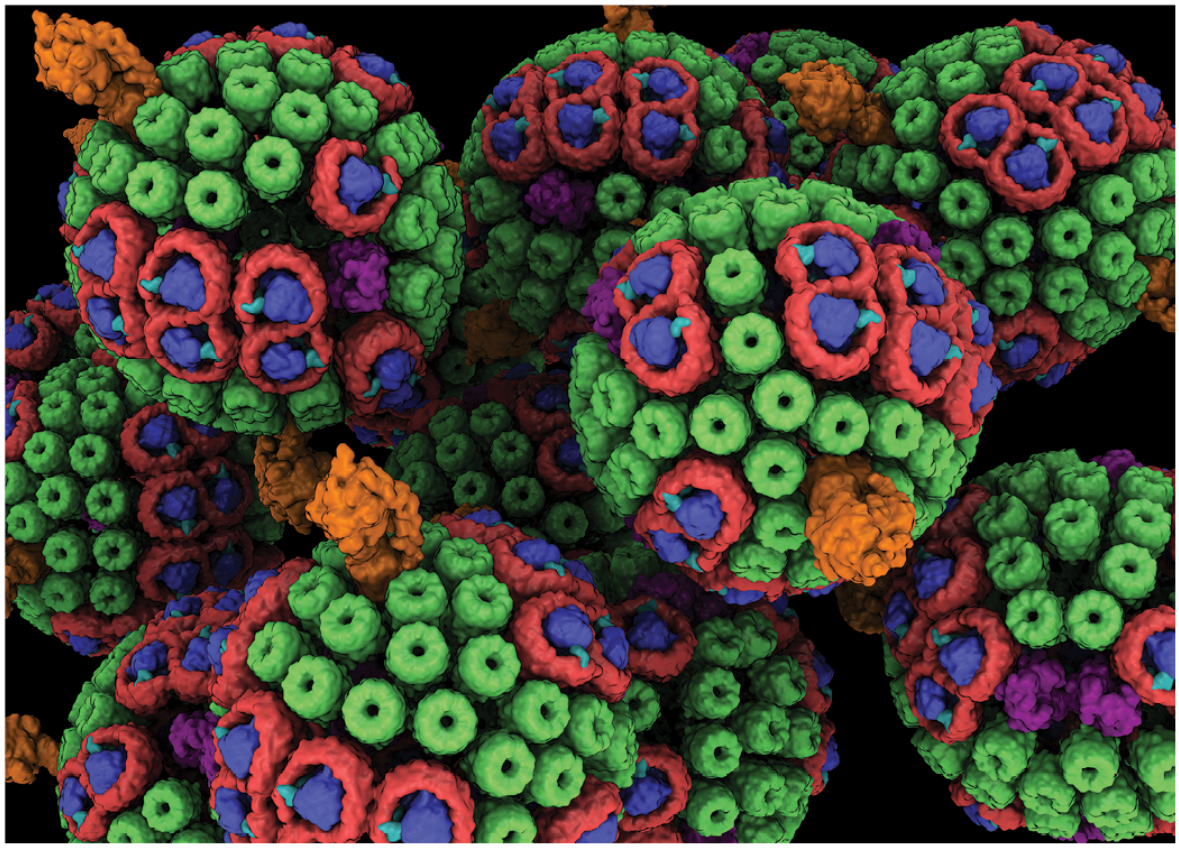
8. Explain why the coupled reaction $ATP \rightarrow ADP + P_i$ in the P-class ion pump mechanism does not involve direct hydrolysis of the phosphoanhydride bond.
9. Describe a negative feedback mechanism for controlling a rising cytosolic Ca^{2+} concentration in cells that require rapid changes in Ca^{2+} concentration for normal functioning. How would a drug that inhibits calmodulin activity affect cytosolic Ca^{2+} concentration regulation by this mechanism? What would be the effect on the function of, for example, a skeletal muscle cell?
10. Certain proton pump inhibitors that inhibit secretion of stomach acid are among the most widely sold drugs in the world today. What pump does this type of drug inhibit, and where is this pump located?
11. The membrane potential in animal cells, but not in plants, depends largely on resting K^+ channels. How do these channels contribute to the resting membrane potential? Why are these channels considered to be nongated channels? How do these channels achieve selectivity for K^+ versus Na^+, which is smaller than K^+?
12. Patch clamping can be used to measure the conductance properties of individual ion channels. Describe how patch clamping can be used to determine whether or not the gene coding for a putative K^+ channel actually codes for a K^+ or a Na^+ channel.
13. Plants use the proton electrochemical gradient across the vacuole membrane to power the accumulation of salts and sugars in the organelle. This accumulation creates

hypertonic conditions in the vacuole. Why does this not result in the plant cell swelling and bursting? Even under isotonic conditions, there is a slow leakage of ions into animal cells. How does the plasma-membrane Na^+/K^+ ATPase enable animal cells to avoid osmotic lysis under isotonic conditions?

14. In the case of the bacterial two-Na^+/one-leucine symporter, what is the key distinguishing feature of the bound Na^+ ions that ensures that other ions, particularly K^+, do not bind?
15. Describe the symport process by which cells lining the small intestine import glucose. What ion is responsible for the transport, and what two particular features facilitate the energetically favored movement of this ion across the plasma membrane?
16. Movement of glucose from one side to the other side of the intestinal epithelium is a major example of transcellular transport. How does the Na^+/K^+ ATPase power the process? Why are tight junctions essential for the process? Why is localization of the transporters specifically in the apical or basolateral membrane crucial for transcellular transport? Rehydration supplements such as sport drinks include a sugar and a salt; why are both important to rehydration?

CHAPTER 12

Cellular Energetics



Angela M. Barragan

Computer generated image of chromatophores in the photosynthetic purple bacterium *Rhodobacter sphaeroides*. Each chromatophore contains membrane proteins that use light to drive the synthesis of ATP. Energy is absorbed by pigments in two types of light-harvesting complexes (green and red) and transferred to reaction centers (light blue) where high-energy electrons are generated. Their energy is used by pumps (purple) to move protons into the chromatophores, generating a proton gradient. This powers ATP synthase (orange) to convert ADP + P_i to ATP.

OUTLINE

12.1 Chemiosmosis, Electron Transport, the Proton-Motive Force, and ATP Synthesis
12.2 First Step of Harvesting Energy from Glucose: Glycolysis
12.3 The Structure of Mitochondria
12.4 The Dynamics of Mitochondria and Mitochondrial-ER Membrane Contact Sites
12.5 The Citric Acid Cycle and Fatty Acid Oxidation
12.6 The Electron-Transport Chain and Generation of the Proton-Motive Force
12.7 Harnessing the Proton-Motive Force to Synthesize ATP
12.8 Chloroplasts and Photosynthesis
12.9 Use of Light Energy to Generate Molecular Oxygen, NADPH, and ATP in Stages 1–3 of Photosynthesis
12.10 ATP and NADPH Drive Carbon Fixation in the Calvin Cycle and Carbohydrate Synthesis in Stage 4 of Photosynthesis

From the growth and division of a cell to the beating of a heart to the electrical activity of a neuron that underlies thinking, life requires energy. Energy is defined as the capacity to do work, and on a cellular level, that work includes conducting and regulating a multitude of chemical reactions and transport processes, growing and dividing, generating and maintaining a highly organized structure, and interacting with other cells. This chapter describes the molecular mechanisms by which cells use chemical nutrients or sunlight as sources of energy, with a special focus on how cells convert these external sources of energy into a biologically universal intracellular chemical energy carrier, **adenosine triphosphate**, or **ATP**. ATP, found in all types of organisms and presumably present in the earliest life forms, is generated by the chemical addition of inorganic phosphate (HPO_4^{2-}, often abbreviated as P_i) to adenosine diphosphate, or ADP, a process called *phosphorylation*. Cells use the

energy released during hydrolysis of the terminal phosphoanhydride bond in ATP (see Figures 1-7 and 2-31) to power many otherwise energetically unfavorable processes. Examples include the synthesis of proteins from amino acids and of nucleic acids from nucleotides (see Chapter 4), the transport of molecules against a concentration gradient by ATP-powered pumps (see Chapter 11), the contraction of muscles (see Chapter 17), and the beating of cilia (see Chapter 18). A key theme of cellular energetics is that proteins use, or *couple*, energy released from one process (e.g., ATP hydrolysis) to drive another process (e.g., movement of molecules across membranes) that otherwise would be thermodynamically unfavorable.

The energy to drive ATP synthesis from ADP ($\Delta G^{\circ\prime} = 7.3$ kcal/mol) derives primarily from two sources: the energy in the chemical bonds of nutrients and the energy in sunlight (Figure 12-1). The two processes primarily responsible for converting these energy sources into ATP are **aerobic oxidation** (also known as **aerobic respiration**), which occurs in mitochondria in nearly all eukaryotic cells (see Figure 12-1, *top*), and **photosynthesis**, which occurs only in **chloroplasts**, found in the leaf cells of plants (see Figure 12-1, *bottom*) and in certain single-celled organisms, such as algae and cyanobacteria. Two additional processes, glycolysis and the citric acid cycle (see Figure 12-1, *top left*), are also important direct or indirect sources of ATP in both animal and plant cells.

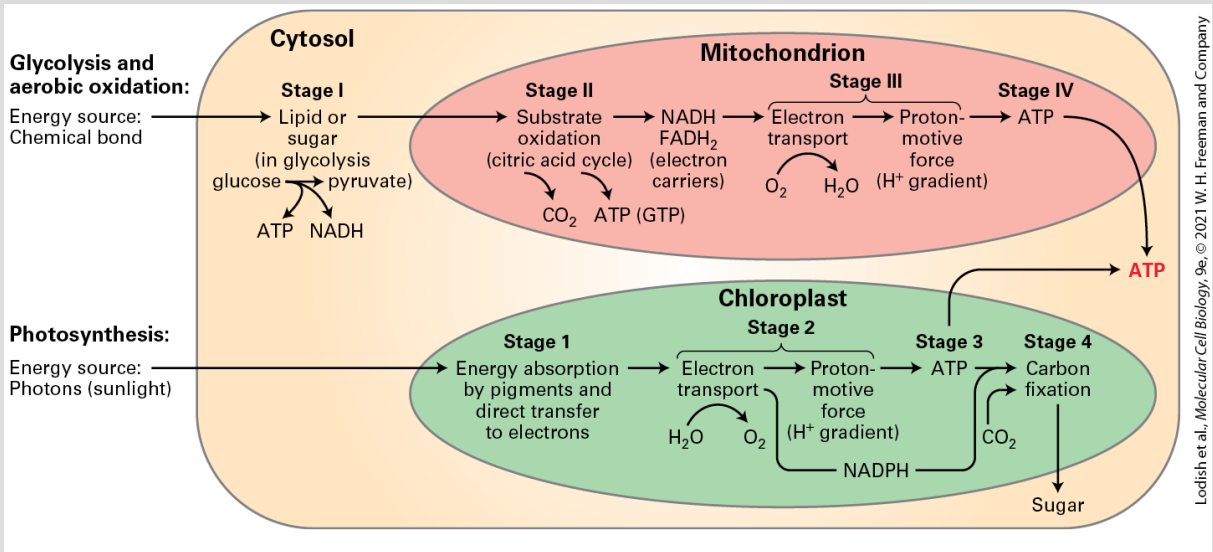


FIGURE 12-1 Overview of aerobic oxidation and photosynthesis. Eukaryotic cells use two fundamental mechanisms to convert external sources of energy into ATP. *Top:* In aerobic oxidation, "fuel" molecules [primarily sugars and fatty acids (lipids)] undergo preliminary processing in the cytosol, such as breakdown of glucose to pyruvate (**stage I**), and are then transferred into mitochondria, where they are converted by oxidation with O_2 to CO_2 and H_2O (**stages II** and **III**) and ATP is generated (**stage IV**). *Bottom:* In photosynthesis, which occurs in chloroplasts, the radiant energy of light is absorbed by specialized pigments (**stage 1**); the absorbed energy is used both to oxidize H_2O to O_2 and to establish conditions (**stage 2**) necessary for the generation of ATP (**stage 3**) and of carbohydrates from CO_2 (carbon fixation, **stage 4**). Both mechanisms involve the production of reduced high-energy electron carriers ($NADH$, $NADPH$, $FADH_2$) and the movement of electrons down an electric potential gradient in an electron-transport chain through specialized membranes. Energy released from these electrons is captured as a proton electrochemical gradient (proton-motive force) that is then used to drive ATP synthesis. Bacteria use comparable processes.

In aerobic oxidation, breakdown products of sugars (carbohydrates) and fatty acids (hydrocarbons) — both derived from the digestion of food in animals — are converted by oxidation with oxygen (O_2) to carbon dioxide (CO_2) and water (H_2O). The energy released from this overall reaction is transformed into the chemical energy of

phosphoanhydride bonds in ATP. This process is analogous to burning wood (carbohydrates) or oil (hydrocarbons) to generate heat in furnaces or motion in automobile engines: both consume O_2 and generate CO_2 and H_2O. The key difference is that cells break down the overall reaction into many intermediate steps, with the amount of energy released in any given step closely matched to the amount of energy that can be stored — for example, as ATP — or that is required for the next intermediate step. If there were not such a close match, either excess released energy would be lost as heat (which would be very inefficient), or there would not be enough energy released in a step to generate energy storage molecules such as ATP or to drive the next step in the process (which would be ineffective).

In photosynthesis, the radiant energy of light is absorbed by pigments such as chlorophyll and used to make ATP and carbohydrates — primarily sucrose and starch. Unlike aerobic oxidation, which uses carbohydrates and O_2 to generate CO_2, photosynthesis uses CO_2 as a substrate and generates O_2 and carbohydrates as products.

This reciprocal relationship between aerobic oxidation in mitochondria and photosynthesis in chloroplasts underlies a profound symbiotic relationship between photosynthetic and non-photosynthetic organisms. The oxygen generated during photosynthesis is the source of virtually all the oxygen in the air on

our planet, and the carbohydrates produced are the ultimate source of energy for virtually all non-photosynthetic organisms on Earth. (An exception is bacteria living in deep-sea hydrothermal vents — and the organisms that feed on them — which obtain energy for converting CO_2 into carbohydrates by oxidation of geologically generated reduced inorganic compounds released by the vents.)

In this chapter, we explore the structures and functions of mitochondria and chloroplasts and the molecular mechanisms underlying aerobic oxidation and photosynthesis. We will first consider the mechanism, known as **chemiosmosis** (also called *chemiosmotic coupling*), used by mitochondria and chloroplasts in eukaryotic cells and by bacteria to generate ATP from ADP and P_i.

12.1 Chemiosmosis, Electron Transport, the Proton-Motive Force, and ATP Synthesis

Chemiosmosis is based on a simple, yet profound, concept. In Chapter 2 (Section 2.4) we described the three forms of potential, or stored, energy in biology: chemical bond energy, the energy stored in concentration gradients, and that stored in electrical gradients. All three forms are interconvertible. For example, as we learned in Chapter 11, the hydrolysis of ATP, which releases its chemical potential energy, can drive, via ATPase pumps, the formation of chemical concentration gradients and electric voltage gradients across biological membranes. Thus it makes sense that the movement of molecules across membranes and down a chemical concentration gradient or movement of ions (e.g., protons, H^+) down an electric voltage gradient release energy. The released energy could be dissipated as heat or harnessed to drive the synthesis of ATP from ADP and P_i. As we also learned in Chapter 2, in redox reactions electrons can release energy as they move from a higher energy state to a lower energy state (down their electric potential gradient), for example, when they move from atoms or molecules having more negative reduction potentials to those with more positive reduction potentials. The released energy can be used to drive chemical reactions and generate electrical and chemical gradients across membranes.

In chemiosmosis, a proton electrochemical gradient across a membrane is first generated by energy released as electrons travel down their electric potential gradient through an **electron-transport chain**. Electron-transport chains will be described in detail below. The energy stored in the proton electrochemical gradient, called the **proton-motive force**, is then used to power the synthesis of ATP (Figure 12-2) or other energy-requiring processes when the protons move back across the membrane down the electrochemical gradient. As protons move down their electrochemical gradient through the ATP synthesizing enzyme called **ATP synthase**, ATP is generated from ADP and P_i, a process that is the reverse of that mediated by the ATP-powered ion pumps discussed in Chapter 11. In ATP-powered pumps, the energy released by ATP hydrolysis is used to move substances up an electrochemical gradient; in ATP synthase, protons moving down their electrochemical gradient release energy used to make ATP from ADP and P_i. We will review in detail the mechanism by which ATP synthase exploits a proton-based electrochemical gradient to make ATP later in this chapter.

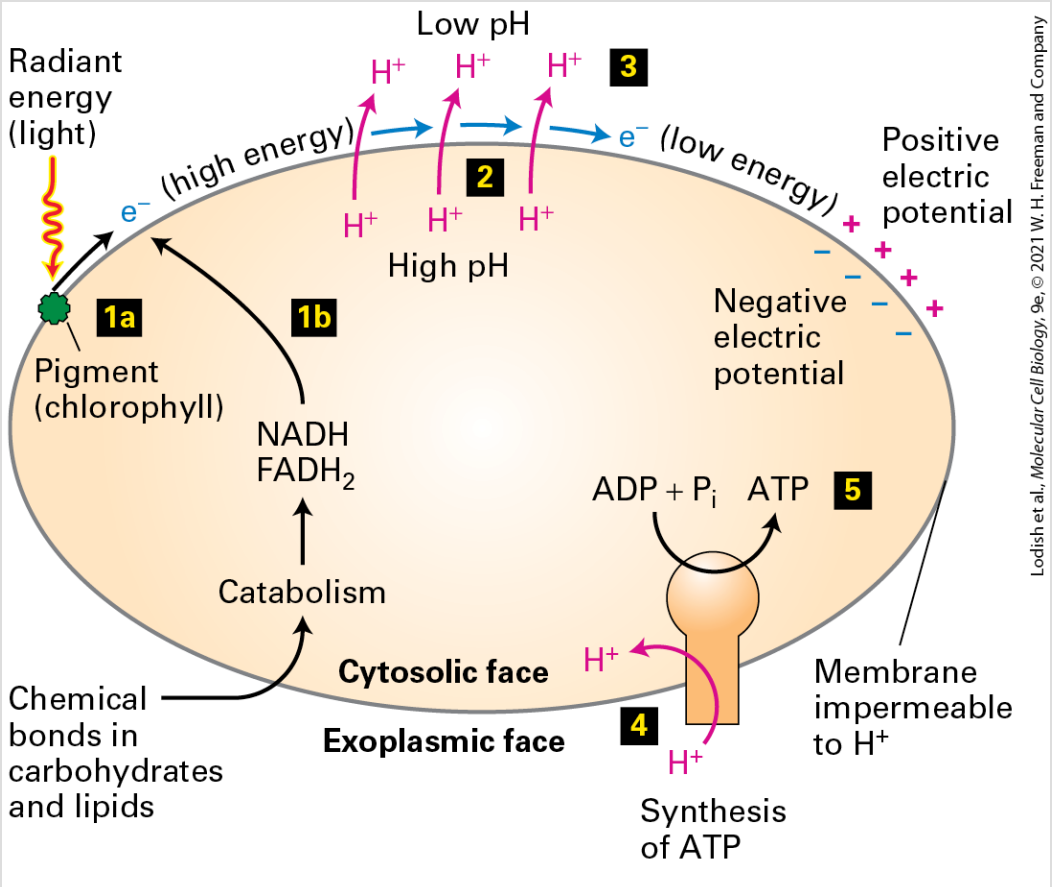


FIGURE 12-2 The proton-motive force powers ATP synthesis. Transmembrane proton concentration and electrical (voltage) gradients, collectively called the *proton-motive force,* are generated during aerobic oxidation and photosynthesis, in both eukaryotes and prokaryotes (bacteria). High-energy electrons (e^-) generated by light absorption by pigments (e.g., chlorophyll, step **1a**), or held in the reduced form of electron carriers (e.g., NADH, $FADH_2$, step **1b**) made during the catabolism of sugars and lipids, pass down an electron-transport chain (blue arrows, step **2**), releasing energy throughout the process. The released energy is used to pump protons across the membrane (red arrows, step **3**), generating the proton-motive force (pH and electric voltage gradients across the membrane). In chemiosmotic coupling, the energy released when protons flow down the gradient through ATP synthase (step **4**) drives the synthesis of ATP (step **5**). The proton-motive force can also power other processes, such as the transport of metabolites across the membrane against their concentration gradient and rotation of bacterial flagella.

KEY CONCEPTS OF SECTION 12.1

Chemiosmosis, Electron Transport, the Proton-Motive Force, and ATP Synthesis

- In a process known as *aerobic oxidation*, cells convert the energy released by the oxidation of glucose or fatty acids into synthesis of the terminal phosphoanhydride bond of ATP.
- Chemiosmosis is based on the interconversion of three forms of biological potential energy: chemical bond energy, chemical gradients across membranes, and electrical (voltage) gradients across membranes.
- The energy released as electrons travel down their electric potential gradient through an electron-transport chain generates a proton electrochemical gradient across a membrane located either in a mitochondrion, a chloroplast, or a bacterial plasma membrane.
- The energy stored in the proton electrochemical gradient is called the *proton-motive force*.
- When protons move back across the membrane down the electrochemical gradient, they release energy.
- The protons move back across the membrane through the ATP-synthesizing enzyme called *ATP synthase*, which harnesses the released energy to make ATP from ADP and P_i.

12.2 First Step of Harvesting Energy from Glucose: Glycolysis

In an automobile engine, hydrocarbon fuel is oxidatively and explosively converted in an essentially one-step process to mechanical work (i.e., driving a piston) plus the products CO_2 and H_2O. The process is relatively inefficient in that substantial amounts of the chemical energy stored in the fuel are wasted, as they are converted to unused heat, and substantial amounts of fuel are only partially oxidized and are sometimes released as toxic exhaust. In the competition to survive, organisms cannot afford to squander their often limited energy sources on an equivalently inefficient process and have therefore evolved a more efficient mechanism for converting fuel into work. That mechanism, known as *aerobic oxidation*, provides the following advantages:

- By dividing the energy conversion process into multiple steps that generate several energy-carrying intermediates, chemical bond energy is efficiently channeled into the synthesis of ATP, with little energy lost as heat.
- Different fuels (sugars and fatty acids) are reduced to common intermediates that can then share subsequent pathways for combustion and ATP synthesis.

Because the total energy stored in the bonds of the initial fuel molecules is substantially greater than that required to drive the

synthesis of a single ATP molecule ($\sim 7.3\,\mathrm{kcal/mol}$), many ATP molecules are produced for each fuel molecule consumed. An important feature of ATP production from the breakdown of nutrient fuels into CO_2 and H_2O is a set of reactions, called **respiration**, involving a series of oxidation and reduction reactions called an *electron-transport chain* (see Figure 12-1, *top*). The combination of these reactions with phosphorylation of ADP to form ATP is called **oxidative phosphorylation** and occurs in mitochondria in nearly all eukaryotic cells. When oxygen is available and is used as the final recipient of the electrons transported via the electron-transport chain, the respiratory process that converts nutrient energy into ATP is called *aerobic oxidation* or *aerobic respiration*. Aerobic oxidation is an especially efficient way to maximize the conversion of nutrient energy into ATP because O_2 is a relatively strong oxidant. If some molecule other than O_2 — for example, the weaker oxidants sulfate ($SO_4{}^{2-}$) or nitrate (NO^{3-}) — is the final recipient of the electrons in the electron-transport chain, the process is called **anaerobic respiration**. Anaerobic respiration is typical of some prokaryotic microorganisms. Although there are exceptions, most known multicellular (metazoan) eukaryotic organisms use aerobic oxidation to generate most of their ATP.

In our discussion of aerobic oxidation, we will be tracing the fate of the two main cellular fuels: sugars (principally glucose) and fatty acids. Under certain conditions — for example, starvation conditions

— amino acids also feed into these metabolic pathways. We first consider glucose oxidation, then turn to fatty acids.

The complete aerobic oxidation of one molecule of glucose yields 6 molecules of CO_2, and the energy released is coupled to the synthesis of as many as 30 molecules of ATP. The overall reaction is

$$C_6H_{12}O_6 + 6\,O_2 + 30\,P_i^{2-} + 30\,ADP^{3-} + 30\,H^+ \rightarrow 6\,CO_2 + 30\,ATP^{4-}$$

Glucose oxidation in eukaryotes takes place in four stages (see Figure 12-1, *top*):

Stage I: Glycolysis In the cytosol, one 6-carbon glucose molecule is converted by a series of reactions to two 3-carbon **pyruvate** molecules; a net of 2 ATPs are produced for each glucose molecule metabolized.

Stage II: Citric Acid Cycle In the **mitochondrion**, the oxidation of pyruvate, generated by glycolysis, to CO_2 is coupled to the generation of the high-energy electron carriers NADH and $FADH_2$, which store the energy for later use. These two carriers can be considered the sources of high-energy electrons.

Stage III: Electron-Transport Chain High-energy electrons flow down their electric potential gradient from NADH and $FADH_2$ to O_2 via membrane proteins that convert the energy released into a proton-motive force (H^+ electrochemical gradient). The energy

released from the electrons pumps protons across a membrane, thus generating the proton gradient.

Stage IV: ATP Synthesis The proton-motive force powers the synthesis of ATP as protons flow down their concentration and electrical (voltage) gradients through the ATP-synthesizing enzyme ATP synthase, which is embedded in a mitochondrial membrane. For each original glucose molecule, an estimated 28 additional ATPs are produced by this mechanism of oxidative phosphorylation (stages II–IV).

In this section, we discuss stage I: the biochemical pathways that break down glucose into pyruvate in the cytosol. We also discuss how these pathways are regulated, and we contrast the metabolism of glucose under anaerobic and aerobic conditions. The ultimate fate of pyruvate, once it enters mitochondria, is discussed in Section 12.5.

During Glycolysis (Stage I), Cytosolic Enzymes Convert Glucose to Pyruvate

Glycolysis, the first stage of glucose oxidation, occurs in the cytosol in both eukaryotes and prokaryotes; it does not require molecular oxygen (O_2) and is thus an anaerobic process. Glycolysis is an example of **catabolism**, the biological breakdown of complex substances into simpler ones. A set of 10 water-soluble cytosolic

enzymes catalyze the reactions constituting the *glycolytic pathway* (*glyco*, “sweet”; *lysis*, “split”), in which one molecule of glucose is converted to two molecules of pyruvate (Figure 12-3). Below we will consider some of the reasons why so many steps are used to efficiently convert glucose to pyruvate.

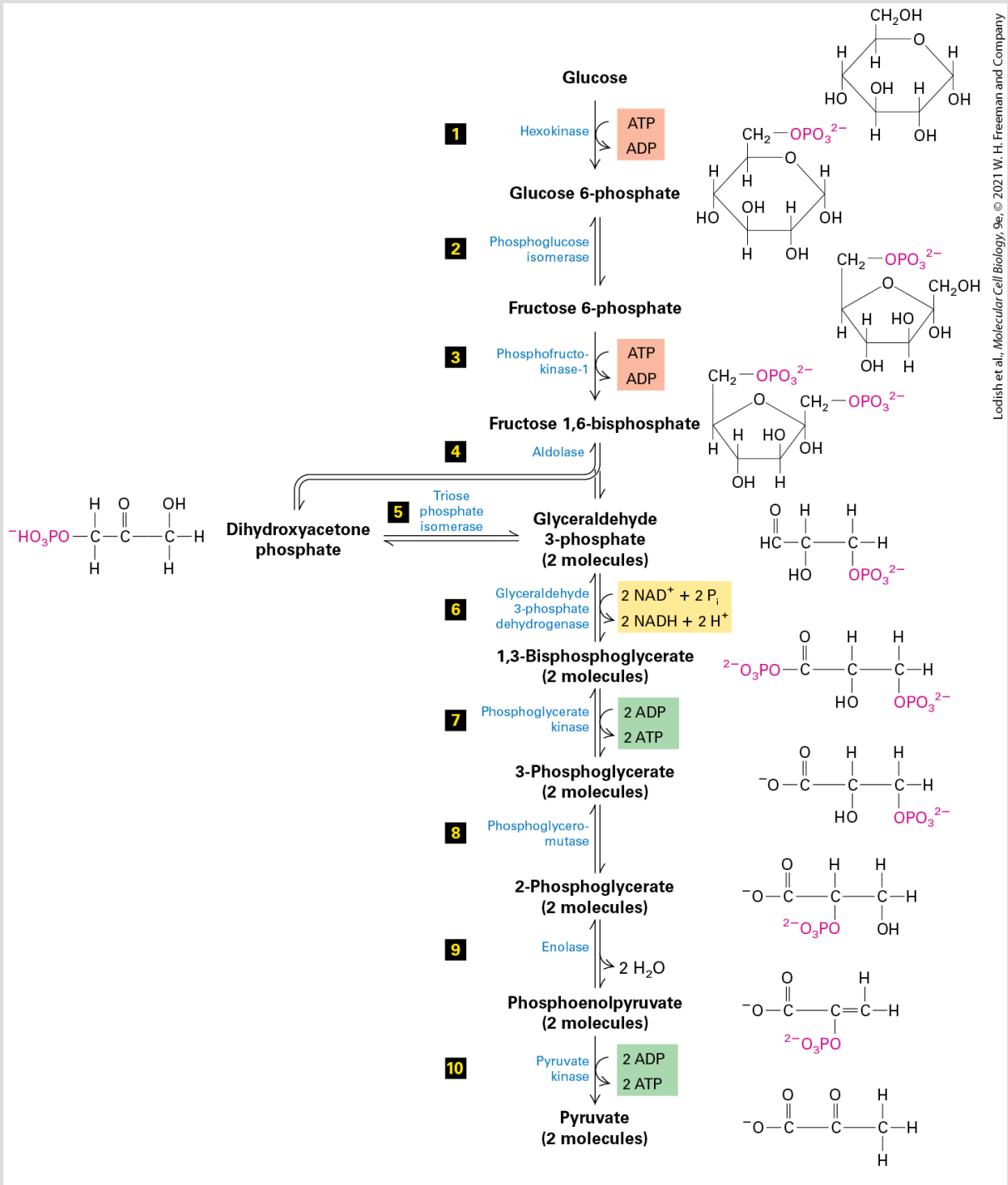
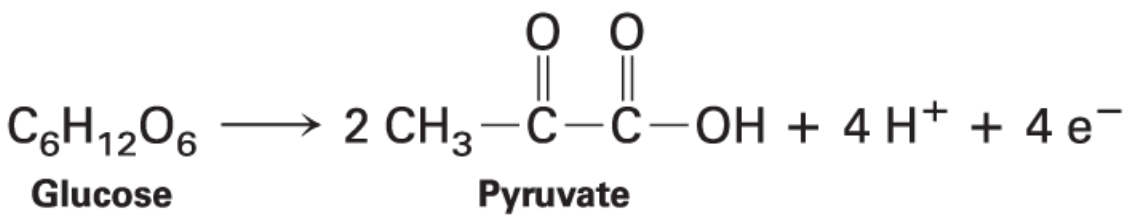


FIGURE 12-3 The glycolytic pathway. A series of 10 reactions degrades glucose to pyruvate. Two reactions consume ATP, forming ADP and phosphorylated sugars (red), two generate ATP from ADP by substrate-level phosphorylation (green), and one yields NADH by reduction of NAD^+ (yellow). Note that all the intermediates between glucose and pyruvate are

phosphorylated compounds. Steps **1**, **3**, and **10**, with single arrows, are essentially irreversible (have large negative ΔG values) under ordinary conditions in cells.

All the reaction intermediates in glycolysis produced by these enzymes are water-soluble, phosphorylated compounds called *metabolic intermediates*. In addition to chemically converting one glucose molecule into two pyruvates, the glycolytic pathway generates four ATP molecules by phosphorylation of four ADPs (steps **7** and **10**). ATP is formed directly through the enzyme-catalyzed joining of ADP with a P_i that is derived from phosphorylated metabolic intermediates; this process is called **substrate-level phosphorylation** (to distinguish it from the *oxidative phosphorylation* that generates ATP in stages III and IV). Substrate-level phosphorylation in glycolysis, which does not involve the use of a proton-motive force, requires the prior addition (in steps **1** and **3**) of two phosphates from two ATPs. By adding a little energy in the beginning of the pathway, glycolysis can recover more energy downstream. As a consequence, glycolysis yields the net production of only two ATP molecules per glucose molecule.

The balanced chemical equation for the conversion of glucose to pyruvate shows that four hydrogen atoms (four protons and four electrons) are also released:



(For convenience, we show pyruvate here in its un-ionized form, pyruvic acid, although at physiological pH it would be largely dissociated.) All four electrons and two of the four protons are transferred (see Figure 12-3, step **6**) to two molecules of the oxidized form of **nicotinamide adenine dinucleotide** ($\mathbf{NAD^+}$) to produce the reduced form of the coenzyme, NADH (see Figure 2-33a):

$$2\,H^+ + 4\,e^- + 2\,NAD^+ \rightarrow 2\,NADH$$

Later we will see that the energy carried by the high-energy electrons in NADH and the analogous electron carrier $FADH_2$, the reduced form of the coenzyme **flavin adenine dinucleotide (FAD)** (see Figure 2-33b), can be used to make additional ATPs via the electron-transport chain. The overall chemical equation for this first stage of glucose metabolism is

$$C_6H_{12}O_6 + 2\,NAD^+ + 2\,ADP^{3-} + 2\,P_i{}^{2-} \rightarrow 2\,C_3H_4O_3 + 2\,NADH +$$

Why are there so many enzymatic steps and so many metabolic intermediates in the conversion of one glucose to two pyruvates?

One reason is that some of the intermediates can be diverted to the biosynthesis of other molecules needed by cells, such as the conversion of dihydroxyacetone phosphate to glycerol. Another was mentioned previously: by dividing the process into multiple steps, small portions of the high chemical bond energy in glucose can be efficiently channeled in separate steps into the synthesis of ATP (steps **7** and **10**) and NADH (step **6**), with little energy lost as heat. In addition, multiple intermediate steps afford the opportunity to regulate the activities of these steps and related cellular processes either independently or coordinately, thus permitting fine tuning of the metabolic pathway.

After glycolysis, only a fraction of the energy available in glucose has been extracted and converted to ATP and NADH. The rest remains trapped in the covalent bonds of the two pyruvate molecules. The ability to efficiently convert the energy remaining in pyruvate to ATP depends on the presence of molecular oxygen. As we will see when we consider the electron-transport chain (Section 2.5), energy conversion is substantially more efficient under aerobic conditions than under anaerobic conditions.

The Rate of Glycolysis Is Adjusted to Meet the Cell's Need for ATP

To maintain appropriate levels of ATP, cells must control the rate of glucose catabolism. The operation of the glycolytic pathway (stage I), as well as the citric acid cycle (stage II), is continuously regulated, primarily by allosteric mechanisms (see Chapter 3 for general principles of allosteric control). Three allosteric enzymes involved in glycolysis play key roles in regulating the entire glycolytic pathway. *Hexokinase* (see Figure 12-3, step **1**) is inhibited by its reaction product, glucose 6-phosphate. Hexokinase is involved in the regulation of energy metabolism in at least two additional ways by acting as a glucose sensor. Hexokinase isoforms in mammals and plants, but not yeast, bind to and are active on the cytoplasmic surface of the **mitochondrial outer membrane**, apparently coordinating the activities of the cytoplasmic glycolytic system and mitochondrial activity, including respiration, and facilitating efficient, global ATP production. In mammals, plants, and yeast, some hexokinase isoforms can enter the nucleus and act as co-factors for transcription factors to control energy metabolism by regulating the synthesis of several key enzymes. The second allosteric enzyme is *Pyruvate kinase* (see Figure 12-3, step **10**), which is inhibited by ATP, so glycolysis slows down if too much ATP is present. The third enzyme, *phosphofructokinase-1* (see Figure 12-3, step **3**), is the principal rate-limiting enzyme of the glycolytic pathway; it is allosterically controlled by several molecules (Figure 12-4).

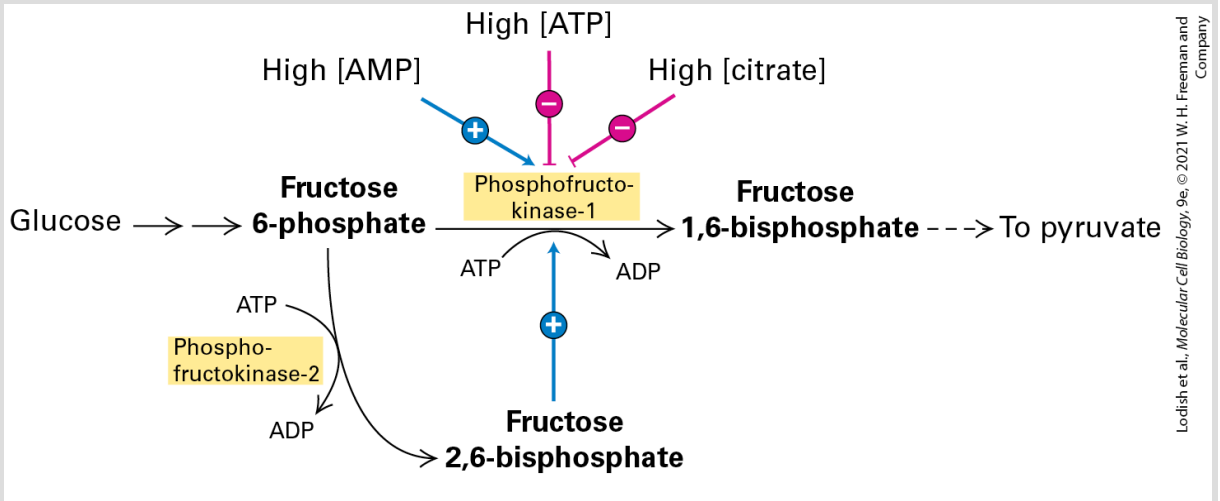


FIGURE 12-4 Allosteric regulation of glucose metabolism. The key regulatory enzyme in glycolysis, phosphofructokinase-1, is allosterically activated by AMP and fructose 2,6-bisphosphate, whose concentrations are elevated when the cell's energy stores are low. The enzyme is inhibited by ATP and citrate, both of which are elevated when the cell is actively oxidizing glucose to CO_2 (i.e., when energy stores are high).

For example, phosphofructokinase-1 is allosterically *inhibited* by ATP and allosterically *activated* by adenosine monophosphate (AMP). As a result, the rate of glycolysis is very sensitive to the cell's **energy charge**, a measure of the fraction of total adenosine phosphates that have "high-energy" phosphoanhydride bonds, which is equal to $[(\mathrm{ATP}) + 0.5(\mathrm{ADP})]/[(\mathrm{ATP}) + (\mathrm{ADP}) + (\mathrm{AMP})]$. The allosteric inhibition of phosphofructokinase-1 by ATP may seem unusual because ATP is also a substrate of this enzyme. But the affinity of the substrate-binding site for ATP is much higher (has a lower K_m) than that of the allosteric site. Thus at low concentrations, ATP binds to the catalytic site, but not to the inhibitory allosteric site, and enzymatic catalysis proceeds at near maximal rates. At high concentrations, ATP also binds to the

allosteric site, inducing a conformational change that reduces the affinity of the enzyme for its other substrate, fructose 6-phosphate, and thus reduces the rate of this reaction and the overall rate of glycolysis.

Another important allosteric activator of phosphofructokinase-1 is *fructose 2,6-bisphosphate*. This metabolite is formed from fructose 6-phosphate by an enzyme called *phosphofructokinase-2*. Fructose 6-phosphate accelerates the formation of fructose 2,6-bisphosphate, which in turn activates phosphofructokinase-1. This type of control is known as *feed-forward activation*, in which a high abundance of a metabolite (here, fructose 6-phosphate) accelerates its own subsequent metabolism. Fructose 2,6-bisphosphate allosterically activates phosphofructokinase-1 in liver cells by decreasing the inhibitory effect of high ATP concentrations and by increasing the affinity of phosphofructokinase-1 for one of its substrates, fructose 6-phosphate. In Chapter 21 we will see how the hormone insulin influences energy metabolism via fructose 2,6-bisphosphate. Phosphofructokinase-1 is also inhibited by high concentrations of citrate. Later in this chapter we will see that citrate is a product generated during stage II of glucose oxidation.

The three glycolytic enzymes that are regulated by allostery catalyze reactions with large negative $\Delta G^{\circ\prime}$ values — reactions that are essentially irreversible under ordinary conditions. These enzymes are therefore particularly suitable for regulating the entire glycolytic

pathway. Additional control is exerted by glyceraldehyde 3-phosphate dehydrogenase, which catalyzes the reduction of NAD^+ to NADH (see Figure 12-3, step **6**). As we shall see, NADH is a high-energy electron carrier used subsequently during oxidative phosphorylation in mitochondria. If cytosolic NADH builds up owing to a slowdown in mitochondrial oxidation, step **6** becomes thermodynamically less favorable.

Glucose metabolism is controlled differently in various mammalian tissues to meet the metabolic needs of the organism as a whole. For instance, when glucose is not available in the diet, it is necessary for the liver to provide that glucose by releasing it into the bloodstream. To do this, the liver converts the polymer glycogen, a storage form of glucose (see Chapter 2), directly to glucose 6-phosphate (without involvement of hexokinase, see Figure 12-3, step **1**). Under these conditions, there is a reduction in fructose 2,6-bisphosphate levels and decreased phosphofructokinase-1 activity (see Figure 12-4). As a result, glucose 6-phosphate derived from glycogen is not metabolized to pyruvate; rather, it is converted to glucose by a phosphatase and released into the blood to nourish the brain and red blood cells, which depend primarily on glucose for their energy. In all cases, the activity of these regulated enzymes is controlled by the level of small-molecule metabolites, generally by allosteric interactions, or by hormone-mediated phosphorylation and dephosphorylation reactions. (Chapters 15 and 21 provide more

detailed discussions of hormonal control of glucose metabolism in liver and muscle.)

Glucose Is Fermented When Oxygen Is Scarce

Many eukaryotes, including humans, are *obligate aerobes*: they thrive only in the presence of molecular oxygen and can metabolize glucose (or related sugars) completely to CO_2, with the concomitant production of a large amount of ATP. Most eukaryotes, however, can generate some ATP by anaerobic metabolism. A few eukaryotes are *facultative anaerobes*: they grow in either the presence or the absence of oxygen. Annelids (segmented worms), mollusks, and some yeasts, for example, can survive without oxygen, relying on the ATP produced by fermentation.

In the absence of oxygen, yeasts convert the pyruvate produced by glycolysis to one molecule each of ethanol and CO_2; in these reactions, two NADH molecules are oxidized to NAD^+ for every two pyruvates converted to ethanol, thereby regenerating the supply of NAD^+, which is necessary for glycolysis to continue (Figure 12-5a, *left*). This anaerobic catabolism of glucose, called **fermentation**, is the basis of beer and wine production.

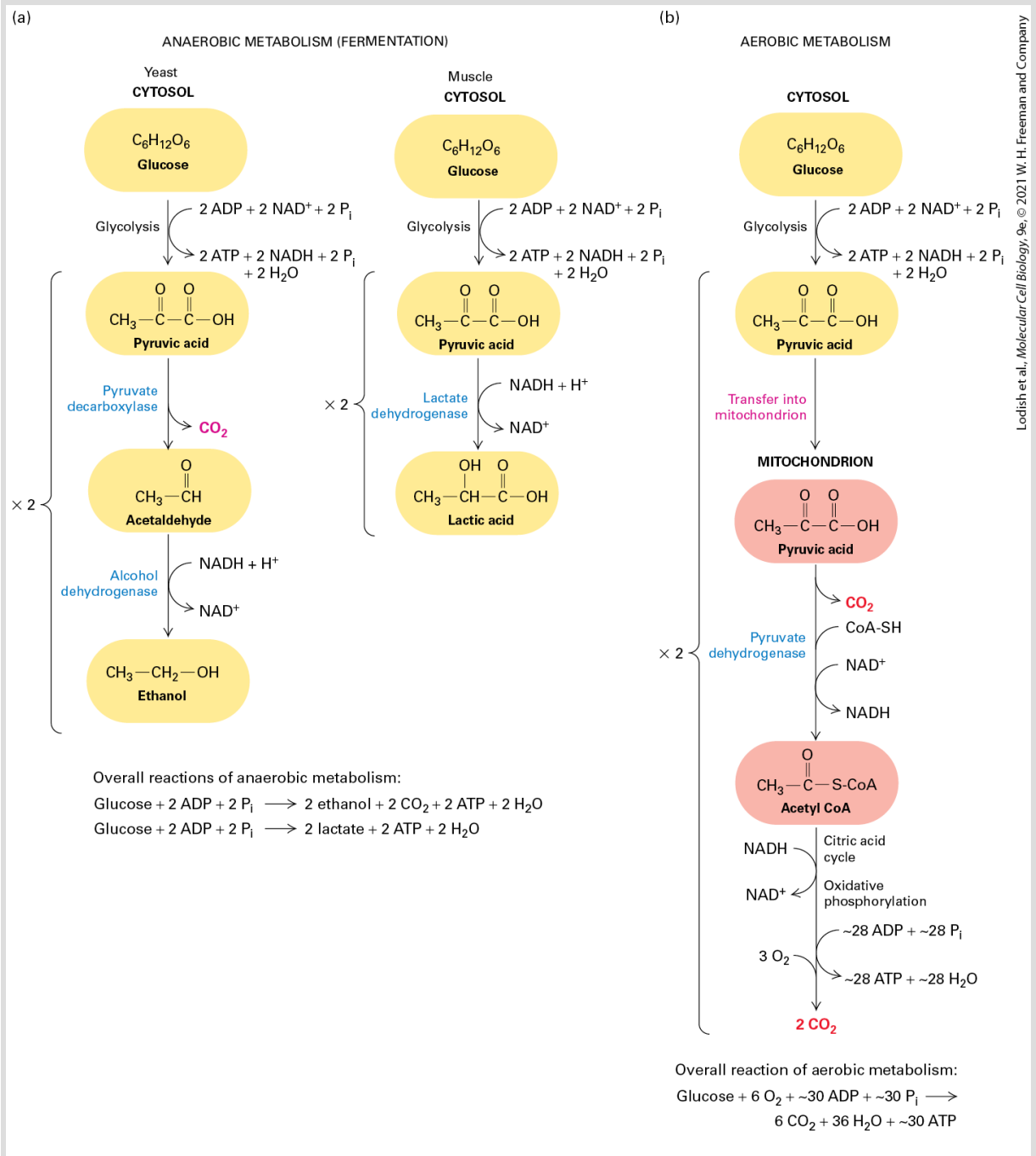


FIGURE 12-5 Anaerobic versus aerobic metabolism of glucose. The ultimate fate of pyruvate formed during glycolysis depends on the presence or absence of oxygen. (a) In the absence of oxygen, pyruvate is only partially degraded and no further ATP is made. However, two electrons are transferred from each NADH molecule produced during glycolysis to an acceptor molecule to regenerate NAD^+, which is required for continued glycolysis. In yeast (*left*), acetaldehyde is the electron acceptor and ethanol is the product. This process is called *alcoholic fermentation*. When oxygen is scarce in muscle cells (*right*), NADH reduces pyruvate

to form lactic acid, regenerating NAD^+, a process called *lactic acid fermentation*. (b) In the presence of oxygen, pyruvate is transported into mitochondria, where it is first converted by pyruvate dehydrogenase into one molecule of CO_2 and one of acetic acid, the latter linked to coenzyme A (CoA-SH) to form acetyl CoA, concomitant with reduction of one molecule of NAD^+ to NADH. Further metabolism of acetyl CoA and NADH generates approximately an additional 28 molecules of ATP per glucose molecule oxidized.

Fermentation also occurs in animal cells, although lactic acid, rather than alcohol, is the product. During prolonged contraction of mammalian skeletal muscle cells — for example, during exercise — oxygen can become scarce within the muscle tissue. As a consequence, glucose catabolism is limited to glycolysis, and muscle cells convert two pyruvates to two molecules of lactic acid by a reduction reaction that also oxidizes two NADHs to two NAD^+s (Figure 12-5a, *right*). A monocarboxylic transporter can then transport the lactate out of the cell. Once it is secreted into the blood, some of the lactic acid passes into the liver, where it is reoxidized to pyruvate and either further metabolized to CO_2 aerobically or converted back to glucose. Much lactate is metabolized to CO_2 by the heart, which is highly perfused by blood and can continue aerobic metabolism at times when exercising, oxygen-poor skeletal muscles secrete lactate. If too much lactic acid accumulates in the blood, the acid causes an unhealthy decrease in the pH of the blood (lactic acidosis). Lactic acid bacteria (the organisms that spoil milk) and other prokaryotes also generate ATP by the fermentation of glucose to lactic acid.

Fermentation is a much less efficient way to generate ATP than aerobic oxidation and therefore occurs in animal cells only when oxygen is scarce. In the presence of oxygen, pyruvate formed by glycolysis is transported into mitochondria, where it is oxidized by O_2 to CO_2 and H_2O via the series of reactions outlined in Figure 12-5b. This aerobic metabolism of glucose, which occurs in stages II–IV of the process outlined in Figure 12-1, generates an estimated 28 additional ATP molecules per original glucose molecule, far outstripping the ATP yield from anaerobic glucose metabolism (fermentation).

To understand how ATP is generated so efficiently by aerobic oxidation, we must consider first the structure and function of the organelle responsible, the mitochondrion. The structure, dynamics, and functions of mitochondria are the subjects of the next two sections.

KEY CONCEPTS OF SECTION 12.2

First Step of Harvesting Energy from Glucose: Glycolysis

- In a process known as *aerobic oxidation*, cells convert the energy released by the oxidation of glucose or fatty acids into formation of the terminal phosphoanhydride bond of ATP.
- The complete aerobic oxidation of each molecule of glucose produces 6 molecules of CO_2 and approximately 30 ATP molecules. The entire process, which starts in the cytosol and is completed in the mitochondrion, can be divided into four stages: (I) degradation of glucose to pyruvate in the cytosol (glycolysis); (II) pyruvate oxidation to CO_2 in the mitochondrion coupled to generation of the high-energy electron carriers NADH and $FADH_2$ (via the citric acid cycle); (III) electron transport to generate a

proton-motive force together with conversion of molecular oxygen to water; and (IV) ATP synthesis (see Figure 12-1). From each glucose molecule, two ATPs are generated by stage I and approximately 28 from stages II–IV.

- In glycolysis (stage I), cytosolic enzymes convert glucose to two molecules of pyruvate and generate two molecules each of NADH and ATP (see Figure 12-3).
- The rate of glucose oxidation via glycolysis is regulated by the inhibition or stimulation of several enzymes, depending on the cell's need for ATP (see Figure 12-4). Glucose is stored, for example, as glycogen, when ATP is abundant.
- In the absence of oxygen (anaerobic conditions), cells can metabolize pyruvate to lactic acid or (in the case of yeast) to ethanol and CO_2, in the process of converting NADH back to NAD^+, which is necessary for continued glycolysis. In the presence of oxygen (aerobic conditions), pyruvate is transported into the mitochondrion, where it is metabolized to CO_2, in the process generating abundant ATP (see Figure 12-5).

12.3 The Structure of Mitochondria

Oxygen-producing photosynthetic cyanobacteria first appeared about 2.7 billion years ago. The subsequent buildup of oxygen in Earth's atmosphere during the next approximately 1 billion years opened the way for organisms to evolve the very efficient aerobic oxidation pathway, which in turn permitted the evolution, especially during what is called the *Cambrian explosion*, of large and complex body forms and associated metabolic activities. In eukaryotic cells, aerobic oxidation is carried out by mitochondria (stages II–IV in Figure 12-1). In effect, mitochondria are ATP-generating factories, taking full advantage of this plentiful oxygen. We first describe their structure, dynamics, and diverse functions that extend well beyond making ATP, and then the reactions they employ to degrade pyruvate and make ATP.

Mitochondria Are Abundant, Multifunctional Organelles

An individual mitochondrion (Figure 12-6a) is about the size of an *Escherichia coli* bacterium, which is not surprising because, as we shall see below, bacteria are thought to be the evolutionary precursors of mitochondria. Most eukaryotic cells contain many mitochondria, which may be fused to one another into large networks and may collectively occupy as much as 25 percent of the

volume of the cytoplasm. The numbers of mitochondria in a cell — hundreds to a few thousands in most mammalian cells, as many as half a million per cell in mammalian eggs — are typically regulated to match the cell's requirements for ATP (e.g., specialized stomach cells that use a lot of ATP for acid secretion have many mitochondria).

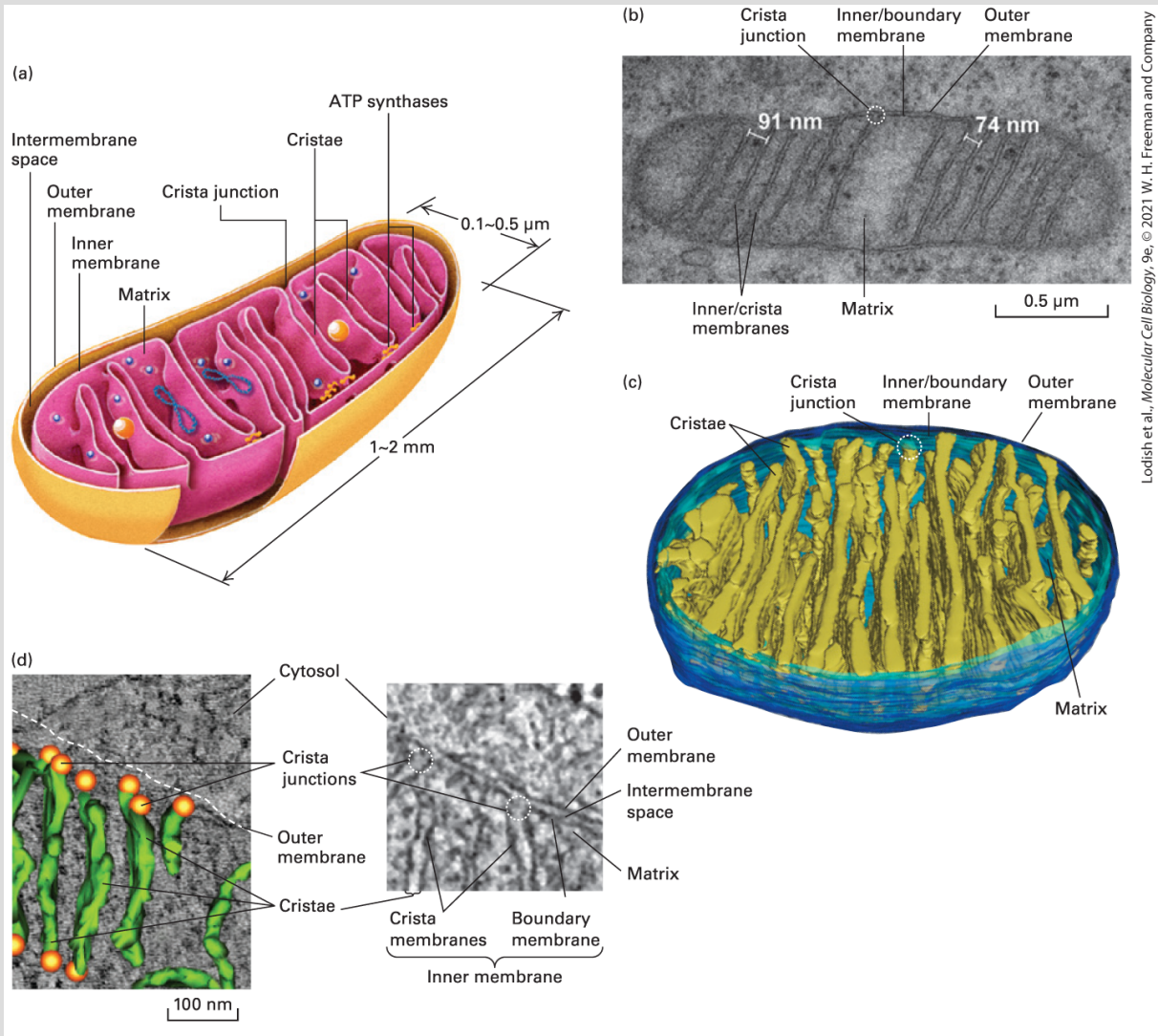


FIGURE 12-6 Internal structure of a mitochondrion. (a) Schematic diagram showing the principal mitochondrial membranes and compartments. The smooth outer membrane forms the outside boundary of the mitochondrion. The inner membrane is apparently a

single continuous membrane that has three distinct domains: boundary membrane, cristae, and crista junctions. The boundary membrane is flat and lies immediately below and adjacent to the outer membrane. The cristae are tubelike or flat, pancake-like invaginations that extend from the boundary membrane into the center of the mitochondrion. The sharp bends that form the connection between the boundary membrane and the cristae are called *crista junctions*. The intermembrane space is that fluid-filled space between the inner and outer membranes. The matrix is the fluid-filled space within the inner membrane (on the opposite side of the inner membrane from the intermembrane space). The ATP synthases (also called F_0F_1 *complexes*, small yellow spheres), which synthesize ATP, are aligned in rows as dimers at the bends in the cristae membrane (described later in Figure 12-34e) and protrude into the matrix. The matrix contains the mitochondrial DNA (blue strands), ribosomes (small blue spheres), and granules (large yellow spheres). (b) Transmission electron micrograph of a thin section of a human cell line — HeLa cells — showing a mitochondrion's inner and outer membranes with cristae in parallel, ladder-like stacks. One crista junction is highlighted with a white dotted circle. The ladder-like stacks of cristae can also be visualized using superresolution fluorescence microscopy (see Figure 12-8). (c) Computer-generated model of a section of a mitochondrion from chicken brain. This model is based on a three-dimensional electron microscopic image calculated from a series of two-dimensional electron micrographs (such as that in panel b) recorded at regular intervals. This technique is analogous to a three-dimensional x-ray tomogram or CAT scan used in medical imaging. Note the tightly packed cristae (yellow-green), the inner membrane (light blue), the outer membrane (dark blue) and the crista junction (white dotted circle). (d) Cristae and crista junctions from human fibroblasts were visualized and modeled using electron microscopy and tomography. The right panel shows one of the multiple sections through the mitochondrion imaged using transmission electron microscopy, with the mitochondrial membranes clearly distinguished. The sharp bends in the inner membrane at the crista junctions (dotted circles) that separate the crista membranes from the inner boundary membrane are seen clearly. The left panel shows a three-dimensional tomographic model of the laminar cristae seen edge on (green) and crista junctions (orange spheres) overlaid with the EM image.

[Part (b) T. Stephan et al., 2019, "Live-Cell STED Nanoscopy of Mitochondrial Cristae." *Sci. Rep.* **9**:12419 (2019). https://doi.org/10.1038/s41598-019-48838-2. Part (c) T. G. Frey and G. A. Perkins. Part (d) D. C. Jans et al., 2013, "STED Super-Resolution Microscopy Reveals an Array of MINOS Clusters Along Human Mitochondria," *Proc. Natl. Acad. Sci. USA* **110**:8936–8941, Fig. 6, Fig. 6A and C.]

Although our focus in this chapter is on the structure of mitochondria and the mechanisms by which they generate ATP, it is important to recognize that mitochondria participate in a wide variety of other critical cellular processes (Table 12-1). Mitochondria are involved with the biosynthesis of many small molecules. In several cases, some steps in these biosynthetic pathways take place within the mitochondria and some outside (extramitochondrial steps), requiring precursors and products to be shuttled into and out of the mitochondria via transport proteins. For example, in many eukaryotes, mitochondria host the first rate-limiting step in heme biosynthesis. The product of this first step is then exported for additional modification in the cytoplasm, producing an intermediate that must then be imported into the mitochondria for the terminal steps. The biosynthesis in the cytoplasm of a wide variety of small molecules depends on mitochondria. Mitochondria provide small organic molecules to the cytoplasm — for example, citrate, isocitrate, malate, formate, and α-ketoglutarate, that can be used to generate NADPH (an energy source), acetyl CoA (a carbon source), or other precursors for the extramitochondrial biosynthesis of molecules such as glutathione, purines, fatty acids, and cholesterol.

TABLE 12-1 • Multiple Functions of Mitochondria

Biosynthesis or processing of small molecules

Fatty acids
Steroid hormones
Pyrimidines
Iron-sulfur clusters
Heme
Phospholipids (phosphatidylethanolamine, phosphatidylglycerol, cardiolipin)
Ubiquinone
Amino acids (synthesis, interconversion, and catabolism)

Other mitochondrial functions
Oxidative phosphorylation and ATP synthesis Reactive oxygen species (ROS) homeostasis Ion homeostasis (e.g., calcium) Ammonia detoxification Fatty acid oxidation Thermogenesis (heat generation) in brown fat Contributions to innate immunity and inflammation Regulated cell death pathways (e.g., apoptosis)

Cellular processes influenced by mitochondrial-ER membrane contact sites (MCS)
Mitochondrial shape and dynamics (fusion, fission, autophagy) PINK1/Parkin-dependent mitophagy (initiated at MCSs) Calcium transport into mitochondria (see Figure 15-29) Calcium homeostasis and calcium-mediated signaling Glucose and energy metabolism Mitochondrial import from the ER of lipids, including phosphatidylserine and possibly cholesterol Mitochondrial biosynthesis of lipids, including phosphatidylethanolamine and steroid hormones Responses to stress Cell survival via regulated cell death (see Chapter 22) Inflammatory responses via the inflammasome and innate immune responses (see Chapter 24) Pathways implicated in viral infections (cytomegalovirus, hepatitis C virus) Neurodegenerative pathology (Alzheimer's and Parkinson's diseases)

Mitochondria Have Two Structurally and Functionally Distinct Membranes

The details of mitochondrial structure can be observed with electron microscopy (Figures 1-20 and 12-6) and superresolution fluorescence microscopy (Figure 12-8). Each mitochondrion has two distinct, concentric membranes: the **mitochondrial inner membrane** and **mitochondrial outer membrane**. The outer mitochondrial membrane defines the smooth outer perimeter. The inner mitochondrial membrane lies underneath the outer membrane. The inner mitochondrial membrane is a single continuous membrane that itself can be considered to have three compositionally and structurally distinct domains. The *boundary membrane* (or inner boundary membrane) is the flat inner mitochondrial membrane that lies immediately inside and adjacent to the outer membrane. The **mitochondrial cristae** are the numerous invaginations of the inner membrane that extend from the boundary membrane at the perimeter into the center of the mitochondrion. The invaginating cristae greatly expand the surface area of the inner mitochondrial membrane. As we shall see below, the cristae's membranes are the site of ATP synthesis — thus the invaginations increase the mitochondrion's capacity to generate ATP. The connection between the inner boundary membrane and a crista is called a *crista junction* (Figure 12-6).

The lengths of the cristae and their structures (which may be tubular in shape or flat and pancake-like) can vary within a mitochondrion. They tend to form stacks that in cross section look like the steps of a ladder (Figure 12-6b and c, also see Figures 1-21 and 12-8). There can be regions within the mitochondria that have densely stacked cristae and regions that have less dense cristae or are even devoid of cristae. The crista junctions and the edges and tips of the cristae are highly curved. The curvature of the crista junctions (see Figures 12-6 and 12-30e) is due to a protein complex called *MICOS* (*mi*tochondrial *co*ntact *s*ite and cristae organizing system), which has an integral membrane protein subunit that homo-oligomerizes and bends the inner membrane to produce high curvature. MICOS also mediates close juxtaposition of the outer membrane and inner membrane by binding to outer membrane–associated proteins. The binding partners for MICOS in the outer membrane include the Tom and SAM protein complexes (these complexes, discussed in Chapter 13, are involved in import of proteins into the mitochondria). MICOS is necessary to maintain normal structures of cristae in mitochondria. Additionally, MICOS appears to function as a diffusion barrier to prevent mixing of the distinct proteins and lipids in the boundary membrane and crista membranes. We will discuss the molecular basis of the curvature at the edges and tips of the cristae in Section 12.6.

The outer and inner membranes topologically define two fluid-filled submitochondrial compartments: the **mitochondrial**

intermembrane space, between the outer and inner membranes, and the **mitochondrial matrix**, or central compartment, which forms the lumen within the inner membrane (see Figure 12-6a). Many of the proteins directly involved with transforming the energy of nutrients into the energy stored in ATP, such as the proteins of the electron-transport chain and ATP synthase, are located in the inner mitochondrial membrane. In typical liver mitochondria, for example, the surface area of the inner membrane, including cristae, is about five times that of the outer membrane. In fact, the total area of all inner mitochondrial membranes in liver cells is about 17 times that of the plasma membranes. The mitochondria in heart and skeletal muscle contain three times as many cristae as are found in typical liver mitochondria — presumably reflecting the greater demand for ATP by muscle cells.

Over a thousand different types of polypeptides are required to make and maintain mitochondria and permit them to function. Defective functioning of mitochondria-associated proteins — due, for example, to inherited genetic mutations — leads to over 250 human diseases. The most common of these are electron-transport chain diseases, which result from mutations in any one of 150 genes and exhibit a very wide variety of clinical abnormalities affecting muscles, the heart, the nervous system, and the liver, among other physiological systems. Other mitochondria-associated diseases include Miller syndrome, which results in multiple anatomic malformations, and connective tissue defects.

The most abundant protein in the outer mitochondrial membrane is a mitochondrial β-barrel **porin** called *VDAC (v*oltage-*d*ependent *a*nion *c*hannel), a multifunctional transmembrane channel protein that is similar in structure to bacterial porins (see Figure 10-18). Ions and most small hydrophilic molecules (up to about 5000 Da) can readily pass through these channel proteins when they are open. Although there may be metabolic regulation of the opening of mitochondrial porins and thus of the flow of metabolites across the outer membrane, the inner membrane is the major permeability barrier between the cytosol and the mitochondrial matrix, controlling the rate of mitochondrion-mediated oxidation and ATP generation.

Proteins constitute 76 percent of the total mass of the inner mitochondrial membrane — a higher fraction than in any other cellular membrane. Many of these proteins are key participants in oxidative phosphorylation. They include ATP synthase, proteins responsible for electron transport, and a wide variety of transport proteins that permit the movement of metabolites between the cytosol (via the intermembrane space) and the mitochondrial matrix. The human genome encodes 48 members of one family of mitochondrial transport proteins. One of these, the ADP/ATP carrier, is an antiporter that moves newly synthesized ATP out of the matrix and into the inner membrane space (and subsequently the cytosol) in exchange for ADP originating from the cytosol. Without this essential antiporter, the energy trapped in the chemical bonds

of mitochondrial ATP made in the matrix would not be available to the rest of the cell.

Keep in mind that plants, as well as animals, have mitochondria and perform aerobic oxidation. In plants, stored carbohydrates, mostly in the form of starch, are hydrolyzed to glucose. Glycolysis then produces pyruvate that is transported into mitochondria, as in animal cells. Mitochondrial oxidation of pyruvate and concomitant formation of ATP occur in photosynthetic cells during dark periods when photosynthesis is not possible, and in roots and other non-photosynthetic tissues at all times. We will discuss photosynthesis in chloroplasts in Section 12.7.

The inner mitochondrial membrane and matrix are the sites of most reactions involved in the oxidation of pyruvate and fatty acids to CO_2 and H_2O and the coupled synthesis of ATP from ADP and P_i (see Figure 12-13). Each of these reactions occurs in a discrete membrane or space in the mitochondrion.

Mitochondria Contain DNA and Evolved from a Single Endosymbiotic Event Involving Alphaproteobacterium

Although the vast majority of DNA in most eukaryotes is found in the nucleus, some DNA is present within the mitochondria of

animals, plants, and fungi and within the chloroplasts of plants. Many lines of evidence indicate that mitochondria and chloroplasts evolved from bacteria that were engulfed into ancestral cells, forming **endosymbionts** (Figure 12-7). Over evolutionary time, most of the bacterial genes in the endosymbiont bacteria were lost from the organelles' DNA — some were lost because their function was replicated by nuclear-encoded genes and others were simply transferred to nuclear chromosomes. However, mitochondria and chloroplasts in today's eukaryotes retain DNA encoding several proteins essential for organelle function as well as the ribosomal and transfer RNAs required for the intramitochondrial and intrachloroplast synthesis of those proteins. Thus eukaryotic cells have multiple genetic systems: a predominant nuclear system and secondary systems with their own DNA, ribosomes (see Figure 3-49), and tRNAs in mitochondria and chloroplasts. As detailed in Chapter 13, the majority of proteins inside mitochondria and chloroplasts are encoded by nuclear DNA, synthesized on cytosolic ribosomes, and subsequently imported into the organelle.

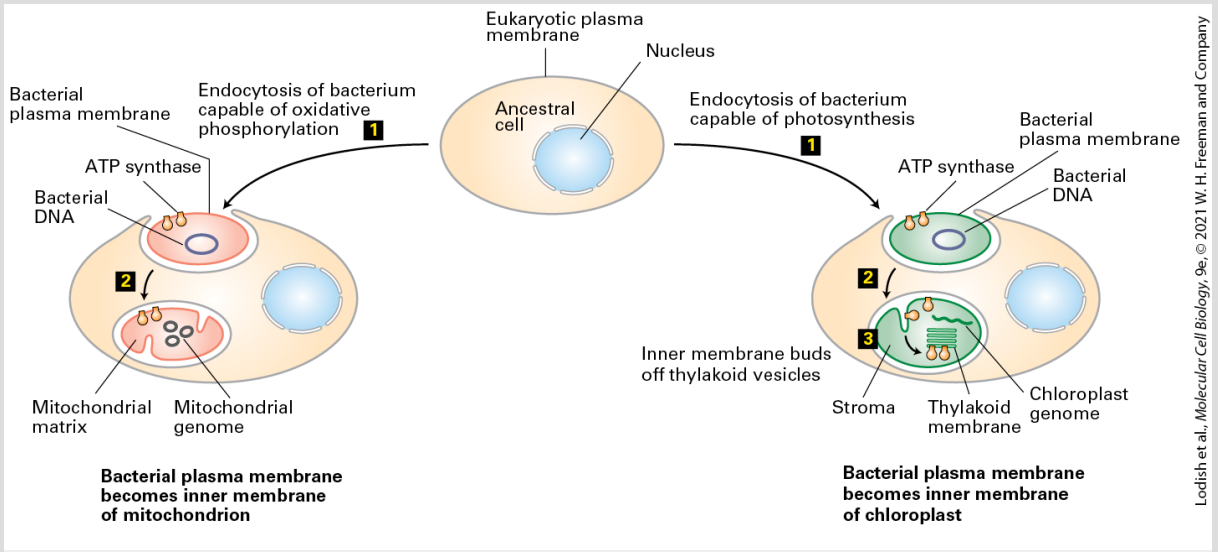


FIGURE 12-7 Endosymbiont hypothesis for the evolutionary origin of mitochondria and chloroplasts. Endocytosis of a bacterium by an ancestral, precursor eukaryotic cell (step 1) would generate an organelle with two membranes, the outer membrane derived from the precursor's plasma membrane and the inner one from the bacterial membrane (step 2). There is uncertainty as to whether the ancestral cell already contained a nucleus (as shown here) or developed a nucleus concomitant with or subsequent to engulfing the bacterial precursor to the mitochondrion (left pathway). Proteins localized to the ancestral bacterial membrane would retain their orientation, such that the portion of the protein once facing the extracellular space would now face the intermembrane space. For example, the F_1 subunit of ATP synthase, localized to the cytosolic face of the bacterial membrane, would face the matrix of the evolving mitochondrion (*left*) or chloroplast (*right*). Budding of vesicles from the inner chloroplast membrane, such as occurs during development of chloroplasts in contemporary plants, would generate the thylakoid membranes with the F_1 subunit remaining on the cytosolic face, facing the chloroplast stroma (step 3). The organelle DNAs are indicated. Membrane surfaces facing a shaded area are cytosolic faces; surfaces facing an unshaded area are exoplasmic faces.

Analysis of mitochondrial DNA (mtDNA) sequences from multiple eukaryotes provides strong support for the idea that the mitochondrion had a single origin, most likely a bacterial symbiote whose closest contemporary relatives are alphaproteobacteria (also

called α purple, see Figure 1-1). Because these are obligate intracellular parasites, the ancestor of mitochondria may have had an intracellular lifestyle that facilitated its evolution into an intracellular symbiote.

In organisms whose mtDNA includes only a limited number of genes, the same set of mitochondrial genes is retained, regardless of the phyla that include these organisms. There are several hypotheses for why these genes were never successfully transferred to the nuclear genome: their encoded polypeptides are too hydrophobic to be imported from the cytosol by crossing the outer mitochondrial membrane, the large size of rRNAs may interfere with their import into mitochondria, and differential regulation of their expression in response to conditions within individual mitochondria may be advantageous. If these genes were located in the nucleus, it would be difficult to tailor the regulation of their expression to accommodate specifically the needs of only one or a few mitochondria whose status required an alteration in gene expression.

The Size, Structure, and Coding Capacity of mtDNA Vary Considerably Among Organisms

Surprisingly, the size of the mtDNA, the number and nature of the proteins it encodes, and even the mitochondrial genetic code itself can vary greatly between different organisms. The mtDNAs of most multicellular animals are approximately 16-kb circular molecules that encode intron-less genes compactly arranged on both DNA strands. Vertebrate mtDNAs encode the two rRNAs found in mitochondrial ribosomes, the 22 tRNAs used to translate mitochondrial mRNAs, and 13 proteins involved in electron transport and ATP synthesis. The smallest mitochondrial genomes in *Plasmodium*, intracellular parasites that cause malaria in humans (Figure 1-25) are only about 6 kb, encoding three proteins and the mitochondrial rRNAs.

Plant mitochondrial genomes are many times larger than those of metazoans, in part due to the presence of non-protein-coding introns, pseudogenes, mobile DNA elements, and pieces of foreign (chloroplast, nuclear, and viral) DNA, evidence of many DNA changes during evolution of these plants. For instance, *Arabidopsis thaliana*, a member of the mustard weed family, has 366 kb of mtDNA and cucurbit plants (e.g., melon and cucumber) have about 2 Mb of mtDNA.

Several proteins encoded by mtDNA in some species are encoded by nuclear DNA in other, closely related species. A striking example involves the *coxII* gene, which encodes subunit 2 of cytochrome *c* oxidase in the mitochondrial electron-transport chain (described in

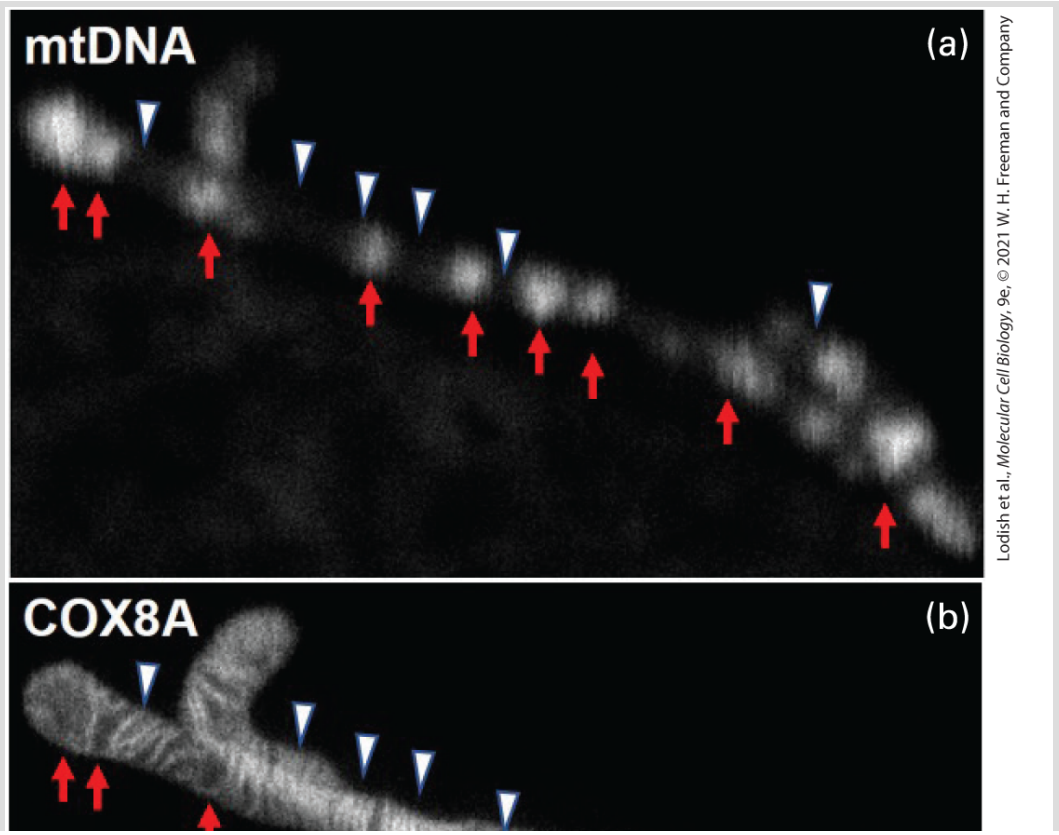
detail below). This gene is found in mtDNA in all multicellular plants studied except for certain related species of legumes, including the mung bean and the soybean, in which the *coxII* gene is nuclear.

Most mtDNA encoded and intramitochondrially synthesized proteins identified thus far are subunits of multimeric complexes used in electron transport or ATP synthesis that we will describe shortly. Thus most of the mitochondrial proteome, such as those proteins involved in the processes listed in Table 12-1, are encoded by nuclear genes, synthesized on cytosolic ribosomes, and imported into the organelle by processes discussed in Chapter 13.

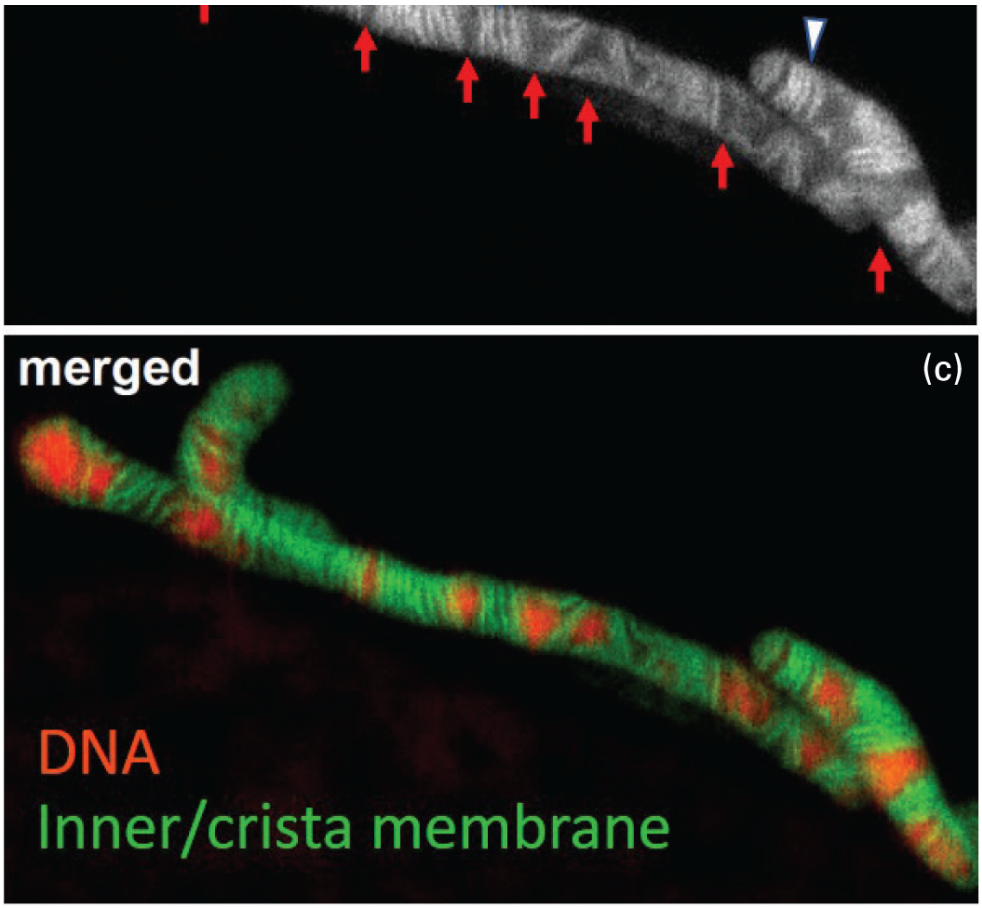
Mitochondrial DNA Is Located in the Matrix and Transferred During Mitosis to Daughter Cells by Cytoplasmic Inheritance

The mitochondrial DNA (mtDNA) is located in the mitochondrial matrix, and in the rapidly growing cells shown in Figure 12-8, the mtDNA is often seen in regions devoid of densely stacked cristae. Replication of mtDNA and division of the mitochondria can be followed in live cells using time-lapse microscopy. Such studies show that in most organisms, mtDNA replicates throughout interphase. At mitosis, each daughter cell receives approximately the same number of mitochondria from the cytoplasm of the

parental cell (a processed called *cytoplasmic inheritance*), but because there is no mechanism for apportioning exactly equal numbers of mitochondria to the daughter cells, some cells contain more mtDNA than others. If a mutation occurs in one mtDNA molecule in a cell, after cell division one daughter cell but not the other might inherit the mutant mtDNA molecule, but most of the mtDNA molecules in that cell would not carry the mutation. The total amount of mtDNA in a cell depends on the number of mitochondria, the size of the mtDNA molecules, and the number of mtDNA molecules per mitochondrion. Each of these parameters varies greatly between cell types. A typical human cell has about 1000–2000 mtDNA molecules; however, a human egg has about 500,000 mtDNA molecules, and a sperm has only about 100.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company



EXPERIMENTAL FIGURE 12-8 Dual staining reveals multiple mitochondrial DNA molecules in the mitochondrial matrix. Live human HeLa cells expressing a fluorescently labeled mitochondrial inner membrane protein (COX8A) that preferentially localizes in the cristae were treated with the double-stranded DNA binding fluorescent dye PicoGreen and then examined using (a) confocal microscopy to visualized the stained mitochondrial DNA and (b) superresolution microscopy to visualize the inner/cristae membranes (COX8A fluorescence). White arrowheads and red arrows highlight regions in this single mitochondrion with either little or intense mtDNA staining, respectively, in panel (a), and corresponding regions with either relatively dense stacks of cristae or regions relatively devoid of cristae stacks in panel (b). In general, the mtDNA in the matrix was concentrated in regions relatively devoid of dense cristae stacks. (c) Merged image of DNA staining (red) and Cox8A fluorescence (inner/crista membrane, green).

[T. Stephan et al., 2019, "Live-Cell STED Nanoscopy of Mitochondrial Cristae." *Sci. Rep.* **9**:12419 (2019). https://doi.org/10.1038/s41598-019-48838-2.]

The dramatic difference in the amount of mtDNA per cell in eggs and sperm has a profound effect on the inheritance of mtDNA. In mammals and most other multicellular organisms, the sperm contributes little (if any) cytoplasm to the zygote, and virtually all the mitochondria in the embryo are derived from those in the egg, not the sperm. Studies in mice have shown that 99.99 percent of mtDNA is maternally inherited, but a small part (0.01 percent) is inherited from the male parent. In higher plants, mtDNA is inherited exclusively in a uniparental fashion through the female parent (egg), not the male (pollen).

Products of Mitochondrial Genes Are Not Exported

Most, if not all, RNA transcripts of mtDNA and their translation products remain in the mitochondrion in which they are produced, and all mtDNA-encoded proteins are synthesized on mitochondrial ribosomes. Mitochondrial DNA encodes the rRNAs of mitochondrial ribosomes, although most of the ribosomal proteins are imported from the cytosol. In animals and fungi, all the tRNAs used for protein synthesis in mitochondria are also encoded by mtDNAs. However, in plants and many protozoans, most mitochondrial tRNAs

are encoded by the nuclear DNA and imported into the mitochondrion.

The structure of the human mitochondrial ribosome is shown in Figure 3-49. Reflecting the bacterial ancestry of mitochondria, mitochondrial ribosomes resemble bacterial ribosomes and differ from eukaryotic cytosolic ribosomes in their RNA and protein compositions, their size, and their sensitivity to certain antibiotics (see Table 5-2). For instance, chloramphenicol blocks protein synthesis by bacterial and mitochondrial ribosomes from most organisms, but cycloheximide, which inhibits protein synthesis on eukaryotic cytosolic ribosomes, does not affect mitochondrial ribosomes. This sensitivity of mitochondrial ribosomes to the important aminoglycoside class of antibiotics, which includes chloramphenicol, is the main cause of the toxicity in patients that these antibiotics can cause. ■

Mitochondrial Genetic Codes Can Differ from the Standard Nuclear Code

The genetic code used in animal and fungal mitochondria is different from the standard code used in all prokaryotic, eukaryotic nuclear, and plant mitochondrial genes; remarkably, the code even differs among mitochondria from different species. Why and how these differences arose during evolution is a mystery. UGA, for

example, is normally a stop codon, but is read as tryptophan by human and fungal mitochondrial translation systems; AGA and AGG, the standard nuclear codons for arginine, also code for arginine in fungal and plant mtDNA, but they are stop codons in mammalian mtDNA and serine codons in *Drosophila* mtDNA.

Mutations in Mitochondrial DNA Cause Several Genetic Diseases in Humans

The severity of disease caused by a mutation in mtDNA depends on the nature of the mutation and on the proportion of mutant and wild-type mtDNAs present in a particular cell type. Generally when mutations in mtDNA are found, cells contain mixtures of wild-type and mutant mtDNAs — a condition known as *heteroplasmy*. Each time a mammalian somatic or germ-line cell divides, the mutant and wild-type mtDNAs segregate randomly into the daughter cells. Thus the mtDNA genotype, which fluctuates from one generation and from one cell division to the next, can drift toward predominantly wild-type or predominantly mutant mtDNAs. Since all enzymes required for the replication and growth of mammalian mitochondria, such as the mitochondrial DNA and RNA polymerases, are encoded in the nucleus and imported from the cytosol, a mutant mtDNA should not be at a replication disadvantage; mutants that have large deletions of mtDNA might even be at a selective advantage because they can replicate faster.

Recent research suggests that the accumulation of mutations in mtDNA is an important component of aging in mammals. Mutations in mtDNA have been observed to accumulate over time, probably because mammalian mtDNA is not repaired in response to DNA damage.

With few exceptions, all human cells have mitochondria, yet mutations in mtDNA affect only some tissues. Those most commonly affected are tissues that have a high requirement for the ATP produced by oxidative phosphorylation and tissues that require most or all of the mtDNA in the cell to synthesize sufficient amounts of functional mitochondrial proteins. For instance, *Leber's hereditary optic neuropathy* (degeneration of the optic nerve) is caused by a missense mutation in the mtDNA gene encoding subunit 4 of the NADH-CoQ reductase (complex I), a protein required for ATP production by mitochondria (see below). Several large deletions in mtDNA cause another set of diseases, including *chronic progressive external ophthalmoplegia*, characterized by eye defects, and *Kearns–Sayre syndrome*, characterized by eye defects, an abnormal heartbeat, and central nervous system degeneration. A third condition, causing "ragged-red" muscle fibers (with improperly assembled mitochondria) and associated uncontrolled jerky movements, is due to a single mutation in the TΨCG loop of the mitochondrial lysine tRNA. As a result of this mutation, the translation of several mitochondrial proteins is apparently inhibited. ■

KEY CONCEPTS OF SECTION 12.3

The Structure of Mitochondria

- In eukaryotic cells, mitochondria use aerobic oxidation to generate ATP. These multifunctional organelles are also responsible for many other key activities (see Table 12-1), including biosynthesis and metabolism of a wide variety of small molecules and regulated cell death.
- The mitochondrion has two distinct membranes (outer and inner) and two distinct subcompartments (the intermembrane space between the two membranes, and the matrix surrounded by the inner membrane) (see Figure 12-6). Aerobic oxidation occurs in the mitochondrial matrix and on the inner mitochondrial membrane.
- The inner mitochondrial membrane is a single continuous membrane with three compositionally, structurally, and functionally distinct domains: boundary membrane, cristae, and crista junctions.
- There are more than one thousand proteins associated with mammalian mitochondria, most of which are encoded by nuclear genes.
- Mitochondria and chloroplasts contain DNA that encodes a relatively small number of proteins involved in mitochondrial or photosynthetic electron transport and ATP synthesis (13 in mammalian mitochondria), ribosomal RNAs and sometimes tRNAs necessary to translate the organelle mRNAs.
- Mitochondrial DNA (mtDNA) and chloroplast DNA most likely evolved from bacteria that formed a symbiotic relationship with ancestral cells (see Figure 12-7).
- Most of the genes originally in the bacterial precursors to mitochondria and chloroplasts were either lost because their functions were redundant with nuclear genes or moved to the nuclear genome over evolutionary time, leaving different gene sets in the organelle DNAs of different organisms.
- Animal mtDNAs are circular molecules, reflecting their probable bacterial origin. Plant mtDNAs and chloroplast DNAs are generally longer than mtDNAs from other eukaryotes, largely because they contain more noncoding regions and repetitive sequences.
- Mitochondrial DNA (mtDNA) is located in the mitochondrial matrix and is often seen in regions devoid of densely stacked cristae.
- Because most mitochondrial DNA (mtDNA) is inherited from egg cells rather than sperm, mutations in mtDNA exhibit a maternal cytoplasmic pattern of inheritance. Similarly, chloroplast DNA is exclusively inherited from the maternal parent.

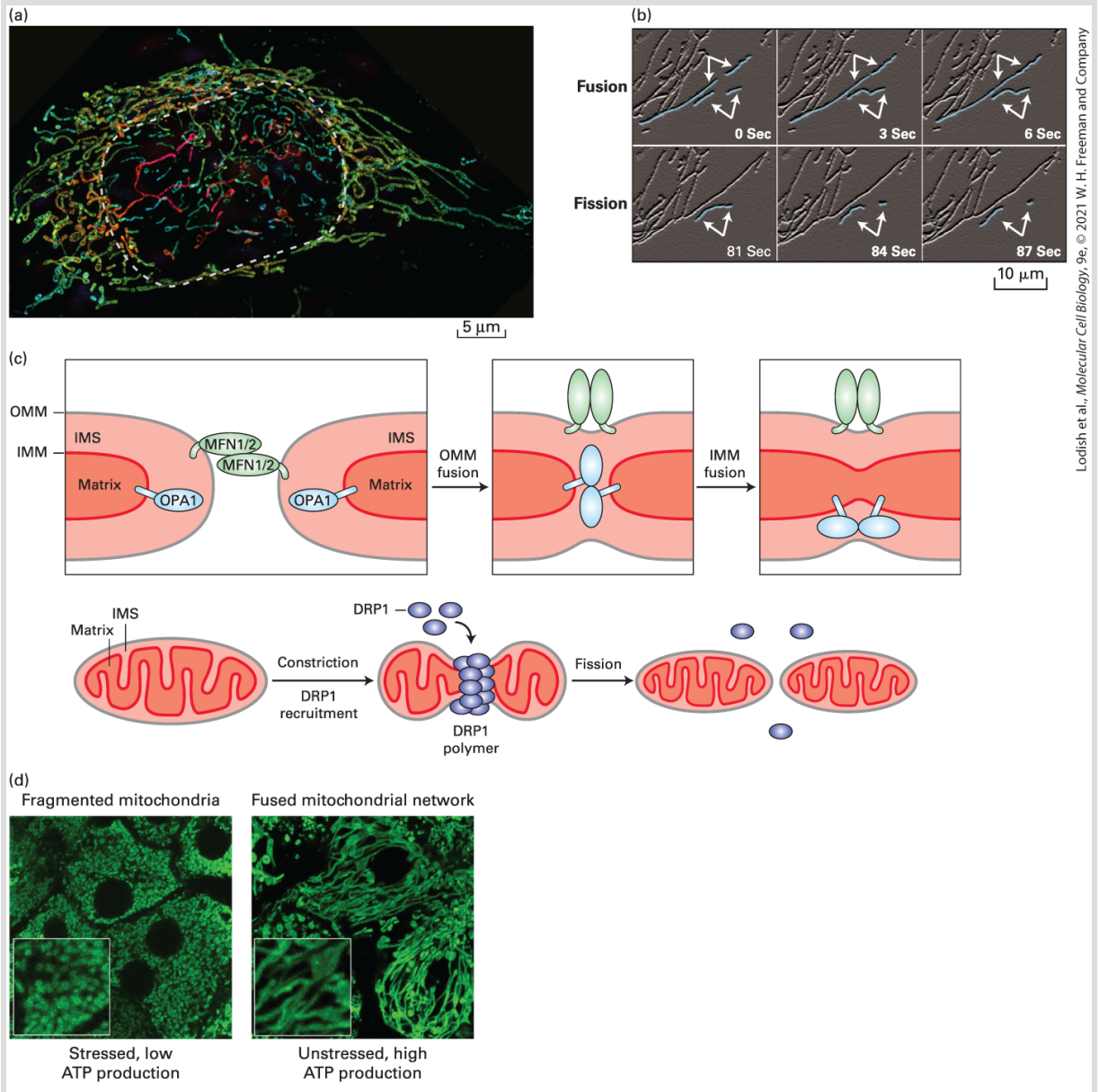
- Mitochondrial ribosomes resemble bacterial ribosomes in their structure and in their sensitivity to drugs such as chloramphenicol (sensitive) and cycloheximide (resistant).
- The genetic code of animal and fungal mtDNA differs slightly from that of bacterial and nuclear genomes and varies among different animals and fungi. In contrast, plant mtDNAs appear to conform to the standard genetic code.
- Several human neuromuscular disorders result from mutations in mtDNA. Patients generally have a mixture of wild-type and mutant mtDNA in their cells (heteroplasmy): the higher the fraction of mutant mtDNA, the more severe the mutant phenotype.

12.4 The Dynamics of Mitochondria and Mitochondrial-ER Membrane Contact Sites

Mitochondria are dynamic organelles that move about through the cytoplasm, merge with one another (fuse) to form large networks and break away from these networks (fission). Mitochondrial functions can be influenced by these dynamic behaviors and by the abilities of their outer membranes to attach to and interact with the plasma membrane and membranes of other organelles. Here we describe several features of mitochondrial dynamics and mitochondrial-endoplasmic reticulum (ER) membrane contact sites.

Mitochondria Are Dynamic Organelles

Many methods, including analysis of fluorescently labeled mitochondria in live cells, have been used to show that mitochondria are highly dynamic. They can move about the cell within the cytoplasm and undergo frequent **fusions** (merging) and **fissions** (breaking apart) that generate tubular, sometimes branched networks (Figures 12-9a and b) which may account for some of the wide variety of mitochondrial morphologies seen in different types of cells. Damaged or otherwise malfunctioning mitochondria can be recognized as such and consequently degraded as part of a quality control mechanism.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 12-9 Mitochondria undergo rapid fusion and fission. (a) A human HeLa cell labeled with a mitochondrion-specific fluorescent dye (MitoTracker Green) was imaged using three-dimensional structured illumination fluorescence microscopy (a 6.1-µm-thick section through the cell is shown). The network of fused and branched mitochondria is seen in the cytoplasm, with only a few mitochondria observed above or below the nucleus (outlined by thin gray dashed line). The identity of the striations seen within the mitochondria in this image is not known. The mitochondria are shown in artificial colors to indicate their positions relative to the surface to which the cell is attached (blue is closest to and red farthest from the surface). (b) Mitochondria labeled with a fluorescent protein in a live normal mouse embryonic fibroblast were observed using time-lapse

fluorescence microscopy. Several mitochondria undergoing fusion (*top*) or fission (*bottom*) are artificially highlighted in blue and with arrows. (c) Mitochondrial fusion (*top*) and fission (*bottom*) are mediated by a set of GTPase enzymes (MFN1, MFN2, OPA1, and DRP1). The integral membrane proteins MFN1 and MFN2 $(\mathrm{MFN1/2})$ mediate outer mitochondrial membrane (OMM) fusion, which is followed by fusion of the inner mitochondrial membranes (IMM) mediated by the integral membrane protein OPA1. The matrix and inner membrane space (IMS) remain distinct. The soluble cytosolic GTPase DRP1 is recruited to a constricted site on the surface of a mitochondrion, where DRP1 polymers sever the membrane, resulting in fission. A variety of post-translational modifications of DRP1 regulate fission. (d) *Left*: Rat liver cells (hepatocytes) one day after being removed from the liver and placed in cell culture, are stressed and depolarized (lack some of the morphological and biochemical properties of epithelial cells; see Chapter 20), have low levels of oxidative phosphorylation and ATP production, and have fragmented mitochondria (visualized by staining with MitoTracker Green). *Right*: After growth in culture for 6 days, the hepatocytes become polarized, their mitochondria fuse, forming an extensive network, and the cells exhibit high levels of oxidative phosphorylation and ATP production. Insets show higher magnification views of the mitochondria.

[(a) Reprinted with permission from Nature Publishing Group, from Shao et al., 2011, "Super-Resolution 3D Microscopy of Live Whole Cells Using Structured Illumination," *Nature Methods* **8**:121044–121046; permission conveyed through Copyright Clearance Center, Inc. (b) Republished with permission from Elsevier, modified from Chan D. C., 2006, "Mitochondria: Dynamic Organelles in Disease, Aging, and Development," *Cell* **125**(7):1241–1252; permission conveyed through Copyright Clearance Center, Inc. (c) Information from P. Mishra and D. C. Chan, 2014, *Nat. Rev. Mol. Cell Biol.* **15**:634–646. (d) From D. Fu et al., 2013, "Coordinated Elevation of Mitochondrial Oxidative Phosphorylation and Autophagy Help Drive Hepatocyte Polarization," *Proc. Nat'l Acad. Sci. USA* **110**(18):7288–7293, Fig. 3 Day 1 and Day 6; courtesy Jennifer Lippincott-Schwartz.]

Intracellular Mitochondrial Movements

Intracellular movement helps to establish the proper geographical distribution of mitochondria within cells. As we shall see below,

cells require that activities of mitochondria and other structures, including other organelles, must be coordinated. For example, in neurons (see Chapter 23) that have extensions called *axons* that can be centimeters to even a meter long, mitochondria must be transported from the region near the nucleus (cell body) to the ends of the axon where contact is made with other neurons to provide energy (ATP) and to regulate calcium levels (see below and Chapter 15). Proteins on the cytoplasmic surface of the outer mitochondrial membrane mediate mitochondrial interactions with motor proteins such as kinesin (Chapter 18) that drive intracytoplasmic movement via microtubules (Figure 18-28), including long distance movement through the axons of neurons. One such protein is called *Miro*, a GTPase that has calcium binding EF hands (Figure 3-7) that appear to confer calcium sensitivity to mitochondrial movement. Miro indirectly links mitochondria to kinesin, and it also is associated with MICOS in the inner membrane and thus influences cristae structure and function. Miro has also been implicated in mitochondrial-endoplasmic reticulum interactions and in Parkinson's disease, both described below. In addition to intracellular movement, mtDNA, and indeed intact mitochondria, have reported to be transferred from one cell to another via membrane tubules called *tunneling nanotubes*, which are described in Chapter 20.

Mitochondrial Fusion and Fission

When individual mitochondria fuse, each of the two membranes fuses (inner with inner, and outer with outer) and each of their distinct compartments intermix (matrix with matrix, intermembrane space with intermembrane space). A set of four evolutionarily conserved GTP-hydrolyzing (GTPase) enzymes — MFN1 (mitofusin 1), MFN2 (mitofusin 2), OPA1, and DRP1 — plays critical roles in mediating these membrane fusions and fissions (Figure 12-9c). These enzymes are members of the dynamin family of GTPases. Dynamin, the first-identified member of this family, mediates a comparable membrane fission reaction required during the pinching off of endocytic vesicles from the plasma membrane (see Chapter 14). Mutations in several of the genes encoding these GTPases can disrupt mitochondrial functions, such as maintenance of proper inner membrane structure and electric potential, the function of which is described in Section 12.6. These mutations result in truncated or enlarged mitochondria and can cause human disease. The inherited, autosomal dominant neuromuscular disease Charcot-Marie-Tooth subtype 2A is caused by loss-of-function mutations in MFN2 that lead to defects in peripheral nerve function and progressive muscle weakness, mainly in the feet and hands. Mutations in OPA1 are associated with autosomal dominant optic atrophy that influences nerves in the eye's retina.

What is the value of mitochondrial fission and fusion? It has been suggested that fusion helps to maintain a relatively homogeneous population of mitochondria within a cell. Should some subset of individual mitochondria suffer deleterious modification or loss of

important components, fusion with other mitochondria would permit restoration by sharing of those components. Studies of cells and organisms with mutations in genes encoding the fusion machinery suggest that fusion also plays a role in the proper localization of mitochondria within cells, maintaining the proper morphology and cristal organization, distribution of mitochondrial DNA, and maintenance of fully functional electron transport.

A number of functions are served by mitochondrial fission. For example, mitochondrial fission is particularly active when cells divide (particularly during the G_2 and M phases of the cell cycle; see Chapter 19). As a consequence, the multiple discrete mitochondria generated by fission are readily distributed into both daughter cells. In addition, fission provides a powerful mechanism of quality control by culling defective segments of mitochondria from the interconnected, healthy mitochondrial network. Should a portion of a large mitochondrial network become damaged or dysfunctional — for example, by the generation of high levels of reactive oxygen species (discussed later) or by mutations in the mitochondrial DNA — fission can separate the compromised segments from healthy segments.

Mitophagy Provides Mitochondrial Quality Control

Cells have the capacity to recognize damaged or dysfunctional segments of the mitochondrial network and, after they detach from

the network, can surround them with a membrane and then deliver them to lysosomes for degradation. This destruction of mitochondria, which is called **mitophagy** ("eating mitochondria"), is a subset of the general process called *autophagy* ("self eating") by which cells engulf in membranes and degrade organelles and portions of the cytosol (see Chapter 14). For example, when the environment of a cell has inadequate oxygen (hypoxia), enzymatic modification of proteins on the outer mitochondrial membrane induces mitophagy.

Hereditary early onset Parkinson's disease is caused by mutations in two gene encoding proteins that can mediate mitophagy. The proteins are PINK1 (a kinase) and Parkin (an E3 ubiquitin ligase that ubiquitinylates a substrate protein, thus targeting it for destruction by proteasomes; see Chapter 3). In healthy mitochondria, PINK1 resides in the mitochondrial matrix. In damaged or dysfunctional mitochondria, PINK1 resides on the outer mitochondrial membrane, where it recruits cytosolic Parkin. The Parkin is activated and polyubiquitinylates outer membrane proteins, thus targeting them for degradation and inducing mitophagy. There also are PINK1/Parkin-independent mechanisms that mediate mitophagy. Some studies suggest that mitochondrial stress (due to aging, pathogens, etc.) can induce inflammation and contribute to pathology and that mitophagy may mitigate mitochondrial stress-induced inflammation by removing defective segments of mitochondria. ■

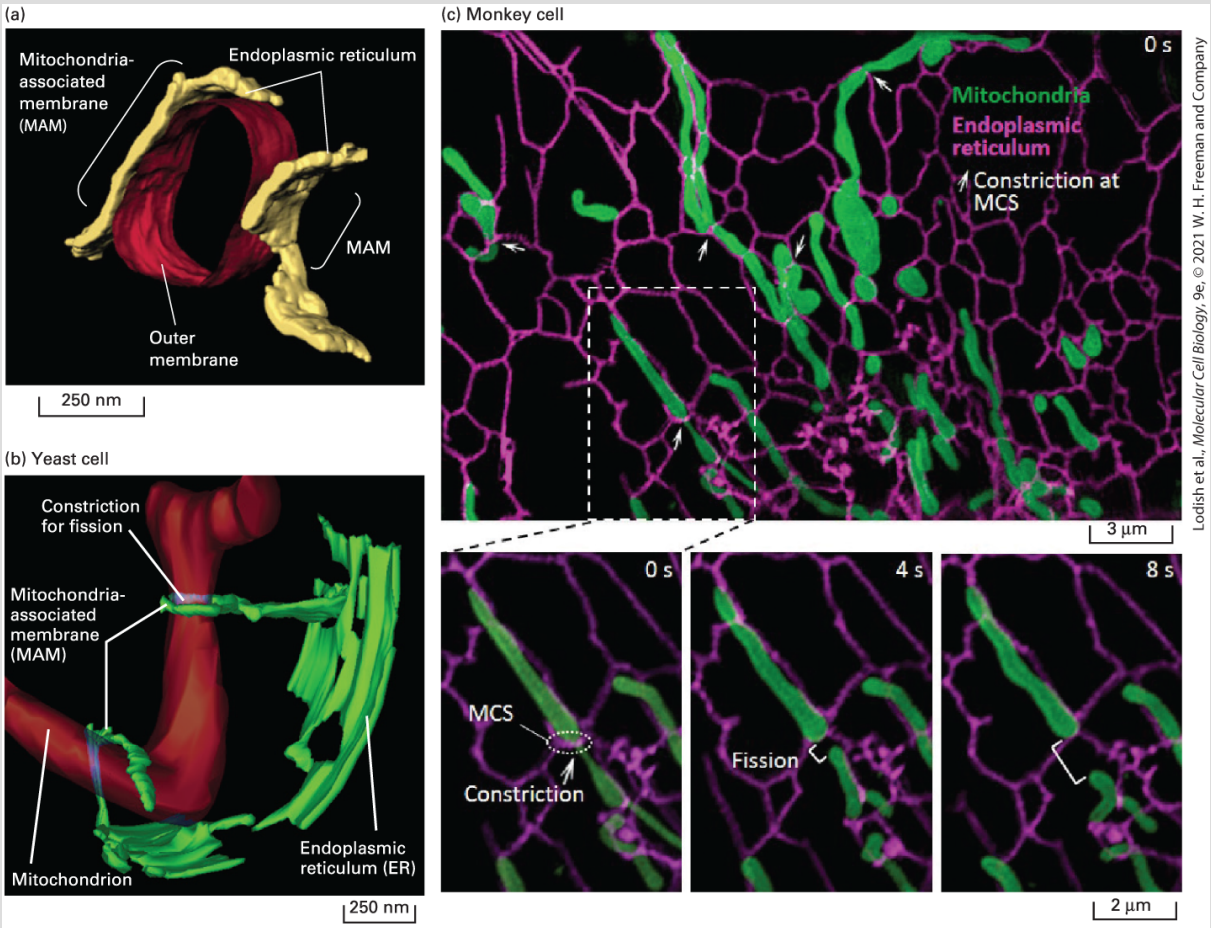
Mitochondrial dynamism has additional features that are worth mentioning. Mitochondrial structure and function can change in response to the metabolic state of the cells. For example, the isolation of rat liver cells (hepatocytes) from the body and their transfer into cell culture stresses the cells and results in their depolarization (loss of some of their epithelial cell properties; see Chapter 20). These cells exhibit fragmentation of their mitochondria (Figure 12-9d, *left*) accompanied by low generation of ATP via oxidative phosphorylation. As the cells adjust to growth in cell culture, they become polarized (have a morphology and biochemistry more like that of hepatocytes in a liver), increase mitochondrial generation of ATP via oxidative phosphorylation, and exhibit an extensive network of fused mitochondria (Figure 12-9d, *right*).

Mitochondrial Function and Dynamics Can Depend on Direct Contacts with Other Organelles

The activities of mitochondria and other organelles must be coordinated with one another for efficient and effective cell function. Such coordination requires inter-organelle communication, including transfer of ions, small molecules, and macromolecules between organelles. Four mechanisms for such transfers include: release into the cytosol and organelle uptake of

ions and small molecules by transporters, pumps, and channels (Chapter 11); transfer of proteins across organelle membranes via pore-like integral membrane proteins (Chapter 13); transport via small membrane vesicles (Chapter 14); and direct, protein-mediated organelle-organelle binding via **membrane contact sites (MCSs)**. Mitochondria establish MCSs with many other organelles, including the endoplasmic reticulum (ER), lysosomes (vacuoles in yeast), peroxisomes, the plasma membrane, the Golgi, endosomes, and even lipid droplets (Chapter 10). Here we will focus on mitochondrial-ER MCSs. The portions of the ER membranes that form special contact regions with the mitochondria are called **mitochondria-associated membranes**, or **MAMs**.

Mitochondrial-ER MCSs can be visualized using electron microscopy and fluorescence microscopy (Figure 12-10). The MAMs' lipid and protein composition differs somewhat from that of the rest of the ER. At the MCSs, tethering proteins hold the MAM and the outer mitochondrial membrane about 10–30 nm apart. Those tethering proteins in yeast have been identified as a protein complex called *ERMES*. Mammals do not have the ERMES complex; they use what appears to be a variety of other proteins to tether mitochondria to the ER (see Figure 15-29).



EXPERIMENTAL FIGURE 12-10 At mitochondrial-ER membrane contact sites (MCSs) specialized regions of the endoplasmic reticulum called *mitochondria-associated membranes (MAMs)* directly contact mitochondria and influence mitochondrial shape, function, and sites of fission. (a) Three-dimensional model of a segment of a mitochondrion (red, only outer membrane shown) and the adjacent ER (yellow) determined from a line of cultured avian lymphoma cells using EM tomography (assembly of a three-dimensional image from consecutive individual sections). At the MCS, the MAM is the portion of the ER membrane that is closely apposed (10–30 nm distance) to the outer mitochondrial membrane. The regions of the MCSs are indicated by brackets. (b) A three-dimensional model of a mitochondrion (red) and adjacent MAMs (green) from a yeast cell using EM tomography. The two MAM domains are derived from ER tubules that in some cases can wrap around the mitochondrion, in the top case forming a clamp-like structure at the MCS that appears to constrict the mitochondrion in preparation for fission. (c) Many mitochondrial fission events occur at the mitochondrial-ER MCSs. These live-cell fluorescence microscopic images of a Cos-7 monkey cell show such a fission event observed

over the 8 seconds of imaging. We see mitochondria (green) and tubular ER (magenta) in a single field of view at 4-second intervals, visualized using a superresolution (97 nm) and high-speed method called *grazing incidence-structured illumination microscopy*. The arrows in the upper panel point to sites of constriction in the mitochondria at MCSs that will develop into sites of DRP1-mediated fission. The upper panel shows the cell at 0 seconds and the area outlined by a white dashed box is shown in the lower panels at three times (0, 4, and 8 seconds), which illustrate the rapid fission (white brackets indicate separation of the released mitochondrial ends) initiated at the MCS (white dotted oval in the left panel). To visualize the mitochondria and ER, the Cos-7 cells were transfected with cDNA vectors encoding two fluorescent proteins that specifically accumulate in either the outer mitochondrial membranes (mEmerald-Tomm20) or the lumen of the ER (mCherry-KDEL).

[Part (a) G. Csordás et al., 2006, *J. Cell Biol.* **174**:915–921; https://doi.org/10.1083/jcb.200604016. Part (b) Republished with permission from American Association for the Advancement of Science, from J. R. Friedman et al., 2011, "ER Tubules Mark Sites of Mitochondrial Division," *Science* **334**(6054):358–362; permission conveyed through the Copyright Clearance Center, Inc. Part (c) Reprinted with permission from Elsevier, from Y. Guo et al., 2018, "Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution on Millisecond Timescales," *Cell* **175**(5):1430–1442; permission conveyed through Copyright Clearance Center, Inc.]

Mitochondrial-ER MCSs significantly contribute to many cellular processes (see Table 12-1). For example, these MCSs play important roles in mitochondrial dynamics. A majority of mitochondrial fission events occur at the mitochondrial-ER MCSs, although most of the MCSs are not associated with sites of fission. MCSs apparently initiate mitochondrial constriction by inducing an increased local density of filamentous actin polymers that, with the aid of myosin motor proteins, contract (Chapter 17). The contractions squeeze down on and thus constrict the mitochondrion. The change of shape of the mitochondrial outer membrane at the constriction recruits

DPR1, which completes membrane fission (see Figure 12-9c). In yeast, MAM tubules have been seen to loop completely around mitochondria (Figure 12-10b). In mammalian cells, the MAMs contact the mitochondria at fission sites, but they usually do not appear to loop fully around the mitochondria (Figure 12-10c).

MCSs also serve as interfaces for the transfer of ions (e.g., Ca^{2+}) and small molecules (e.g., glycerophospholipids) between the ER and mitochondria. These MCSs play an integral role in intracellular calcium and energy metabolism. Variations in the concentrations of calcium ions in intracellular compartments — cytosolic calcium ($[Ca^{2+}]_c$), mitochondrial calcium ($[Ca^{2+}]_m$), and calcium in the ER ($[Ca^{2+}]_{er}$) — are employed to control a wide variety of activities within cells, a process called *calcium signaling* (see Chapter 15). Calcium is also important for extracellular processes, such as the activity of some blood-clotting proteins. Intramitochondrial calcium ions play an important role in controlling mitochondrial function, and MCSs mediate this control by delivering calcium from the ER to mitochondria. For example, an increase in $[Ca^{2+}]_m$ in the matrix can increase mitochondrial production of ATP. Increased $[Ca^{2+}]_m$ directly increases the activities of three mitochondrial enzymes that produce NADH from NAD^+: pyruvate dehydrogenase (see Figure 12-5), and α-ketoglutarate and isocitrate dehydrogenases (see Figure 12-13 in the following section). As we shall see later in this chapter, NADH provides high-energy electrons for ATP synthesis. Thus continuous low-level release of Ca^{2+} from ER MAMs into

mitochondria is necessary for ATP synthesis when cells are in a basal, or resting, state. Increased delivery of Ca^{2+} via MCSs can occur when cells require more ATP — for example, when muscle cells are stimulated to contract. Strikingly, calcium signaling is used both to induce muscle contraction (see Chapter 17) and coordinately to increase mitochondrial ATP synthesis to provide the energy to fuel that contraction. When $[Ca^{2+}]_m$ is elevated, mitophagy can be induced. Indeed, mitochondrial calcium overload can activate regulated cell death pathways (Chapter 22). Thus the control of $[Ca^{2+}]_m$ can literally control the life and death of cells.

KEY CONCEPTS OF SECTION 12.4

The Dynamics and Functions of Mitochondria

- Mitochondria are dynamic organelles, undergoing intracellular transport, as well as fusion and fission reactions that are regulated by the state of the cell. In many cells, the fused mitochondria form a large, interconnected branched tubular network. A family of GTPases mediate mitochondrial membrane fusion and fission (see Figure 12-9). Mutations in the genes encoding some of these GTPases cause human diseases.
- Mitochondrial fission and fusion are thought to play roles in maintaining a relatively homogeneous population of mitochondria, distributing mitochondria among the daughter cells during cell division, and establishing a system of quality control to permit culling of defective mitochondria from healthy mitochondria.
- Defective mitochondria or segments of mitochondria are destroyed by a processed called *mitophagy*. Two proteins that can mediate mitophagy, PINK1 and Parkin, are encoded by genes that, when mutated, can be responsible for hereditary early onset Parkinson's disease.
- Membrane contact sites (MCSs) are regions of organelles that directly bind to one another via tethering proteins. Mitochondria form MCSs with many different organelles.
- Mitochondrial-endoplasmic reticulum (ER) MCSs significantly influences many cellular functions, including mitochondrial shape and dynamics (see Table 12-1). The portions

of the ER membranes that are in contact with mitochondria at MCSs are called *mitochondria-associated membranes (MAMs)* (see Figure 12-10).

- A majority of mitochondrial fission events occur at the mitochondrial-ER MCSs, although most of the MCS are not associated with sites of fission.
- MCSs also serve as interfaces for the transfer of ions (e.g., Ca^{2+}) and small molecules (e.g., glycerophospholipids) between the ER and mitochondria. These MCSs play an integral role in intracellular calcium and energy metabolism. Calcium influx into mitochondria from MAMs can stimulate ATP synthesis and, in the context of mitochondrial calcium overload, initiates a program of regulated cell death.

12.5 The Citric Acid Cycle and Fatty Acid Oxidation

We now continue our detailed discussion of glucose oxidation and ATP generation, exploring what happens to the pyruvate generated during glycolysis (stage I, see Figures 12-1 and 12-3) after it is transported into the mitochondrial matrix. The last three of the four stages of glucose oxidation (Figure 12-11) are:

- **Stage II:** Stage II can be subdivided into two distinct parts: (1) the conversion of pyruvate to acetyl CoA and CO_2, followed by (2) oxidation of acetyl CoA to CO_2 in the citric acid cycle. These oxidations are coupled to reduction of NAD^+ to NADH and of FAD to $FADH_2$. These two carriers can be considered the sources of high-energy electrons. (Fatty acid oxidation follows a similar route, with conversion of fatty acyl CoA to acetyl CoA.) Most of the reactions occur in or on the inner membrane facing the matrix.
- **Stage III:** Electron transfer from NADH and $FADH_2$ to O_2 via an electron-transport chain within the inner membrane converts the energy carried in those electrons into an electrochemical gradient across that membrane, called the *proton-motive force*.
- **Stage IV:** The energy of the proton-motive force is harnessed for ATP synthesis in the inner mitochondrial membrane. Stages III and IV together are called *oxidative phosphorylation*.

In the First Part of Stage II, Pyruvate Is Converted to Acetyl CoA and High-Energy Electrons

Within the mitochondrial matrix, pyruvate reacts with coenzyme A, forming CO_2, acetyl CoA, and NADH (see Figure 12-11, stage II). This reaction, catalyzed by *pyruvate dehydrogenase*, is highly exergonic ($\Delta G^{\circ\prime} = -8.0\,\text{kcal/mol}$) and essentially irreversible. Influx of calcium into the mitochondria at mitochondrial-ER MCSs (from the ER MAM) increases the activity of pyruvate dehydrogenase, driving the formation of acetyl CoA.

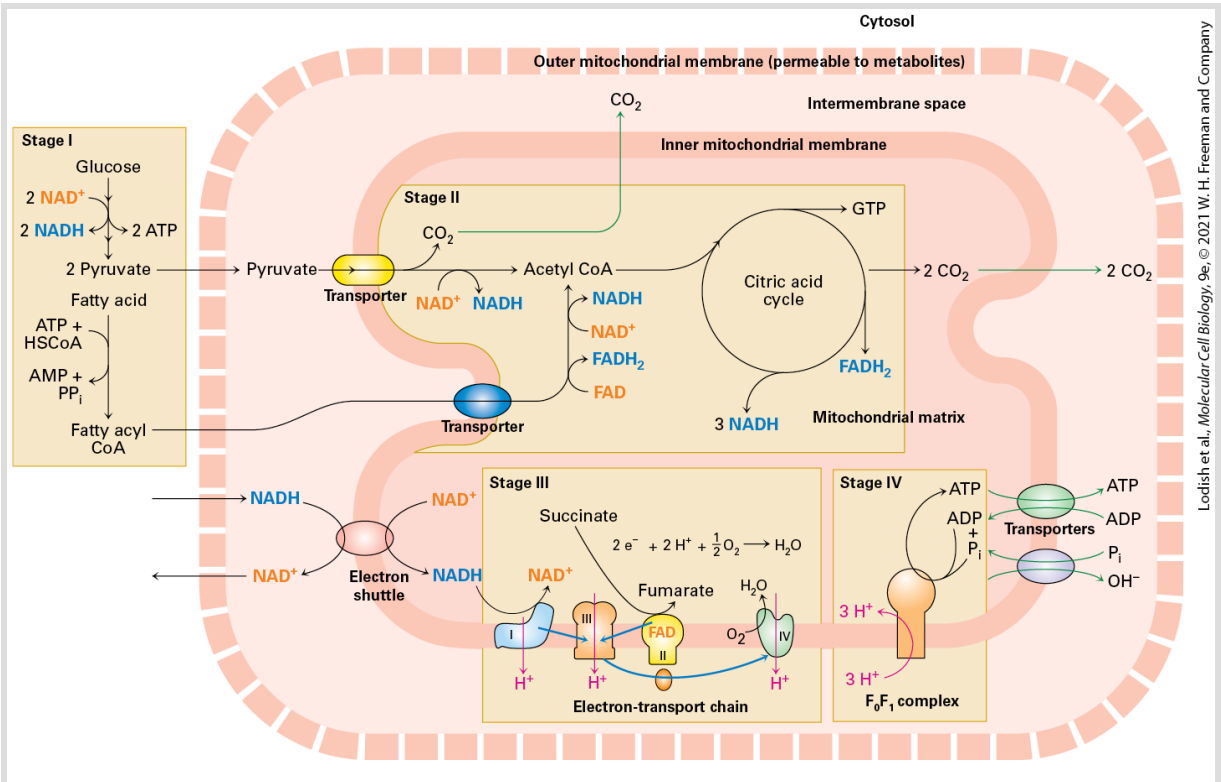


FIGURE 12-11 Summary of aerobic oxidation of glucose and fatty acids. Stage I: In the cytosol, glucose is converted to pyruvate (glycolysis) and fatty acid to fatty acyl CoA.

Pyruvate and fatty acyl CoA then move into the mitochondrion. Mitochondrial porins make the outer membrane permeable to these metabolites, but specific pyruvate and fatty acid transport proteins (yellow and blue colored ovals, respectively) in the inner membrane are required to import these molecules into the matrix. Fatty acyl groups are transferred from fatty acyl CoA to an intermediate carrier, transported across the inner membrane, and then reattached to CoA on the matrix side. **Stage II:** In the mitochondrial matrix, pyruvate and fatty acyl CoA are converted to acetyl CoA and then oxidized, releasing CO_2. Pyruvate is converted to acetyl CoA with the formation of NADH and CO_2; two carbons from fatty acyl CoA are converted to acetyl CoA with the formation of $FADH_2$ and NADH. Oxidation of acetyl CoA in the citric acid cycle generates NADH and $FADH_2$, GTP, and CO_2. **Stage III:** Electron transport reduces O_2 to H_2O and generates a proton-motive force. Electrons (blue) from reduced coenzymes NADH and $FADH_2$, are transferred via electron-transport complexes (colored shapes) to O_2 concomitant with transport of H^+ ions (red) from the matrix to the intermembrane space, generating the proton-motive force. Electrons from NADH flow from complex I (blue) to complex III (orange), bypassing complex II (yellow). Electrons from succinate flow through $FAD/FADH_2$ within complex II to complex III (orange), bypassing complex I. **Stage IV:** ATP synthase, also called the F_0F_1 *complex* (orange), harnesses the proton-motive force to synthesize ATP in the matrix. Antiporter proteins (purple and green ovals) transport ADP and P_i into the matrix and export hydroxyl groups and ATP. NADH generated in the cytosol is not transported directly to the matrix because the inner membrane is impermeable to NAD^+ and NADH; instead, a shuttle system (red) transports electrons from cytosolic NADH to NAD^+ in the matrix. O_2 diffuses into the matrix, and CO_2 diffuses out.

Acetyl CoA is a molecule that consists of a two-carbon acetyl group covalently linked to a longer molecule known as *coenzyme A (CoA)* (Figure 12-12). It plays a central role in the oxidation of pyruvate, fatty acids, and amino acids. In addition, it is an intermediate in numerous biosynthetic reactions, including the transfer of an acetyl group to histones (Figure 8-33) and many other mammalian proteins, and the synthesis of lipids such as cholesterol (Figure 10-26). In respiring mitochondria, however, the two-carbon acetyl

group of acetyl CoA is almost always oxidized to CO_2 via the citric acid cycle. Note that the two carbons in the acetyl group come from pyruvate; the third carbon of pyruvate is released as carbon dioxide.

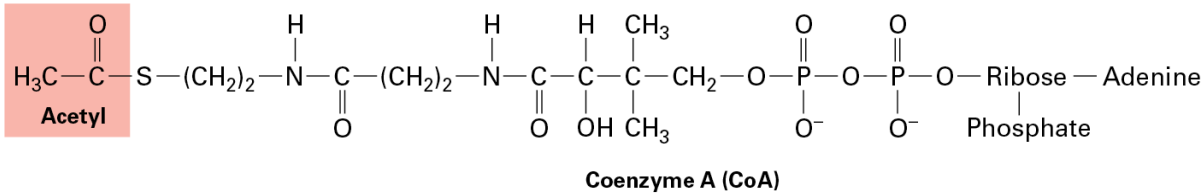


FIGURE 12-12 The structure of acetyl CoA. This compound, consisting of an acetyl group covalently linked to a coenzyme A (CoA) molecule, is an important intermediate in the aerobic oxidation of pyruvate, fatty acids, and many amino acids. It also contributes acetyl groups to many biosynthetic pathways.

In the Second Part of Stage II, the Citric Acid Cycle Oxidizes the Acetyl Group in Acetyl CoA to CO_2 and Generates High-Energy Electrons

Nine sequential reactions operate in a cycle to oxidize the acetyl group of acetyl CoA to CO_2 (see Figures 12-11 and 12-13, stage II). This cycle is referred to by several names: the **citric acid cycle**, the *tricarboxylic acid (TCA) cycle*, and the *Krebs cycle*. The net result is that for each acetyl group entering the cycle as acetyl CoA, two molecules of CO_2, three of NADH, and one each of $FADH_2$ and GTP are produced. NADH and $FADH_2$ are high-energy electron carriers

that will play a major role in stage III of mitochondrial oxidation: electron transport.

As shown in Figure 12-13, the cycle begins with condensation of the two-carbon acetyl group from acetyl CoA and the four-carbon molecule *oxaloacetate* (step **1**) to yield the six-carbon *citric acid* (*citrate* when ionized, hence the name *citric acid cycle*). Reactions step **4** and step **5** each release a CO_2 molecule and reduce NAD^+ to NADH. The source of the oxygen for generating the CO_2 molecules in these reactions is water (H_2O), not molecular oxygen (O_2), and the enzymatic activities of the enzymes catalyzing reactions step **4** and step **5** are increased by the influx of calcium into the mitochondrion from the ER's MAM. Reduction of NAD^+ to NADH occurs during reaction step **9**; thus three NADHs are generated per turn of the cycle. In reaction step **7**, two electrons and two protons are transferred to FAD, yielding the reduced form of this coenzyme, $FADH_2$. Reaction step **7** is distinctive not only because it is an intrinsic part of the citric acid cycle (stage II) but also because it is catalyzed by a membrane-attached enzyme that, as we shall see, also plays an important role in stage III. In reaction step **6**, hydrolysis of the high-energy thioester bond in succinyl CoA is coupled to synthesis of one GTP by substrate-level phosphorylation. Because GTP and ATP are interconvertible,

$$\mathrm{GTP} + \mathrm{ADP} \rightleftharpoons \mathrm{GDP} + \mathrm{ATP}$$

this can be considered an ATP-generating step. Reaction step **9** regenerates oxaloacetate, so the cycle can begin again. Note that molecular O_2 does not participate in the citric acid cycle.

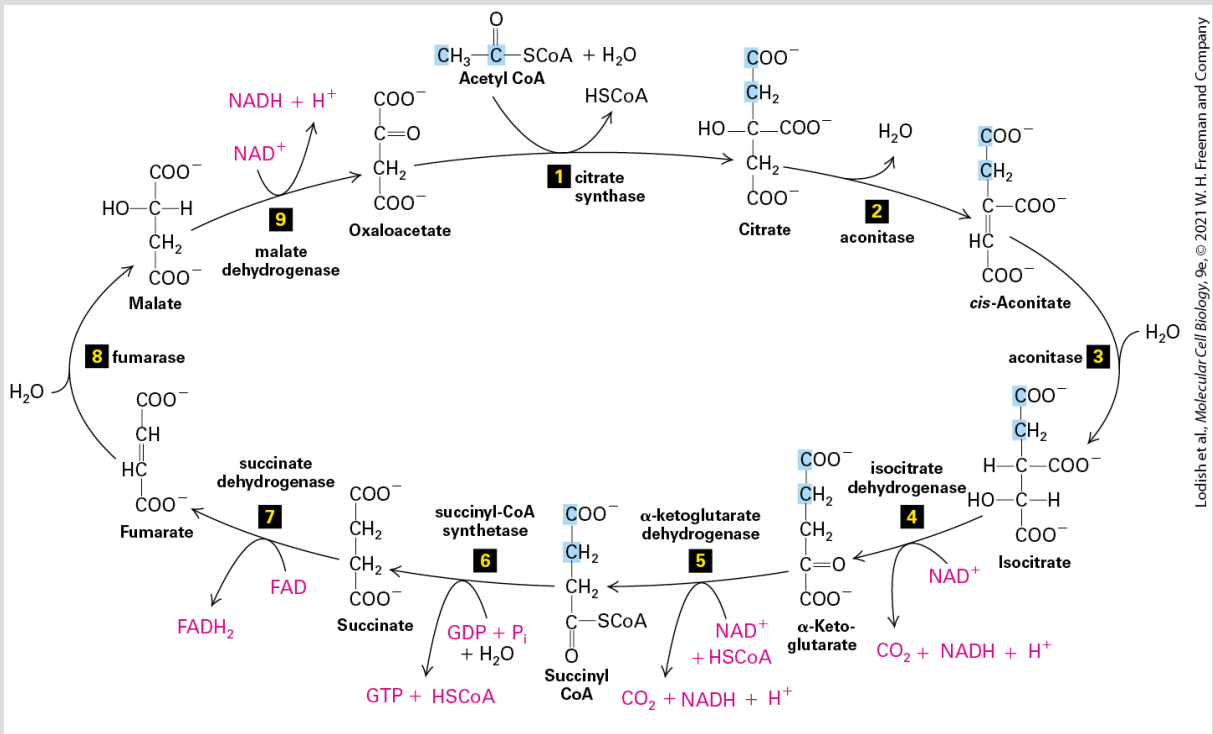


FIGURE 12-13 The citric acid cycle. Acetyl CoA is metabolized to CO_2 and the high-energy electron carriers NADH and $FADH_2$. In reaction **1**, a two-carbon acetyl residue from acetyl CoA condenses with the four-carbon molecule oxaloacetate to form the six-carbon citrate. In the remaining reactions (**2–9**), each molecule of citrate is eventually converted back to oxaloacetate, losing two CO_2 molecules in the process. In each turn of the cycle, four pairs of electrons are removed from carbon atoms, forming three molecules of NADH from three of NAD^+, one molecule of $FADH_2$ from FAD, and one molecule of GTP, from GDP and P_i. The two carbon atoms that enter the cycle with acetyl CoA are highlighted in blue through succinyl CoA. In succinate and fumarate, which are symmetric molecules, they can no longer be specifically denoted. Isotope-labeling studies have shown that these carbon atoms are *not* lost in the turn of the cycle in which they enter; on average, one will be lost as CO_2 during the next turn of the cycle and the other in subsequent turns.

Most enzymes and small molecules involved in the citric acid cycle are soluble in the aqueous mitochondrial matrix. These include CoA, acetyl CoA, succinyl CoA, NAD^+, and NADH, as well as most of the citric acid cycle enzymes. *Succinate dehydrogenase* (reaction step **7**), however, is a component of an integral membrane protein in the inner membrane, with its active site facing the matrix. The non-membrane-bound enzymes of the citric acid cycle can be isolated as very large **multiprotein complexes**. It is believed that within such complexes, the reaction product of one enzyme passes directly to the next enzyme without diffusing through the solution (see Figure 3-31).

Because glycolysis of one glucose molecule generates 2 pyruvate molecules, and thus 2 acetyl CoA molecules, the reactions in the glycolytic pathway and citric acid cycle produce 6 CO_2 molecules, 10 NADH molecules, and 2 $FADH_2$ molecules per glucose molecule (Table 12-3). Although these reactions also generate four high-energy phosphoanhydride bonds in the form of 2 ATP and 2 GTP molecules, this represents only a small fraction of the available energy released in the complete aerobic oxidation of glucose. The remaining energy is stored as high-energy electrons in the reduced coenzymes NADH and $FADH_2$, which can be thought of as high-energy electron carriers. The goal of stages III and IV is to recover this energy in the form of ATP.

TABLE 12-3 • Electron-Carrying Prosthetic Groups in the Electron-

Transport Chain

Protein Component	Prosthetic Groups[i]
NADH-CoQ reductase (complex I)	*FMN* Fe-S
Succinate-CoQ reductase (complex II)	FAD Fe-S
$CoQH_2$–cytochrome *c* reductase (complex III)	Heme b_L Heme b_H Fe-S Heme c_1
Cytochrome *c*	Heme *c*
Cytochrome *c* oxidase (complex IV)	$Cu_a{}^{2+}$ Heme *a* $Cu_b{}^{2+}$ Heme a_3

[i]Not included is coenzyme Q, an electron carrier that is not permanently bound to a protein complex.

SOURCE: Data from J. W. De Pierre and L. Ernster, 1977, *Annu. Rev. Biochem.* **46**:201.

Transporters in the Inner Mitochondrial Membrane Help Maintain Appropriate Cytosolic and

Matrix Concentrations of NAD^+ and NADH

In the cytosol, NAD^+ is required for step **6** of glycolysis (see Figure 12-3); in the mitochondrial matrix, NAD^+ is required for the conversion of pyruvate to acetyl CoA and for three steps in the citric acid cycle (step **4**, step **5**, and step **9** in Figure 12-13). In each case, NADH is a product of the reaction. If glycolysis and oxidation of pyruvate are to continue, NAD^+ must be regenerated by oxidation of NADH to ensure that this substrate is available. (Similarly, the $FADH_2$ generated in stage II reactions must be reoxidized to FAD if FAD-dependent reactions are to continue.) As we will see in the next section (Section 12.6), the electron-transport chain *within* the inner mitochondrial membrane converts NADH to NAD^+ and $FADH_2$ to FAD as it reduces O_2 to water and converts the energy stored in the high-energy electrons in the reduced forms of these molecules into a proton-motive force (stage III). Even though O_2 is not involved in any reaction of the citric acid cycle, in the absence of O_2 this cycle soon stops operating because in such anaerobic conditions, the mitochondria cannot regenerate the required NAD^+ and FAD substrates. NAD^+ and FAD concentrations dwindle due to the inability of the electron-transport chain within the mitochondrion to oxidize NADH and $FADH_2$. These observations raise the question of how a supply of NAD^+ in the cytosol is regenerated.

If the NADH from the cytosol could move into the mitochondrial matrix and be oxidized by the electron-transport chain, and if the NAD^+ product could be transported back into the cytosol, regeneration of cytosolic NAD^+ would be simple when O_2 is available. However, the inner mitochondrial membrane is impermeable to NADH. To bypass this problem and permit the electrons from cytosolic NADH to be transferred *indirectly* to O_2 via the mitochondrial electron-transport chain, cells use several *electron shuttles* to transfer electrons from NADH in the cytoplasm to NAD^+ in the matrix. The operation of the most widespread shuttle — the *malate-aspartate shuttle* — is depicted in Figure 12-14.

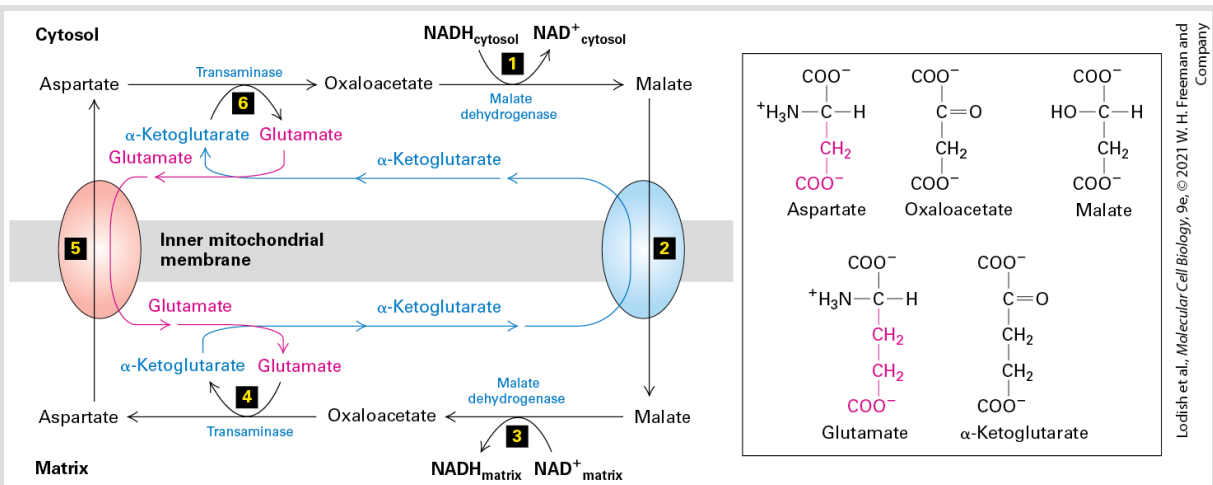


FIGURE 12-14 The malate-aspartate shuttle. This cyclical series of reactions transfers electrons from NADH in the cytosol (via the intermembrane space) across the inner mitochondrial membrane, which is impermeable to NADH itself, to NAD^+ in the matrix. The net result is the replacement of cytosolic NADH with NAD^+ and matrix NAD^+ with NADH. Step **1**: Cytosolic malate dehydrogenase transfers electrons from cytosolic NADH to oxaloacetate, forming malate and NAD^+. Step **2**: An antiporter (blue oval) in the inner mitochondrial membrane transports malate into the matrix in exchange for α-ketoglutarate. Step **3**: Mitochondrial malate dehydrogenase converts malate back to oxaloacetate,

reducing NAD^+ in the matrix to NADH in the process. Step 4: Oxaloacetate, which cannot directly cross the inner membrane, is converted to aspartate by addition of an amino group from glutamate. In this transaminase-catalyzed reaction in the matrix, glutamate is converted to α-ketoglutarate. Step 5: A second antiporter (red oval) exports aspartate to the cytosol in exchange for glutamate. Step 6: A cytosolic transaminase converts aspartate to oxaloacetate and α-ketoglutarate to glutamate, completing the cycle. The blue arrows reflect the movement of the α-ketoglutarate, the red arrows the movement of glutamate, and the black arrows that of $\mathrm{aspartate/malate}$. It is noteworthy that as aspartate and malate cycle clockwise, glutamate and α-ketoglutarate cycle in the opposite direction.

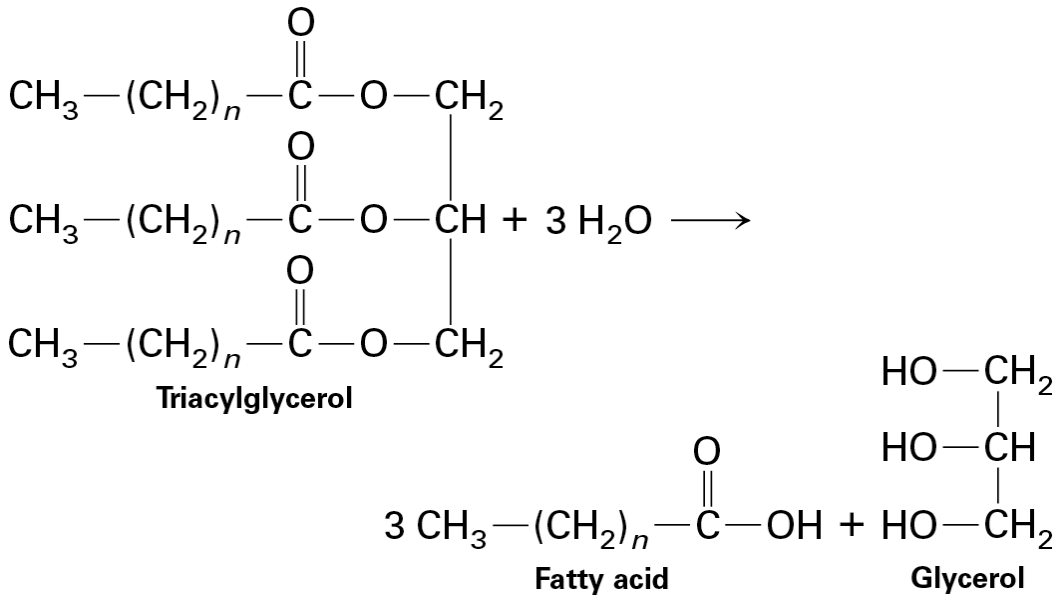
For every complete cycle of the shuttle, there is no overall change in the numbers of NADH and NAD^+ molecules or the intermediates aspartate or malate. In the cytosol, however, NADH is oxidized to NAD^+, which can be used for glycolysis, and in the matrix, NAD^+ is reduced to NADH, which can be used for electron transport:

$$\mathrm{NADH_{cytosol}} + \mathrm{NAD^+{}_{matrix}} \rightarrow \mathrm{NAD^+{}_{cytosol}} + \mathrm{NADH_{matrix}}$$

Mitochondrial Oxidation of Fatty Acids Generates ATP

Up to now, we have focused mainly on the oxidation of carbohydrates, namely glucose, for ATP generation. Fatty acids are another important source of cellular energy. Cells can take up either glucose or fatty acids from the extracellular space with the help of specific transporter proteins (see Chapter 11). Should a cell not need to burn these molecules immediately, it can store them as a polymer

of glucose called *glycogen*, especially in muscle or liver, or as a trimer of fatty acids covalently linked to glycerol, called a **triacylglycerol** or **triglyceride**:

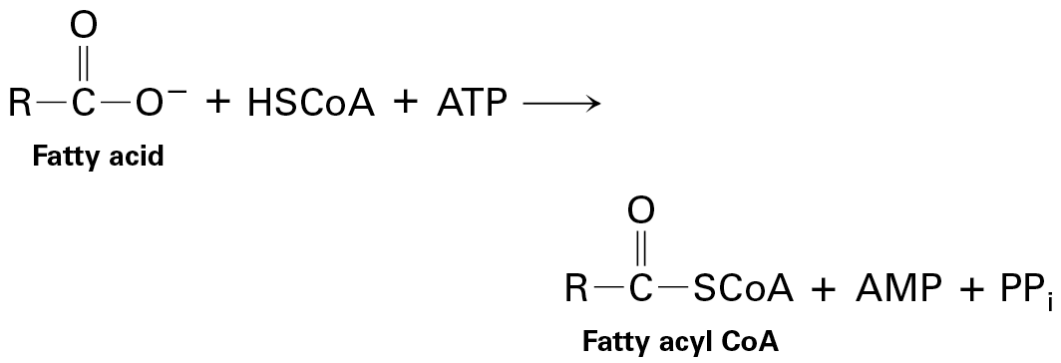


In some cells, excess glucose is converted into fatty acids and then triacylglycerols for storage. However, unlike microorganisms, animals are unable to convert fatty acids to glucose. When the cells need to burn these energy stores to make ATP (e.g., when a resting muscle begins to do work and needs to burn glucose or fatty acids as fuel), enzymes break down glycogen to glucose or hydrolyze triacylglycerols to fatty acids, which are then oxidized to generate ATP.

Fatty acids are the major energy source for some tissues, particularly adult heart muscle. In humans, in fact, more ATP is generated by the oxidation of 1 g of fats than by the oxidation of 1 g

of glucose. The oxidation of 1 g of triacylglycerol to CO_2 generates about six times as much ATP as does the oxidation of 1 g of hydrated glycogen. Thus considering the mass of stored fuel an organism must carry, triglycerides are more efficient than carbohydrates for storage of energy, in part because they are stored in anhydrous form and can yield more energy when oxidized, and in part because they are intrinsically more reduced (have more hydrogens) than carbohydrates. In mammals, the primary site of storage of triacylglycerol is fat (adipose) tissue, whereas the primary sites for glycogen storage are muscle and the liver. In animals, when tissues need to generate a lot of ATP, as in exercising muscle, signals are sent to adipose tissue to hydrolyze triacylglycerols and to release the fatty acids into the circulatory system so that they can move to and be transported into the ATPrequiring tissues.

Just as there are four stages in the oxidation of glucose, there are four stages in the oxidation of fatty acids. To optimize the efficiency of ATP generation, part of stage II (citric acid cycle oxidation of acetyl CoA) and all of stages III and IV of fatty acid oxidation are identical to those of glucose oxidation. The differences lie in cytosolic stage I and in the first part of mitochondrial stage II. In stage I, fatty acids are converted to a fatty acyl CoA in the cytosol in a reaction coupled to the hydrolysis of ATP to AMP and PP_i (inorganic pyrophosphate) (see Figure 12-11):



Subsequent hydrolysis of PP_i to two molecules of P_i releases energy that drives this reaction to completion. To enter the mitochondrial matrix, the fatty acyl group must be covalently transferred to a molecule called *carnitine* and moved across the inner mitochondrial membrane by an acylcarnitine transporter protein (see Figure 12-11, blue oval); then, on the matrix side, the fatty acyl group is released from carnitine and reattached to another CoA molecule. The activity of the acylcarnitine transporter is regulated to prevent oxidation of fatty acids when cells have adequate energy (ATP) supplies.

In the first part of stage II, each molecule of a fatty acyl CoA in the mitochondrion is oxidized in a cyclical sequence of four reactions in which all the carbon atoms are converted, two at a time, to acetyl CoA with generation of $FADH_2$ and NADH (Figure 12-15a). For example, mitochondrial oxidation of each molecule of the 18-carbon stearic acid, $CH_3(CH_2)_{16}COOH$, yields nine molecules of acetyl CoA and eight molecules each of NADH and $FADH_2$. In the second part of stage II, as with acetyl CoA generated from pyruvate, these acetyl groups enter the citric acid cycle and are oxidized to CO_2. As will be

described in detail in the next section, the reduced NADH and $FADH_2$ with their high-energy electrons will be used in stage III to generate a proton-motive force, which in turn is used in stage IV to power ATP synthesis.

Peroxisomal Oxidation of Fatty Acids Generates No ATP

Mitochondrial oxidation of fatty acids is the major source of ATP in mammalian liver cells, and biochemists at one time believed this was true in all cell types. However, rats treated with clofibrate, a drug that affects many features of lipid metabolism, were found to exhibit an increased rate of fatty acid oxidation and a large increase in the number of **peroxisomes** in their liver cells. This finding suggested that peroxisomes, as well as mitochondria, can oxidize fatty acids. These small organelles, 0.2–1 μm in diameter, are lined by a single membrane (see Figure 1-12). They are present in all mammalian cells except erythrocytes and are also found in plant cells, yeasts, and probably most other eukaryotic cells.

Mitochondria preferentially oxidize short-chain [fewer than 8 carbons ($<C_8$)], medium-chain (C_8–C_{12}), and long-chain (C_{14}–C_{20}) fatty acids, whereas peroxisomes preferentially oxidize very long chain fatty acids (VLCFAs$>C_{20}$), which cannot be oxidized by mitochondria. Most dietary fatty acids have long chains, which

means that they are oxidized mostly in mitochondria. In contrast to mitochondrial oxidation of fatty acids, which is coupled to generation of ATP, peroxisomal oxidation of fatty acids is not linked to ATP formation, and energy is released as heat.

The reaction pathway by which fatty acids are degraded to acetyl CoA in peroxisomes is similar to that used in mitochondria (Figure 12-15). However, peroxisomes lack an electron-transport chain, and electrons from the $FADH_2$ produced during the oxidation of fatty acids are immediately transferred to O_2 by *oxidases*, regenerating FAD and forming hydrogen peroxide (H_2O_2). In addition to oxidases, peroxisomes contain abundant *catalase*, which quickly decomposes the H_2O_2, a highly cytotoxic metabolite. NADH produced during peroxisomal oxidation of fatty acids is exported and reoxidized in the cytosol; there is no need for a malate-aspartate shuttle here. Peroxisomes also lack the citric acid cycle, so acetyl CoA generated during peroxisomal degradation of fatty acids cannot be oxidized further; instead, it is transported into the cytosol for use in the synthesis of cholesterol (see Chapter 10) and other metabolites.

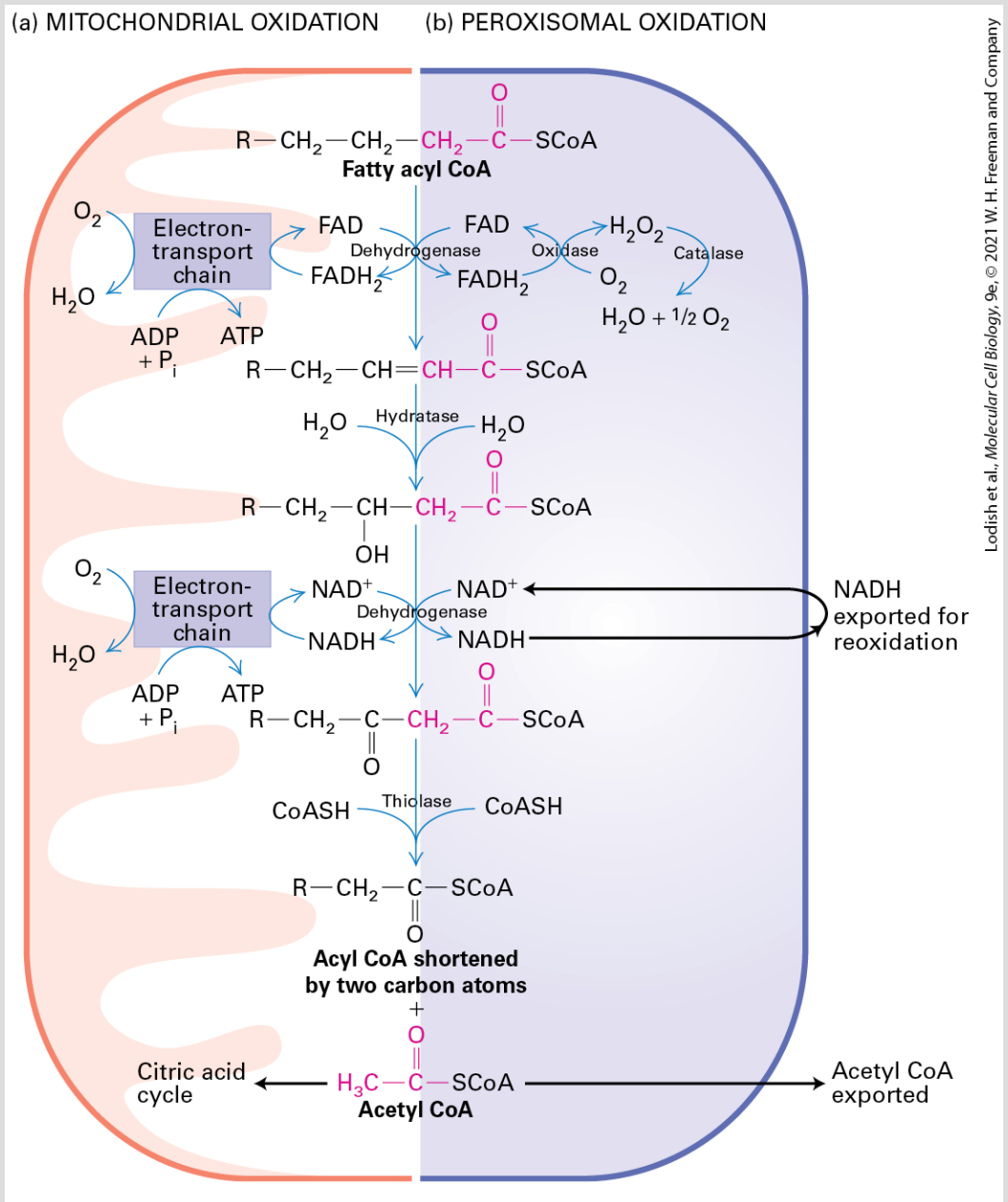


FIGURE 12-15 Oxidation of fatty acids in mitochondria and in peroxisomes. In both mitochondrial oxidation (a) and peroxisomal oxidation (b), fatty acids are converted to acetyl CoA by a series of four enzyme-catalyzed reactions (shown down the center of the figure). A fatty acyl CoA molecule is converted to acetyl CoA and a fatty acyl CoA shortened by two carbon atoms. Concomitantly, one FAD molecule is reduced to $FADH_2$ and one NAD^+ molecule is reduced to NADH. The cycle is repeated on the shortened acyl CoA until fatty acids with an even number of carbon atoms are completely converted to acetyl CoA. In mitochondria (a), electrons from $FADH_2$ and NADH enter the electron-

transport chain and are ultimately used to generate ATP; the acetyl CoA generated is oxidized in the citric acid cycle, resulting in the release of CO_2 and ultimately the synthesis of additional ATP. Because peroxisomes (b) lack the protein complexes composing the electron-transport chain and the enzymes of the citric acid cycle, oxidation of fatty acids in these organelles yields no ATP.

KEY CONCEPTS OF SECTION 12.5

The Citric Acid Cycle and Fatty Acid Oxidation

- In stage II of glucose oxidation, the three-carbon pyruvate molecule is first oxidized to generate one molecule each of CO_2, NADH, and acetyl CoA. The acetyl group of acetyl CoA is then oxidized to CO_2 by the citric acid cycle (see Figure 12-11).
- Each turn of the citric acid cycle releases two molecules of CO_2 and generates three NADH molecules, one $FADH_2$ molecule, and one GTP (see Figure 12-13).
- Most of the energy released in stages I and II of glucose oxidation is temporarily stored in the reduced coenzymes NADH and $FADH_2$, which carry high-energy electrons that subsequently drive the electron-transport chain (stage III).
- Neither glycolysis nor the citric acid cycle directly uses molecular oxygen (O_2).
- The malate-aspartate shuttle regenerates the supply of cytosolic NAD^+ necessary for continued glycolysis (see Figure 12-14).
- Like glucose oxidation, the oxidation of fatty acids takes place in four stages. In stage I, fatty acids are converted to fatty acyl CoA in the cytosol. In stage II, in the mitochondria the fatty acyl CoA is first converted into multiple acetyl CoA molecules, with generation of NADH and $FADH_2$. Then, as in glucose oxidation, the acetyl CoA enters the citric acid cycle. Stages III and IV are identical for fatty acid and glucose oxidation (see Figure 12-11).
- In most eukaryotic cells, oxidation of short- to long-chain fatty acids occurs in mitochondria with production of ATP, whereas oxidation of very long chain fatty acids occurs primarily in peroxisomes and is not linked to ATP production (see Figure 12-15); the energy released during peroxisomal oxidation of fatty acids is converted to heat.

12.6 The Electron-Transport Chain and Generation of the Proton-Motive Force

Most of the energy released during the oxidation of glucose and fatty acids to CO_2 (stages I and II) is converted into high-energy electrons in the reduced coenzymes NADH and $FADH_2$. We now turn to stage III, in which the energy transiently stored in these reduced coenzymes is converted by an electron-transport chain, also known as the **respiratory chain**, into the proton-motive force. We first describe the logic and components of the electron-transport chain. Next we follow the path of electrons as they flow through the chain and describe the mechanism of **proton pumping** across the inner mitochondrial membrane. We conclude this section with a discussion of the magnitude of the proton-motive force produced by electron transport and proton pumping. In Section 12.7, we will see how the proton-motive force is used to synthesize ATP.

Oxidation of NADH and $FADH_2$ Releases a Substantial Amount of Energy

During electron transport, electrons are released from NADH and $FADH_2$ and eventually transferred to O_2, forming H_2O, according to

the following overall reactions:

$$\begin{gathered}\mathrm{NADH} + \mathrm{H}^{+} + {}^{1}/_{2}\mathrm{O}_{2} \rightarrow \mathrm{NAD}^{+} + \mathrm{H}_{2}\mathrm{O}, \\ \Delta G = -52.6 \text{ kcal/mol} \\ \mathrm{FADH}_{2} + {}^{1}/_{2}\mathrm{O}_{2} \rightarrow \mathrm{FAD} + \mathrm{H}_{2}\mathrm{O}, \\ \Delta G = -43.4 \text{ kcal/mol}\end{gathered}$$

Recall that the conversion of 1 glucose molecule to CO_2 via the glycolytic pathway and citric acid cycle yields 10 NADH and 2 FADH_2 molecules (see Table 12-2). Oxidation of these reduced coenzymes has a total $\Delta G^{\circ\prime}$ of -613 kcal/mol $[10(-52.6) + 2(-43.4)]$. Thus of the total potential free energy present in the chemical bonds of glucose (-686 kcal/mol), about 90 percent is conserved in the reduced coenzymes. Why should there be two different coenzymes, NADH and FADH_2? Although many of the reactions involved in glucose and fatty acid oxidation are sufficiently energetic to reduce NAD^{+}, several are not. To capture the energy released by those reactions, they are coupled to reduction of FAD, which requires less energy.

TABLE 12-2 • Net Result of the Glycolytic Pathway and the Citric Acid Cycle

Reaction	CO_2 Molecules Produced	NAD^{+} Molecules Reduced to NADH	FAD Molecules Reduced to FADH_2	ATP (or GTP)

1 glucose molecule to 2 pyruvate molecules	0	2	0	2
2 pyruvates to 2 acetyl CoA molecules	2	2	0	0
2 acetyl CoA to 4 CO_2 molecules	4	6	2	2
Total	6	10	2	4

The energy carried in the reduced coenzymes can be released by oxidizing them. The biochemical challenge faced by the mitochondrion is to transfer, as efficiently as possible, the energy released by this oxidation into the energy in the terminal phosphoanhydride bond in ATP.

$$\mathrm{P_i^{2-}} + \mathrm{H^+} + \mathrm{ADP^{3-}} \rightarrow \mathrm{ATP^{4-}} + \mathrm{H_2O},$$

$$\Delta G = +7.3 \text{ kcal/mol}$$

A relatively simple one-to-one reaction involving reduction of one coenzyme molecule and synthesis of one ATP molecule would be terribly inefficient because the $\Delta G^{\circ\prime}$ for ATP generation from ADP and $\mathrm{P_i}$ is substantially less than that for the coenzyme oxidation, and much energy would be lost as heat. To efficiently recover that energy, the mitochondrion converts the energy of coenzyme

oxidation into a proton-motive force using a series of electron carriers, all but one of which are integral components of the inner membrane (see stage III in Figure 12-11). The proton-motive force can then be used to generate ATP very efficiently.

Electron Transport in Mitochondria Is Coupled to Proton Pumping

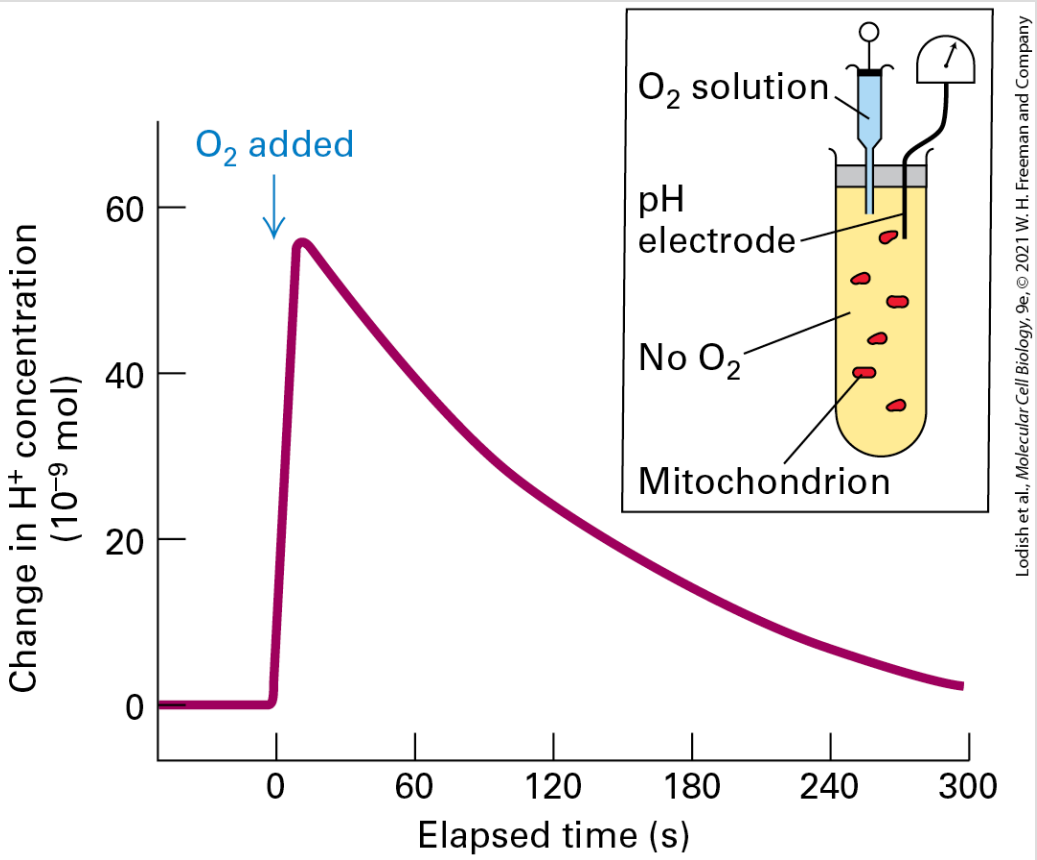
During electron transport from NADH and $FADH_2$ to O_2, protons from the mitochondrial matrix are pumped across the inner membrane. This pumping raises the pH of the mitochondrial matrix relative to the intermembrane space and cytosol and also makes the matrix side of the inner membrane more electrically negative relative to the intermembrane side. In other words, the free energy released during the oxidation of NADH or $FADH_2$ is stored both as a proton concentration gradient and as an electrical voltage gradient across the membrane — collectively known as the *proton-motive force* (see Figure 12-2). As we will see in Section 12.7, the movement of protons back across the inner membrane, driven by this force, is coupled to the synthesis of ATP from ADP and P_i by ATP synthase (stage IV).

The synthesis of ATP from ADP and P_i, driven by the energy released by transfer of electrons from NADH or $FADH_2$ to O_2, is the major source of ATP in aerobic non-photosynthetic cells. Much

evidence shows that in mitochondria and bacteria, this process of oxidative phosphorylation depends on the generation of a proton-motive force across the inner membrane (in mitochondria) or bacterial plasma membrane, with electron transport, proton pumping, and ATP formation occurring simultaneously. In the laboratory, for instance, addition of O_2 and an oxidizable substrate such as pyruvate or succinate to isolated intact mitochondria results in net synthesis of ATP if the inner mitochondrial membrane is intact. In the presence of minute amounts of detergents that make the membrane leaky, electron transport and the oxidation of these metabolites by O_2 still occurs. However, no ATP is made under these conditions because the proton leak prevents the maintenance of the proton-motive force.

The coupling between electron transport from NADH (or $FADH_2$) to O_2 and proton transport across the inner mitochondrial membrane can also be demonstrated experimentally, with isolated, intact mitochondria (Figure 12-16). As soon as O_2 is added to a suspension of mitochondria in an otherwise O_2-free solution that contains NADH, the medium outside the mitochondria transiently becomes more acidic (increased proton concentration) because the mitochondrial outer membrane is freely permeable to protons. (Remember that the malate-aspartate shuttle and other shuttles can convert the NADH in the solution into NADH in the mitochondrial matrix.) Once the O_2 is depleted by its reduction, the excess protons in the medium slowly leak back into the matrix. By measuring the

pH change in such experiments, one can calculate that about 10 protons are transported out of the matrix for every electron pair transferred from NADH to O_2.



EXPERIMENTAL FIGURE 12-16 Electron transfer from NADH to O_2 is coupled to proton transport across the mitochondrial membrane. If NADH is added to a suspension of mitochondria depleted of O_2 (experimental apparatus shown in the inset), no NADH is oxidized. When a small amount of O_2 is added to the system (arrow), there is a sharp rise in the concentration of protons as a function of time in the surrounding medium outside the mitochondria (decrease in pH). Thus the oxidation of NADH by O_2 is coupled to the movement of protons out of the matrix into the intermembrane space and then to the surrounding media. Once the O_2 is depleted, the excess protons slowly move back into the mitochondria (powering the synthesis of ATP), and the pH of the extracellular medium returns to its initial value.

To obtain numbers for $FADH_2$, this experiment can be repeated using succinate instead of NADH as the substrate. (Recall that oxidation of succinate to fumarate in the citric acid cycle generates $FADH_2$; see Figure 12-13.) The amount of succinate added can be adjusted so that the amount of $FADH_2$ generated is equivalent to the amount of NADH in the first experiment. As in the first experiment, addition of O_2 causes the medium outside the mitochondria to become acidic, but less so than with NADH. This difference is not surprising, because electrons in $FADH_2$ have less potential energy (43.4 kcal/mol) than electrons in NADH (52.6 kcal/mol), and thus $FADH_2$ drives the translocation of fewer protons from the matrix and a smaller change in pH.

Electrons Release Energy as They Flow Downhill Through a Series of Electron Carriers

Let's examine more closely the energetically favored movement of electrons from NADH and $FADH_2$ to the final electron acceptor, O_2. For simplicity, we will focus our discussion on NADH. In respiring mitochondria, each NADH molecule releases two electrons to the electron-transport chain; these electrons ultimately reduce one oxygen atom (half of an O_2 molecule), forming one molecule of water:

$$\mathrm{NADH} \rightarrow \mathrm{NAD}^+ + \mathrm{H}^+ + 2\mathrm{e}^-$$
$$2\,\mathrm{e}^- + 2\,\mathrm{H}^+ + {}^1/_2\mathrm{O}_2 \rightarrow \mathrm{H}_2\mathrm{O}$$

As electrons move from NADH to O_2, their electric potential declines by 1.14 V, which corresponds to $26.29\,\mathrm{kcal/mol}$ of electrons transferred, or about $53\,\mathrm{kcal/mol}$ for a pair of electrons. As noted earlier, much of this energy is conserved in the proton-motive force generated across the inner mitochondrial membrane.

Four large multiprotein complexes (complexes I–IV) compose the electron-transport chain in the inner mitochondrial membrane that is responsible for the generation of the proton-motive force (see Figure 12-11, stage III). Each complex contains several **prosthetic groups** that participate in the process of moving electrons from donor molecules to acceptor molecules in coupled oxidation-reduction reactions (see Chapter 2). These small nonpeptide organic molecules or metal ions are tightly and specifically associated with the multiprotein complexes (Table 12-3).

Heme and the Cytochromes

Several types of *heme*, an iron-containing prosthetic group similar to that found in hemoglobin and myoglobin (Figure 12-17a), are tightly bound (covalently or noncovalently) to a set of mitochondrial proteins called **cytochromes**. Each cytochrome is designated by a

letter, such as *a, b, c,* or c_1. Electron flow through the cytochromes occurs by oxidation and reduction of the Fe atom in the center of the heme molecule:

$$Fe^{3+} + e^- \rightleftharpoons Fe^{2+}$$

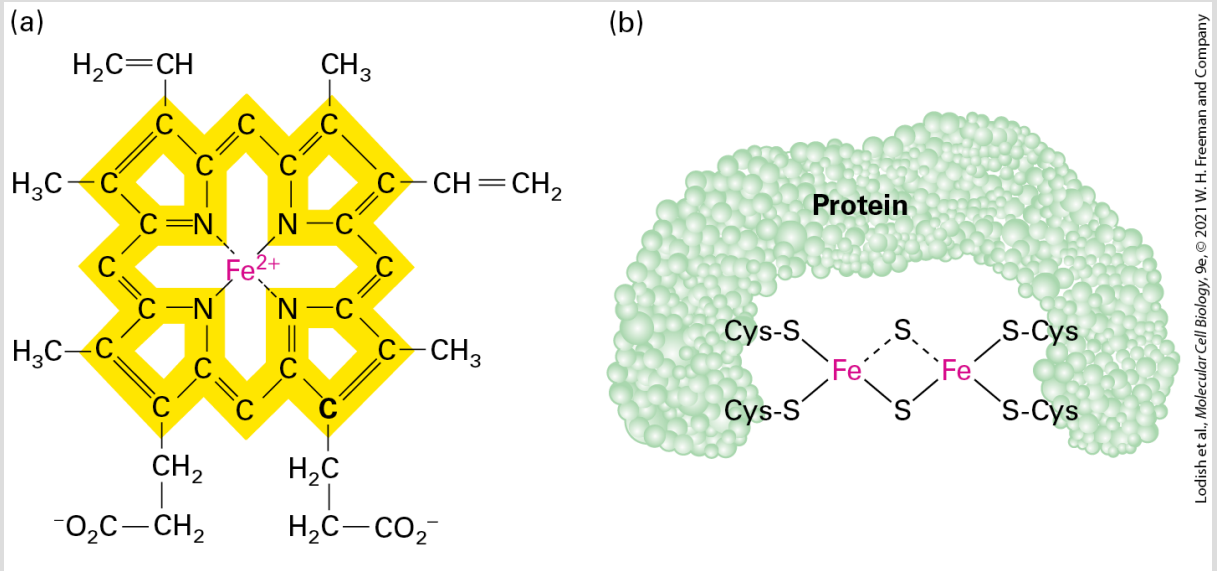


FIGURE 12-17 Heme and iron-sulfur prosthetic groups in the electron-transport chain. (a) Heme portion of cytochromes b_L and b_H, which are components of $CoQH_2$–cytochrome *c* reductase (complex III). The same porphyrin ring (yellow) is present in all hemes. The chemical substituents attached to the porphyrin ring differ in the other cytochromes in the electron-transport chain. All hemes accept and release one electron at a time. (b) Dimeric iron-sulfur cluster (Fe-S). Each Fe atom is bonded to four S atoms: two are inorganic sulfur, and two are in cysteine side chains of the associated protein. All Fe-S clusters accept and release one electron at a time.

Because the heme ring in cytochromes consists of alternating double- and single-bonded atoms, a large number of resonance hybrid forms exist. These forms allow the extra electron delivered to

the cytochrome to be spread throughout the heme carbon and nitrogen atoms as well as the Fe ion.

The various cytochromes each have slightly different heme groups and surrounding atoms, which generate different environments for the Fe ion. Therefore, each cytochrome has a different reduction potential, or tendency to accept an electron — an important property that dictates the unidirectional, energetically downhill electron flow along the chain. Just as water spontaneously flows downhill from a higher to a lower potential energy state — but not uphill — electrons flow in only one direction from one heme (or other prosthetic group) to another due to their differing reduction potentials. (For more on the concept of reduction potential, *E*, see Chapter 2.) As we will see shortly, all the cytochromes except cytochrome *c* are components of integral membrane multiprotein complexes in the inner mitochondrial membrane.

Iron-Sulfur Clusters

Iron-sulfur clusters are nonheme, iron-containing prosthetic groups consisting of Fe atoms bonded both to inorganic sulfur (S) atoms and to S atoms on cysteine residues in a protein (Figure 12-17b). Some Fe atoms in the cluster bear a $+2$ charge; others have a $+3$ charge. However, the net charge of each Fe atom is actually between $+2$ and $+3$, because electrons in their outermost orbitals, together with the extra electron delivered via the transport chain, are

dispersed among the Fe atoms and move rapidly from one atom to another. Iron-sulfur clusters accept and release electrons one at a time.

Coenzyme Q

Also called **ubiquinone, Coenzyme Q (CoQ)** is the only small-molecule electron carrier in the electron-transport chain that is not an essentially irreversibly protein-bound prosthetic group (Figure 12-18). It is a carrier of both protons and electrons. The oxidized quinone form of CoQ can accept a single electron to form a semiquinone, a negatively charged free radical denoted by $CoQ^{\bullet -}$, where the superscript dot ($^{\bullet}$) indicates an unpaired and reactive electron. Addition of a second electron and two protons (thus a total of two hydrogen atoms) to $CoQ^{\bullet -}$ forms dihydroubiquinone ($CoQH_2$), the fully reduced form. Both CoQ and $CoQH_2$ are hydrophobic, and thus are soluble in phospholipids and diffuse freely in the hydrophobic center of the inner mitochondrial membrane. These properties underlie ubiquinone's role in the electron-transport chain: carrying electrons and protons between the membrane-embedded protein complexes of the chain.

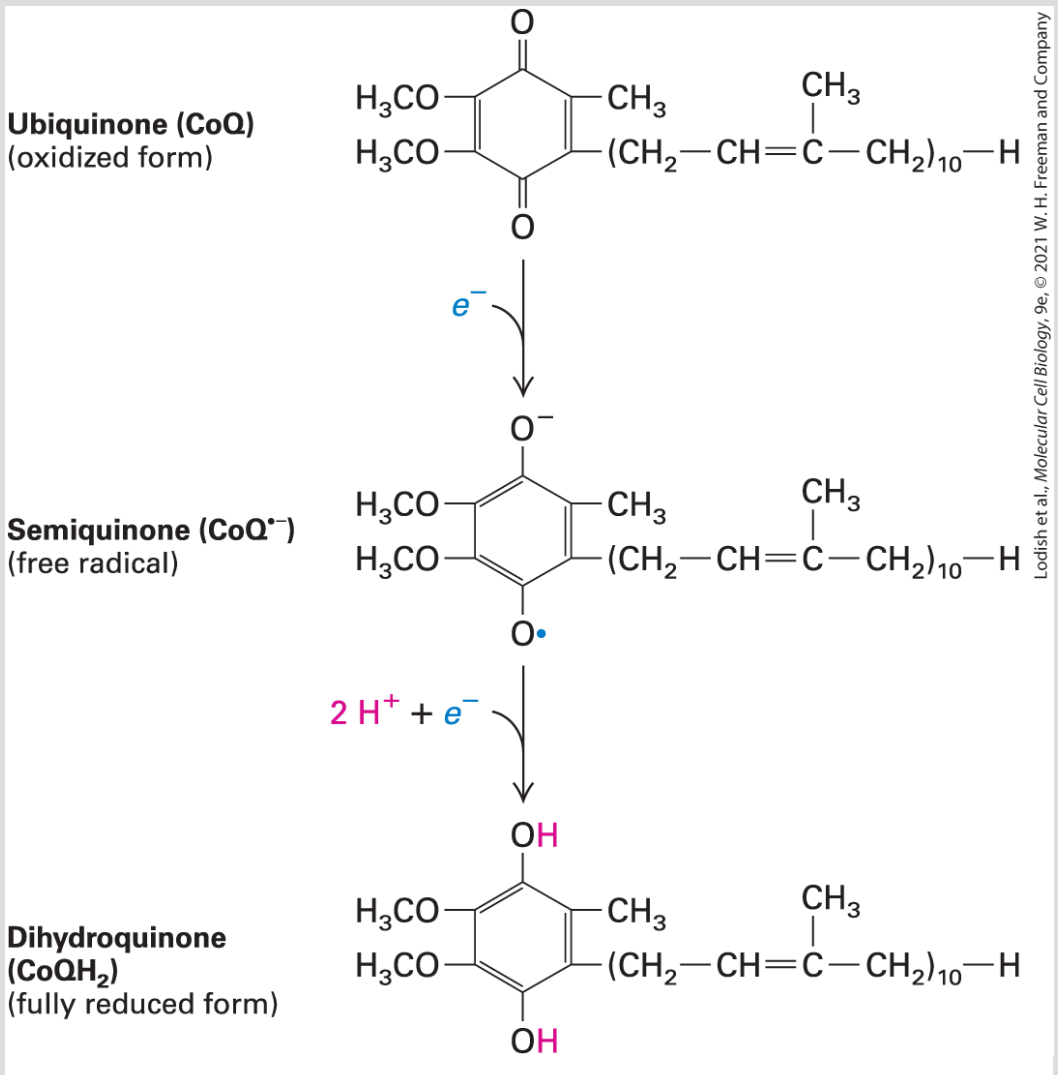


FIGURE 12-18 Oxidized and reduced forms of coenzyme Q (CoQ), which can carry two protons and two electrons. Because of its long hydrocarbon "tail" of isoprene units, CoQ, also called *ubiquinone*, is soluble in the hydrophobic core of phospholipid bilayers and is very mobile. Reduction of CoQ to the fully reduced form, QH_2 (dihydroquinone), occurs in two steps with a half-reduced free-radical intermediate, called *semiquinone*.

Four Large Multiprotein Complexes (I–IV) Couple Electron Transport to

Proton Pumping Across the Inner Mitochondrial Membrane

We now consider in detail the multiprotein complexes that use these prosthetic groups and the paths taken by electrons and protons as they pass through these complexes.

As electrons flow downhill from one electron carrier to the next in the electron-transport chain, the energy released is used to power the pumping of protons out of the matrix against their electrochemical gradient across the inner mitochondrial membrane. Four large multiprotein complexes (Figure 12-19) directly or indirectly couple the movement of electrons to proton pumping: *NADH-CoQ reductase* (complex I, >40 protein subunits), *succinate-CoQ reductase* (complex II, 4 subunits), $CoQH_2$*–cytochrome* c *reductase* (complex III, 11 subunits), and *cytochrome* c *oxidase* (complex IV, 13 subunits). The electrons follow one of two routes through these complexes: I $\rightarrow$ III $\rightarrow$ IV or II $\rightarrow$ III $\rightarrow$ IV. In addition, complexes I, III, and IV all pump protons directly across the inner membrane, whereas complex II does not (see Figure 12-19). Electrons from NADH flow from complex I via $\mathrm{CoQ/CoQH_2}$ to complex III and then, via the soluble protein cytochrome *c* (cyt *c*), to complex IV which reduces molecular oxygen (complex II is bypassed) (see Figure 12-19a). Electrons from $\mathrm{FADH_2}$ flow from complex II via $\mathrm{CoQ/CoQH_2}$ to complex III (see Figure 12-19b) and

then via cytochrome *c* to complex IV to reduce molecular oxygen (complex I is bypassed).

As shown in Figure 12-19, CoQ accepts electrons released from NADH-CoQ reductase (complex I) or succinate-CoQ reductase (complex II) and donates them to $CoQH_2$–cytochrome *c* reductase (complex III). Protons are simultaneously transported from the matrix side of the membrane (also called the *cytosolic side*) to the intermembrane space (also called the *exoplasmic side*). Whenever CoQ accepts electrons, it does so at a binding site on the matrix side of a protein complex, always picking up protons from the medium there. Whenever $CoQH_2$ releases its electrons, it does so at a site on the intermembrane space side of a protein complex, releasing protons into the fluid of the intermembrane space. Thus the transport of each pair of electrons by CoQ is obligately coupled to the movement of two protons from the matrix to the intermembrane space.

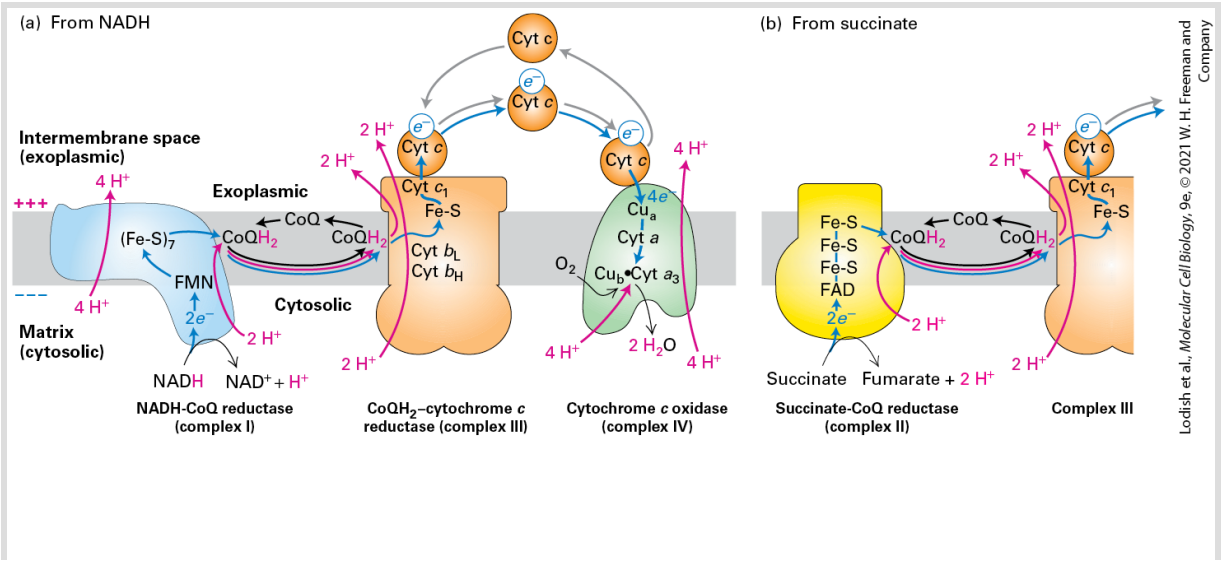


FIGURE 12-19 The mitochondrial electron-transport chain. Electrons (blue arrows) flow through four major multiprotein complexes (I–IV). Electron movement between complexes is mediated either by the lipid-soluble molecule coenzyme Q (CoQ, oxidized form; $\mathrm{CoQH_2}$, reduced form) or the water-soluble protein cytochrome *c* (cyt *c*) in the intermembrane space. Three of the multiprotein complexes use the energy released from the electrons to pump protons (red arrows) from the matrix (cytosolic compartment in bacteria) to the intermembrane space (exoplasmic space of bacteria). (a) Pathway from NADH. Two electrons from NADH $(2/\mathrm{NADH\ oxidized})$ flow through complex I, initially via a flavin mononucleotide (FMN) and then one at a time via seven iron-sulfur clusters (Fe-S), to CoQ, to which two protons bind, forming $\mathrm{CoQH_2}$. Conformational changes in complex I that accompany the electron flow drive proton pumping from the matrix to the intramembrane space. Electrons then flow via the released (and subsequently recycled) $\mathrm{CoQH_2}$ to complex III, and then via cyt *c* to complex IV. Each of four cyt *c* molecules carrying four electrons originating from two NADHs transfers individual electrons to complex IV for the four-electron reduction of one $\mathrm{O_2}$ molecule to two $\mathrm{H_2O}$ molecules and transport of four protons. Thus for every two NADHs oxidized and one $\mathrm{O_2}$ molecule reduced, a total of 20 protons are translocated out of the matrix into the intermembrane space. (b) Pathway from succinate. Two electrons flow from each succinate to complex II via $\mathrm{FAD/FADH_2}$ and iron-sulfur clusters (Fe-S), from complex II to complex III via $\mathrm{CoQ/CoQH_2}$, and then to complex IV via cyt *c*. Electrons released during oxidation of succinate to fumarate in complex II are used to reduce CoQ to $\mathrm{CoQH_2}$ without translocating additional protons. The remainder of electron transport from $\mathrm{CoQH_2}$ proceeds by the same pathway as for the NADH pathway in (a). Thus for every two succinates oxidized and one $\mathrm{O_2}$ molecule reduced,12 protons (8 via the Q-cycle of complex III that is described below, 4 via complex IV) are translocated.

NADH-CoQ Reductase (Complex I)

Electrons are transferred from NADH to CoQ by NADH-CoQ reductase (see Figure 12-19a). Electron microscopy and x-ray crystallography of complex I from bacteria (mass $\sim$500 kDa, with 14 subunits) and from eukaryotes ($\sim$1.7 MDa, with 14 highly conserved core subunits, which are also present in bacteria, plus about 26–32

accessory subunits) has established that it is L-shaped (Figure 12-20a). The membrane-embedded arm of the L is slightly curved, about 180 Å long, and comprises proteins with more than 60 transmembrane α helices. This arm has four subdomains, three of which contain proteins that are members of a family of cation antiporters. The hydrophilic peripheral arm extends over 95 Å away from the membrane into the matrix (cytosolic) space.

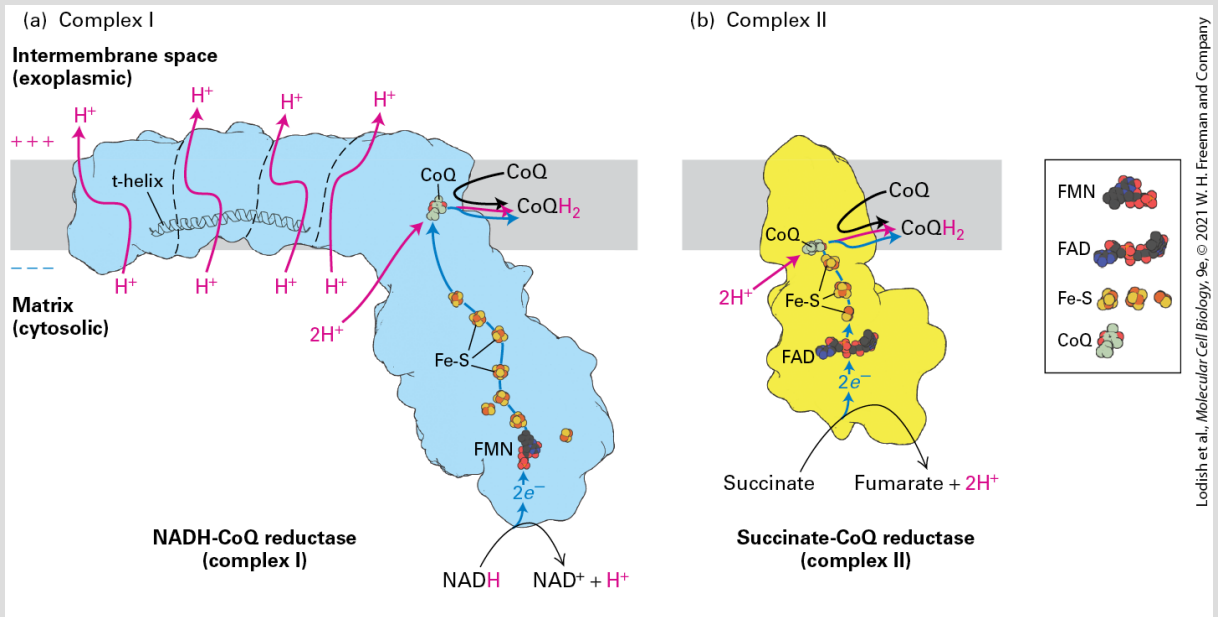


FIGURE 12-20 Electron and proton transport through complexes I and II. (a) Model of complex I based on its three-dimensional structure. The outline of the shape of the core complex I from the yeast *Yarrowia lipolytica*, as determined by x-ray crystallography, is shown in light blue, and the borders separating several distinct structural subunits are indicated by thin dashed black lines. From NADH, electrons flow first to a flavin mononucleotide (FMN) and then, via iron-sulfur clusters (Fe-S, red and yellow balls), to CoQ, to which two protons from the matrix bind (red arrow) to form $\mathrm{CoQH_2}$. Conformational changes due to the electron flow, which may include a horizontal movement of the t-helix, drive proton pumping through the transmembrane subunits from the matrix to the intramembrane space (red arrows). (b) Model of complex II based on its three-dimensional structure. Electrons flow through complex II from succinate to CoQ via $\mathrm{FAD/FADH_2}$ and

iron-sulfur clusters (Fe-S), and from complex II to complex III via $\mathrm{CoQ/CoQH_2}$. Electrons released during oxidation of succinate to fumarate in complex II are used to reduce CoQ to $\mathrm{CoQH_2}$ without translocating additional protons.

[Part (a) Data from V. Zickermann et al., 2015, *Science* **347**:44–49, PDB ID 3m9s. Part (b) data from F. Sun et al., 2005, *Cell* **121**:1043–1057, PDB ID 1zoy.]

$\mathrm{NAD^+}$ is exclusively a two-electron carrier: it accepts (or for NADH releases) a pair of electrons simultaneously. In NADH-CoQ reductase (complex I), the NADH-binding site is at the tip of the peripheral arm (see Figure 12-20a); electrons released from NADH first flow to FMN (flavin mononucleotide), a prosthetic group, or cofactor, related to FAD, then are shuttled about 95 Å down the peripheral arm through a series of iron-sulfur clusters and finally to CoQ, which is bound at a site at least partially in the plane of the membrane. FMN, like FAD, can accept two electrons, but does so one electron at a time.

Each transported electron undergoes a drop in potential of about 360 mV, equivalent to a $\Delta G^{\circ\prime}$ of $-16.6\,\mathrm{kcal/mol}$ for the two electrons transported. Much of this released energy is used to transport four protons across the inner membrane per molecule of NADH oxidized by complex I. These four protons are distinct from the two protons that are transferred to the CoQ as illustrated in Figures 12-18, 12-19a, and 12-20a. The precise mechanism by which the energy released by electron transport in the peripheral arm is used to change the conformation of subunits in the membrane arm and thus mediate the movement of four protons across the membrane is uncertain.

Three protons are likely to pass through the three cation antiporter domains via a zigzag series of polar side chains that can be protonated and that span the membrane. A similar series of side chains is the likely conduit for the fourth proton. A transverse α helix (t-helix) in the membrane arm runs parallel to the plane of the membrane, potentially mechanically linking the antiporter-like domains to the peripheral arm (see Figure 12-20a) and possibly contributing directly to the conformational changes required to convert the energy released by the electron transport in the peripheral arm into proton transport by the transmembrane domains.

The overall reaction catalyzed by this complex is

$$\underset{\text{(Reduced)}}{\text{NADH}} + \underset{\text{(Oxidized)}}{\text{CoQ}} + 6\,\text{H}^+{}_{\text{matrix}} \rightarrow \underset{\text{(Oxidized)}}{\text{NAD}^+} + \text{H}^+{}_{\text{in}} + \underset{\text{(Reduced)}}{\text{CoQH}_2} + 4\,\text{H}^+{}_{\text{interme}}$$

Succinate-CoQ Reductase (Complex II)

Succinate dehydrogenase, the enzyme that oxidizes a molecule of succinate to fumarate in the citric acid cycle (and in the process generates the reduced coenzyme FADH_2), is one of the four subunits of complex II (Figure 12-20b). Thus the citric acid cycle is physically as well as functionally linked to the electron-transport chain. The two electrons released in the conversion of succinate to fumarate are transferred first to FAD in succinate dehydrogenase,

then to iron-sulfur clusters — regenerating FAD — and finally to CoQ, which binds to a cleft on the matrix side of the transmembrane portions of complex II (see Figures 12-19b and 12-20b). The pathway is somewhat reminiscent of that in complex I (see Figure 12-20a).

The overall reaction catalyzed by this complex is

$$\underset{\text{(Reduced)}}{\text{NADH}} + \underset{\text{(Oxidized)}}{\text{CoQ}} + 6\,\text{H}^+{}_{\text{matrix}} \rightarrow \underset{\text{(Oxidized)}}{\text{NAD}^+} + \text{H}^+{}_{\text{in}} + \underset{\text{(Reduced)}}{\text{CoQH}_2} + 4\,\text{H}^+{}_{\text{intermem}}$$

Although the $\Delta G^{\circ\prime}$ for this reaction is negative, the released energy is insufficient for proton pumping in addition to reduction of CoQ to form CoQH_2. Thus no protons are translocated directly across the membrane by succinate-CoQ reductase (complex II), and no proton-motive force is generated in this part of the electron-transport chain. We will see shortly how the protons and electrons in the CoQH_2 molecules generated by complexes I and II contribute to the generation of the proton-motive force.

Complex II generates CoQH_2 from succinate via FAD/FADH_2-mediated redox reactions. Another set of proteins in the matrix and inner mitochondrial membrane performs a comparable set of FAD/FADH_2-mediated redox reactions. As with succinate dehydrogenase, these proteins, described in detail in the following paragraph and Figure 12-15, use FADH_2 to produce CoQH_2, which contributes electrons to the electron-transport chain.

Fatty acyl–CoA dehydrogenase (see Figure 12-15), which is a water-soluble enzyme, catalyzes the first step of the oxidation of fatty acyl CoA in the mitochondrial matrix. There are several fatty acyl–CoA dehydrogenase enzymes with specificities for fatty acyl chains of different lengths. These enzymes mediate the initial step in a four-step process that removes two carbons from the fatty acyl group by oxidizing the carbon in the β position of the fatty acyl chain (thus the entire process is often referred to as *β-oxidation*). These reactions generate acetyl CoA, which in turn enters the citric acid cycle. They also generate an $FADH_2$ intermediate and NADH. The $FADH_2$ generated remains bound to the enzyme during the redox reaction, as is the case for complex II. A water-soluble protein called *electron transfer flavoprotein (ETF)* transfers the high-energy electrons from the $FADH_2$ in the fatty acyl–CoA dehydrogenase to *electron transfer flavoprotein:ubiquinone oxidoreductase (ETF:QO)*, a membrane protein that reduces CoQ to $CoQH_2$ in the inner membrane. This $CoQH_2$ intermixes in the membrane with the other $CoQH_2$ molecules generated by complexes I and II, all contributing to proton transport out of the matrix by complex III.

$CoQH_2$–Cytochrome *c* Reductase (Complex III)

A $CoQH_2$ generated by complex I, complex II, or ETF:QO donates two electrons to $CoQH_2$–cytochrome *c* reductase (complex III), regenerating oxidized CoQ. Concomitantly, it releases into the intermembrane space two protons previously picked up by CoQ on

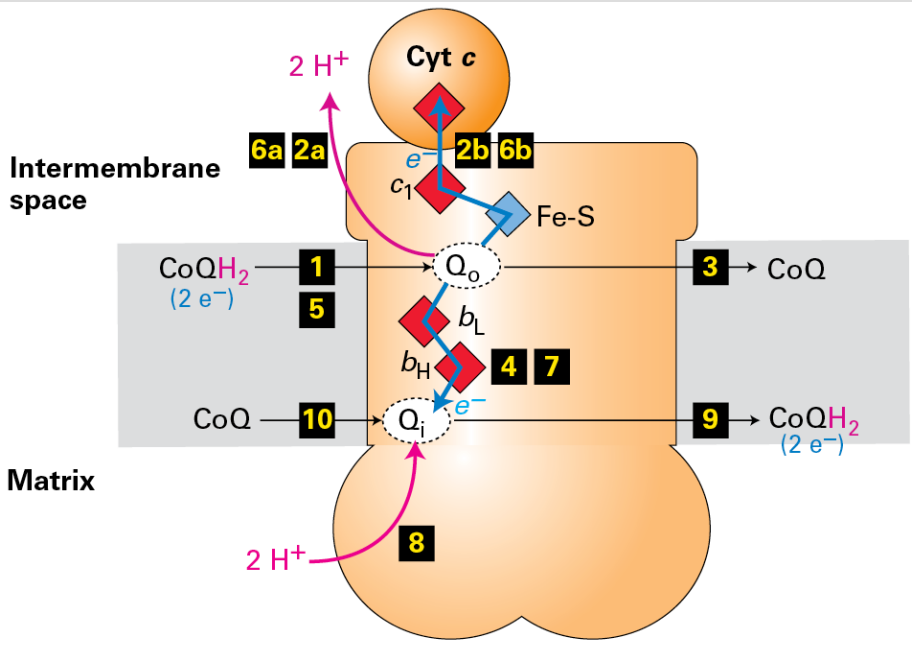
the matrix face, generating part of the proton-motive force (see Figure 12-19). Within complex III, the released electrons are first transferred to an iron-sulfur cluster within the complex and then to cytochrome c_1 or to two *b*-type cytochromes (b_L and b_H, see the discussion of the Q cycle below). Finally, the two electrons are transferred sequentially to two molecules of the oxidized form of cytochrome *c*, a water-soluble peripheral protein that diffuses in the intermembrane space. For each pair of electrons transferred, the overall reaction catalyzed by complex III is

$$\underset{\text{(Reduced)}}{\mathrm{CoQH_2}} + \underset{\text{(Oxidized)}}{2\,\mathrm{Cyt}\,c^{3+}} + 2\,\mathrm{H^+}_{\mathrm{in}} \rightarrow \underset{\text{(Oxidized)}}{\mathrm{CoQ}} + 4\,\mathrm{H^+}_{\mathrm{out}} + \underset{\text{(Reduced)}}{2\,\mathrm{Cyt}\,c^{2+}}$$

The $\Delta G^{\circ\prime}$ for this reaction is sufficiently negative that two protons in addition to those from $\mathrm{CoQH_2}$ are translocated from the mitochondrial matrix across the inner membrane for each pair of electrons transferred; this transfer involves the proton-motive Q cycle, discussed immediately below. The heme protein cytochrome *c* and the small lipid-soluble molecule CoQ play similar roles in the electron-transport chain in that they both serve as mobile electron shuttles, transferring electrons (and thus energy) between the complexes of the electron-transport chain.

The Q Cycle

Experiments have shown that four protons are translocated across the inner mitochondrial membrane per electron pair transported from $CoQH_2$ through complex III. These four protons are those carried on two $CoQH_2$ molecules, which are converted to two CoQ molecules during the cycle. However, during this process another CoQ molecule receives two other protons from the matrix and is converted to one $CoQH_2$ molecule. Thus the net overall reaction involves the conversion of only one $CoQH_2$ molecule to CoQ as two electrons are transferred, one at a time, to two molecules of the acceptor cytochrome *c*. An evolutionarily conserved mechanism, called the **Q cycle**, is responsible for the two-for-one transport of protons and electrons by complex III (Figure 12-21).



$CoQH_2$–cytochrome *c* reductase (complex III)

At Q_o site: $2\ CoQH_2 + 2\ Cyt\ c^{3+} \longrightarrow$
$(4\ H^+, 4\ e^-)$
$2\ CoQ + 2\ Cyt\ c^{2+} + 2\ e^- + 4\ H^+_{\text{intermembrane space/exoplasmic side}}$
$(2\ e^-)$

At Q_i site: $CoQ + 2\ e^- + 2\ H^+_{\text{matrix/cytosolic side}} \longrightarrow CoQH_2$
$(2\ H^+, 2\ e^-)$

Net Q cycle (sum of reactions at Q_o and Q_i):

$CoQH_2 + 2\ Cyt\ c^{3+} + 2\ H^+_{\text{matrix/cytosolic side}} \longrightarrow$
$(2\ H^+, 2\ e^-)$
$CoQ + 2\ Cyt\ c^{2+} + 4\ H^+_{\text{intermembrane space/exoplasmic side}}$
$(2\ e^-)$

Per $2\ e^-$ transferred through complex III to cytochrome *c*, $4\ H^+$ released to the intermembrane space

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 12-21 The Q cycle. The Q cycle of complex III uses the net oxidation of one $CoQH_2$ molecule to transfer four protons into the intermembrane space and two electrons to two cytochrome *c* molecules. The cycle begins (step **1**) when a molecule from the combined pool of reduced $CoQH_2$ in the inner mitochondrial membrane binds to the Q_o site on the *intermembrane space (exoplasmic) side* of the transmembrane portion of complex III. There in step **2a** $CoQH_2$ releases two protons into the intermembrane space and in step **3** two electrons and the resulting CoQ dissociate. In step **2b**, one of the

electrons is transported, via an iron-sulfur protein and cytochrome c_1, directly to cytochrome *c*. (Recall that each cytochrome *c* shuttles one electron from complex III to complex IV.) In step **4**, the other electron moves through cytochromes b_L and b_H and partially reduces an oxidized CoQ molecule already bound to the second, Q_i, site on the *matrix (cytosolic) side* of the complex, forming a CoQ semiquinone anion, $Q^{\bullet -}$. The process is then repeated with step **5**: the binding of a second $CoQH_2$ at the Q_o site, step **6a**: proton release, step **6b**: reduction of another cytochrome *c*, and step **7**: addition of the other electron to the $Q^{\bullet -}$ bound at the Q_i site. In steps **8** and **9** the addition of two protons from the matrix yields a fully reduced $CoQH_2$ molecule at the Q_i site, which then dissociates, that leads to step **10**, the freeing the Q_i to bind a new molecule of CoQ and the Q cycle begins over again. See B. Trumpower, 1990, *J. Biol. Chem.* **265**:11409; and E. Darrouzet et al., 2001, *Trends Biochem. Sci*. **26**:445.

The substrate for complex III, $CoQH_2$, is generated by several enzymes, including NADH-CoQ reductase (complex I), succinate-CoQ reductase (complex II), ETF:QO (during β-oxidation), and as we shall see, by complex III itself.

As shown in Figure 12-21, in one turn of the Q cycle, two molecules of $CoQH_2$ are oxidized to CoQ at the Q_o site near the intermembrane space in complex III and release a total of four protons into the intermembrane space. In addition, at a different CoQ binding site near the matrix side of complex III called the *Q_i site*, one molecule of $CoQH_2$ is regenerated from CoQ and two additional protons from the matrix. The translocated protons are all derived from $CoQH_2$, which obtained its protons from the matrix, as described above. Although seemingly cumbersome, the Q cycle optimizes the number

of protons pumped per pair of electrons moving through complex III. The Q cycle is found in all plants and animals as well as in bacteria. Its formation at a very early stage of cellular evolution was probably essential for the success of all life forms as a way of converting the potential energy in reduced coenzyme Q ($CoQH_2$) into the maximum proton-motive force across a membrane. In turn, this process maximizes the number of ATP molecules synthesized from each electron that moves down the electron-transport chain from NADH or $FADH_2$ to O_2.

How are the two electrons released from $CoQH_2$ at the Q_o site directed to different acceptors, either to Fe-S, cytochrome c_1, and then cytochrome *c* (upward pathway in Figure 12-21), or alternatively to cytochrome b_L, cytochrome b_H, and then CoQ at the Q_i site (downward pathway in Figure 12-21)? The mechanism involves a flexible hinge in the Fe-S–containing protein subunit of complex III. Initially, the Fe-S cluster is close enough to the Q_o site to pick up an electron from $CoQH_2$ bound there. Once this happens, the subunit containing this Fe-S cluster swings the cluster away from the Q_o site to a position near enough to the heme on cytochrome c_1 for electron transfer to occur. With the Fe-S–containing subunit in this alternate conformation, the second electron released from the $CoQH_2$ bound to the Q_o site cannot move to the Fe-S cluster — it is too far away, so it takes an alternative path open to it via a somewhat less thermodynamically favored route to cytochrome b_L and through cytochrome b_H to the CoQ at the Q_i site.

Cytochrome *c* Oxidase (Complex IV)

Cytochrome *c*, after being reduced by one electron from complex III, is reoxidized when it transports its electron to cytochrome *c* oxidase (complex IV) (see Figure 12-19a). Mitochondrial cytochrome *c* oxidases contain 13 different subunits, but the catalytic core of the enzyme consists of only three. The functions of the remaining subunits are not well understood. Bacterial cytochrome *c* oxidases contain only the three catalytic subunits. In both mitochondria and bacteria, four molecules of reduced cytochrome *c* bind, one at a time, to the oxidase. Figure 12-19a shows how an electron is transferred from the heme of each cytochrome *c*, first to the pair of copper ions labeled $Cu_a{}^{2+}$, then to the heme in cytochrome *a*. Next the electron moves to the oxygen reduction center, composed of $Cu_b{}^{2+}$ and the heme in cytochrome a_3, which together bind in a sandwich fashion and sequester the O_2 molecule that will be reduced to two H_2Os. Several lines of evidence suggest that O_2 reaches the reduction center via one or more hydrophobic channels through the protein, into which O_2 flows from the hydrophobic core of the membrane, where it is more soluble than in aqueous solution (see Figure 12-19a).

The four electrons are finally passed by complex IV to O_2, the ultimate electron acceptor, yielding two H_2O, which together with CO_2 is one of the end products of the overall oxidation pathway. Proposed intermediates in oxygen reduction include the peroxide

anion ($O_2{}^{2-}$) and the hydroxyl radical ($OH^{\bullet}$), as well as unusual complexes of iron and oxygen atoms. These intermediates would be harmful to the cell if they escaped from complex IV, but they do so only rarely because they are sequestered by the $Cu_b{}^{2+}$ and the heme in cytochrome a_3. To generate H_2O from the reduced O_2, protons are channeled to the oxygen reduction center. In addition, other protons are transported across the membrane from the matrix to the intermembrane space. During the transport of four electrons through the cytochrome *c* oxidase complex, four protons from the matrix are translocated across the membrane. Thus complex IV transports only one proton per electron transferred, whereas complex III, using the Q cycle, transports two protons per electron transferred. However, the mechanism by which complex IV translocates these protons and energetically couples the translocation to O_2 reduction is not fully understood.

For each four electrons transferred, the overall reaction catalyzed by cytochrome *c* oxidase is

$$\underset{\text{(Reduced)}}{4\,\text{cyt}\,c^{2+}} + 8\,H^{+}{}_{\text{matrix}} + O_2 \rightarrow \underset{\text{(Oxidized)}}{4\,\text{cyt}\,c^{3+}} + 2\,H_2O + 4\,H^{+}{}_{\text{intermembrane sp}}$$

The poison cyanide, which has been used as a chemical warfare agent, by spies to commit suicide when captured, in gas chambers to execute prisoners, and by the Nazis (Zyklon B gas) for the mass murder of Jews and others, is toxic because it binds to the heme a_3

in mitochondrial cytochrome *c* oxidase (complex IV), inhibiting electron transport and thus oxidative phosphorylation and production of ATP. Cyanide is one of many toxic small molecules that interfere with energy production in mitochondria. ■

The Reduction Potentials of Electron Carriers in the Electron-Transport Chain Favor Electron Flow from NADH to O_2

As we saw in Chapter 2, the **reduction potential (*E*)** for a partial reduction reaction

$$\text{Oxidized molecule} + e^- \rightleftharpoons \text{reduced molecule}$$

is a measure of the equilibrium constant of that partial reaction. With the exception of the *b* cytochromes in complex III ($CoQH_2$–cytochrome *c* reductase), the standard reduction potential $E^{\circ\prime}$ of the electron carriers in the electron-transport chain increases steadily from NADH to O_2. For instance, for the partial reaction

$$NAD^+ + H^+ + 2\,e^- \rightleftharpoons NADH$$

the value of the standard reduction potential is $-320\,\mathrm{mV}$, which is equivalent to a $\Delta G^{\circ\prime}$ of $+14.8\,\mathrm{kcal/mol}$ for transfer of two electrons. Thus this partial reaction tends to proceed toward the left: that is, toward the oxidation of NADH to $\mathrm{NAD^+}$.

In contrast, the standard reduction potential for the partial reaction

$$\text{Cytochrome } c_{\mathrm{ox}}\,(\mathrm{Fe^{3+}}) + \mathrm{e^-} \rightleftharpoons \text{ cytochrome } c_{\mathrm{red}}\,(\mathrm{Fe^{2+}})$$

is $+220\,\mathrm{mV}$ ($\Delta G^{\circ\prime} = -5.1\,\mathrm{kcal/mol}$) for transfer of one electron. Thus this partial reaction tends to proceed toward the right: that is, toward the reduction of cytochrome *c* ($\mathrm{Fe^{3+}}$) to cytochrome *c* ($\mathrm{Fe^{2+}}$).

The final reaction in the electron-transport chain, the reduction of $\mathrm{O_2}$ to $\mathrm{H_2O}$

$$2\,\mathrm{H^+} + {}^1\!/_2\mathrm{O_2} + 2\,\mathrm{e^-} \rightarrow \mathrm{H_2O}$$

has a standard reduction potential of $+816\,\mathrm{mV}$ ($\Delta G^{\circ\prime} = -37.8\,\mathrm{kcal/mol}$ for transfer of two electrons), the most positive in the whole series; thus this reaction also tends to proceed toward the right.

As illustrated in Figure 12-22, the steady increase from the top left to the bottom right in $E^{\circ\prime}$ values, and the corresponding decrease in

$\Delta G^{\circ\prime}$ values of the carriers in the electron-transport chain favors the flow of electrons from NADH and $FADH_2$ (generated from succinate) to O_2. The energy released as electrons flow energetically downhill through the electron-transport chain complexes drives the pumping of protons against their concentration gradient across the inner mitochondrial membrane.

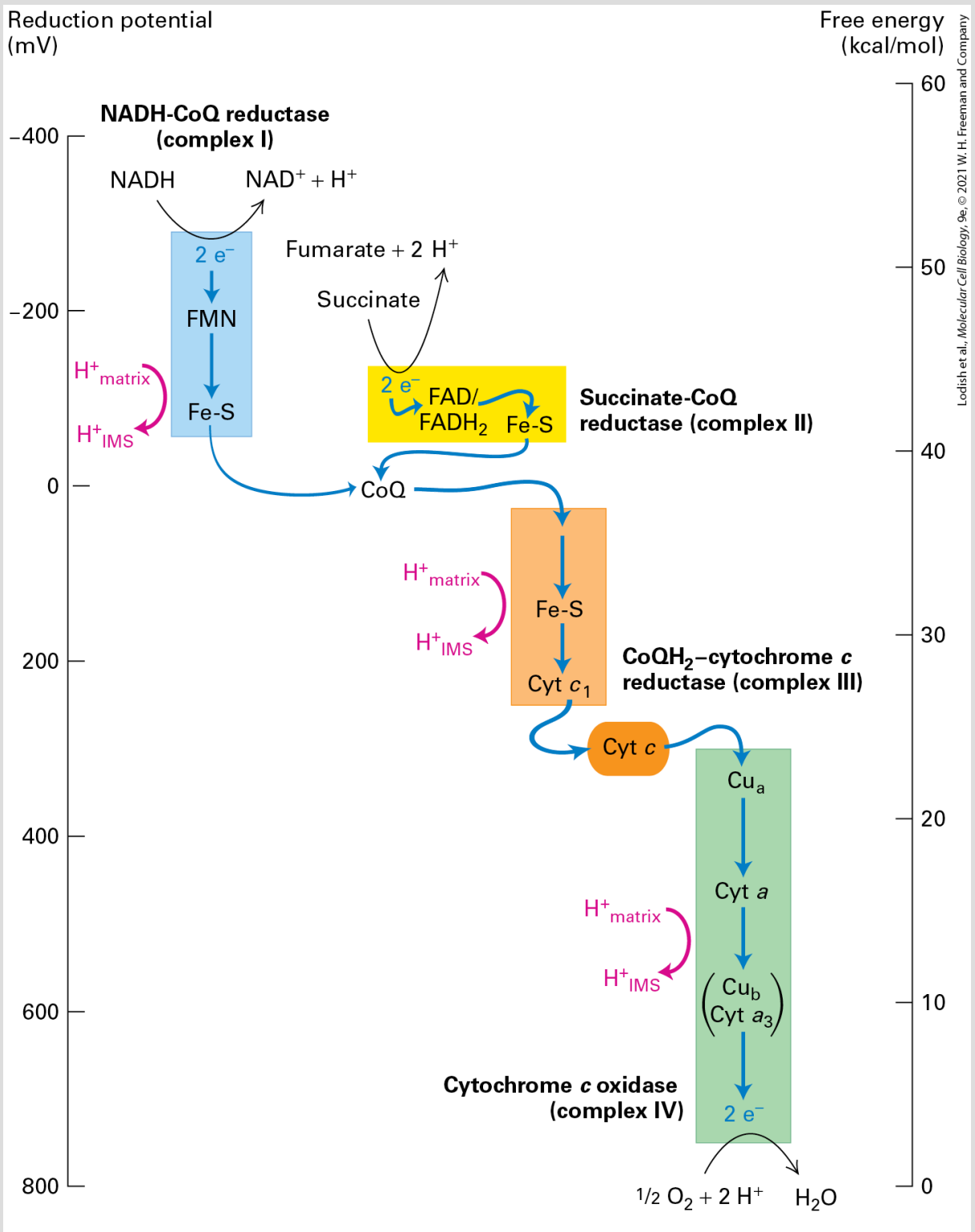


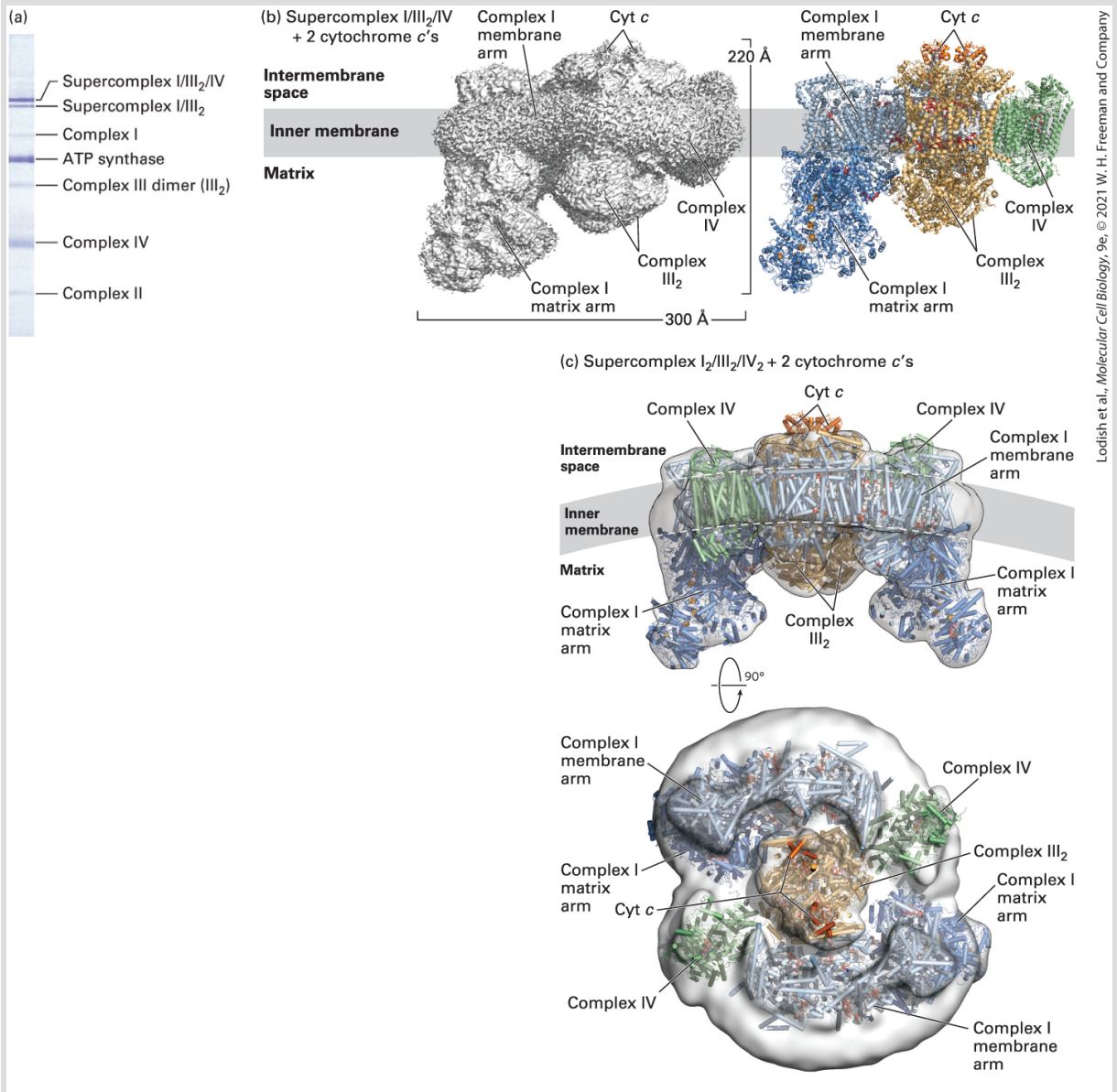
FIGURE 12-22 Changes in reduction potential and free energy during the stepwise flow of electrons through the electron-transport chain. Blue arrows indicate electron flow; red arrows, translocation of protons across the inner mitochondrial membrane. Electrons pass

through the multiprotein complexes from those with a lower reduction potential to those with a higher (more positive) reduction potential (*left* scale), with a corresponding reduction in free energy (*right* scale). The energy released as electrons flow through three of the complexes is sufficient to power the pumping of H^+ ions across the membrane from the matrix to the intermembrane space (IMS), establishing a proton-motive force.

The Multiprotein Complexes of the Electron-Transport Chain Assemble into Supercomplexes

Over 50 years ago, Britton Chance proposed that the electron-transport complexes might assemble into large **supercomplexes**. Doing so would bring the complexes into close and highly organized proximity, which might improve the speed and efficiency of the overall electron-transport process. Indeed, genetic, biochemical, and biophysical studies have provided very strong evidence for the existence of electron-transport chain supercomplexes. For example, polyacrylamide gel electrophoretic methods called *blue native (BN)-PAGE* and *colorless native (CN)-PAGE*, which permit separation of very large macromolecular protein complexes, have allowed detection of multiple inner mitochondrial membrane supercomplexes with various stoichiometries (Figure 12-23a). When one such supercomplex containing one copy each of complexes I and IV and two copies of complex III ($I/III_2/IV$) was isolated with ubiquinone (CoQ) and cytochrome *c* from BN-PAGE gels, it was shown to transfer electrons from NADH to O_2; in other words, this

supercomplex can respire — it is a respirasome. The structures of two of such supercomplexes from human mitochondria determined by cryo-EM are shown in Figure 12-23b ($I/III_2/IV + 2$ cytochrome *c*'s) and 12-23c ($I/III_2/IV + 2$ cytochrome *c*'s). The precise function of supercomplex formation in the context of the very high protein concentration in the inner mitochondrial membrane remains to be established with certainty, but is thought to involve improving the speed and efficiency of electron transport, stabilizing individual multiprotein complexes, or preventing inappropriate protein aggregates.

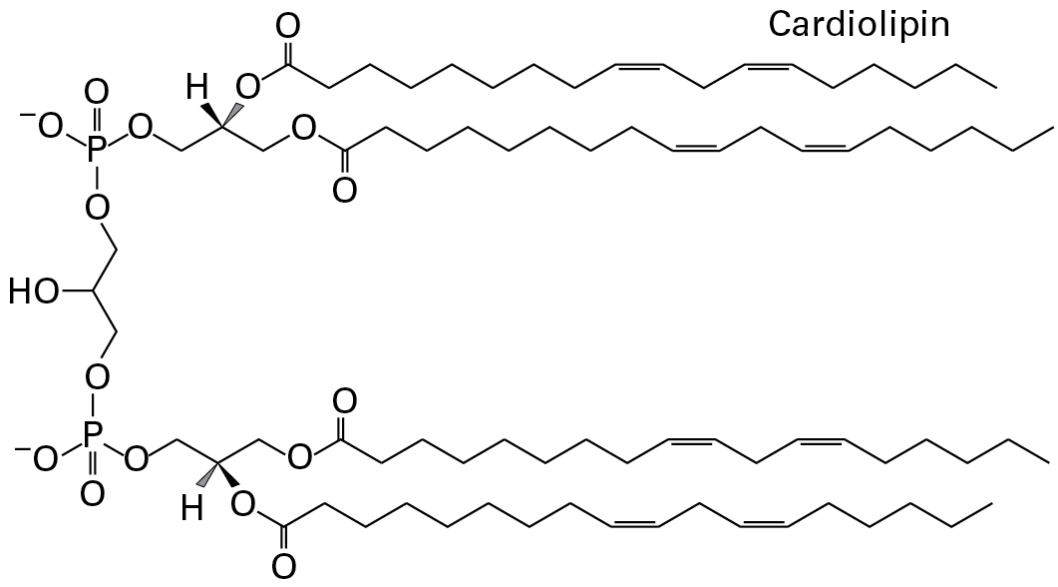


EXPERIMENTAL FIGURE 12-23 Electrophoresis and cryoelectron microscopic imaging have identified and permitted high-resolution structure determination of electron-transport chain supercomplexes containing complexes I, III, and IV. (a) Membrane proteins in isolated bovine heart mitochondria were solubilized with a detergent, and the complexes and supercomplexes were separated by gel electrophoresis (see Chapter 3) using the blue native (BN)-PAGE method. Each blue-stained band within the gel represents the indicated protein complex or supercomplex. The intensity of the blue stain is approximately proportional to the amount of complex or supercomplex present. The supercomplex $\mathrm{I/III_2/IV}$ has an apparent molecular weight of ~$1.7\,\mathrm{MDa}$. (b and c) Supercomplexes $\mathrm{I/III_2/IV} + 2$ cytochrome *c*'s and $\mathrm{I_2/III_2/IV_2} + 2$ cytochrome *c*'s were isolated by sucrose

density gradient centrifugation from detergent extracted mitochondria from human HEK293 cultured cells. Cryo-EM imaging methods were used to determine the structures at overall resolutions of $3.9\,\text{Å}$ ($\mathrm{I}/\mathrm{III}_2/\mathrm{IV} + 2$ cytochrome *c*'s) and $17.4\,\text{Å}$ ($\mathrm{I}_2/\mathrm{III}_2/\mathrm{IV}_2 + 2$ cytochrome *c*'s). (b) The left image shows the three-dimensional surface structure of supercomplex $\mathrm{I}/\mathrm{III}_2/\mathrm{IV} + 2$ cytochrome *c*'s viewed from an orientation parallel to the presumptive plane of the membrane. The right image is a molecular model showing the individual complexes: complex I (the peripheral "matrix" arm is dark blue and the membrane-embedded arm is light blue), dimer of complex III (gold), complex IV (green) and cytochrome c (Cyt c, orange/brown). Additional atoms in associated small molecules (such as the iron sulfur clusters in complex I and lipids) are colored as follows: carbon, white; iron, burnt orange; sulfur, yellow; oxygen, red; phosphorous, orange; and nitrogen, blue. The same color scheme is used in panel (c). The data permitted the structure of most of the supercomplex to be determined. For example, the locations of all 45 protein subunits in complex I and most of their 8499 amino acid residues were identified. (c) The top image shows the translucent electron density of supercomplex $\mathrm{I}_2/\mathrm{III}_2/\mathrm{IV}_2 + 2$ cytochrome *c*'s into which molecular models for each complex and Cyt c were docked. The bottom image was generated by rotating the top image 90° about a horizontal axis. The locations of the cytochrome *c*'s would permit them to rapidly shuttle electrons from complex III to complex IV (see Figure 12-19).

[Part (a) Republished with permission by *The American Society for Biochemistry and Molecular Biology*, from Schafer, E., et al., "Architecture of active mammalian respiratory chain supercomplexes," *J. Biol. Chem.* 2006, Jun 2;**281**(22):15370–15375. Epub 2006 Mar 20; permission conveyed through Copyright Clearance Center, Inc. Parts (b) and (c) Republished with permission of Elsevier, from Guo R, Zong S, Wu M, Gu J, Yang M., Architecture of Human Mitochondrial Respiratory Megacomplex I(2)III(2)IV(2). *Cell*. 2017 Sep 7;**170**(6):1247–1257; permission conveyed through Copyright Clearance Center, Inc.]

The unique phospholipid *cardiolipin* (diphosphatidyl glycerol) appears to play an important role in the assembly and function of these supercomplexes.



Generally found only in the inner mitochondrial membrane, cardiolipin, which normally composes about 20 percent of the lipids in the inner mitochondrial membrane, has been observed to bind to integral membrane proteins of the inner mitochondrial membrane (e.g., complexes I and II). Genetic and biochemical studies in yeast mutants in which cardiolipin synthesis is blocked have established that cardiolipin contributes to the formation and activity of mitochondrial supercomplexes; thus it has been called the glue that holds together the electron-transport chain, though the precise mechanism by which it does so remains to be defined. In addition, there is evidence that cardiolipin may influence the inner membrane's binding and permeability to protons and consequently the proton-motive force. When mitochondria are damaged, cardiolipin can move to the outer membrane, where it can participate in mitophagy. The role of cardiolipin can be seen in patients with Barth's syndrome, a human X-linked genetic disease

caused by defects in an enzyme that determines the structures of the acyl chains on cardiolipin. The reduction in the amounts of cardiolipin in these patients and its abnormal structure result in heart and skeletal muscle defects, growth retardation, and other abnormalities.

Reactive Oxygen Species Are By-Products of Electron Transport

Rather than being converted to water, about 1–2 percent of the oxygen metabolized by aerobic organisms is partially reduced to the superoxide anion radical ($O_2^{\bullet-}$, where the "dot" represents an unpaired electron). Radicals are atoms that have one or more unpaired electrons in an outer (valence) shell, or molecules that contain such an atom. Many, though not all, radicals are generally highly chemically reactive, altering the structures and properties of those molecules with which they react. The products of such reactions are often themselves radicals and can thus propagate a chain reaction that alters many additional molecules. Superoxide and other highly reactive oxygen-containing molecules, both radicals (e.g., $O_2^{\bullet-}$) and non-radicals (e.g., hydrogen peroxide, H_2O_2), are called **reactive oxygen species (ROS)**. ROS are of great interest because they can react with, and thus damage, many key biological molecules, including lipids (particularly unsaturated fatty acids and their derivatives), proteins, and DNA, and thus severely

interfere with their normal functions. At moderate to high levels, ROS contribute to what is often called *cellular oxidative stress* and can be highly toxic. Indeed, ROS are purposefully generated by body-defense cells (e.g., macrophages, neutrophils) to kill pathogens. In humans, excessive or inappropriate generation of ROS has been implicated in many diverse diseases, including heart failure, neurodegenerative diseases, alcohol-induced liver disease, and diabetes, as well as in aging. ■

Although there are several mechanisms for generating ROS in cells, their major source in eukaryotic cells is electron transport in the mitochondria (or in chloroplasts, as described below). Electrons passing through the mitochondrial electron-transport chain can have sufficient energy to reduce molecular oxygen (O_2) to form **superoxide anions** ($O_2^{\bullet -}$, Figure 12-24, *top*). This can occur, however, only when molecular oxygen comes in close contact with the reduced electron carriers (iron, FMN, $CoQH_2$) in the chain. Usually such contact is prevented by sequestration of the carriers within the proteins involved. However, there are some sites (particularly in complex I and $CoQ^{\bullet -}$, see Figure 12-18) and some conditions (e.g., high $NADH/NAD^+$ ratio in the matrix, high proton-motive force when ATP is not generated) when electrons can more readily leak out of the chain and reduce O_2 to $O_2^{\bullet -}$.

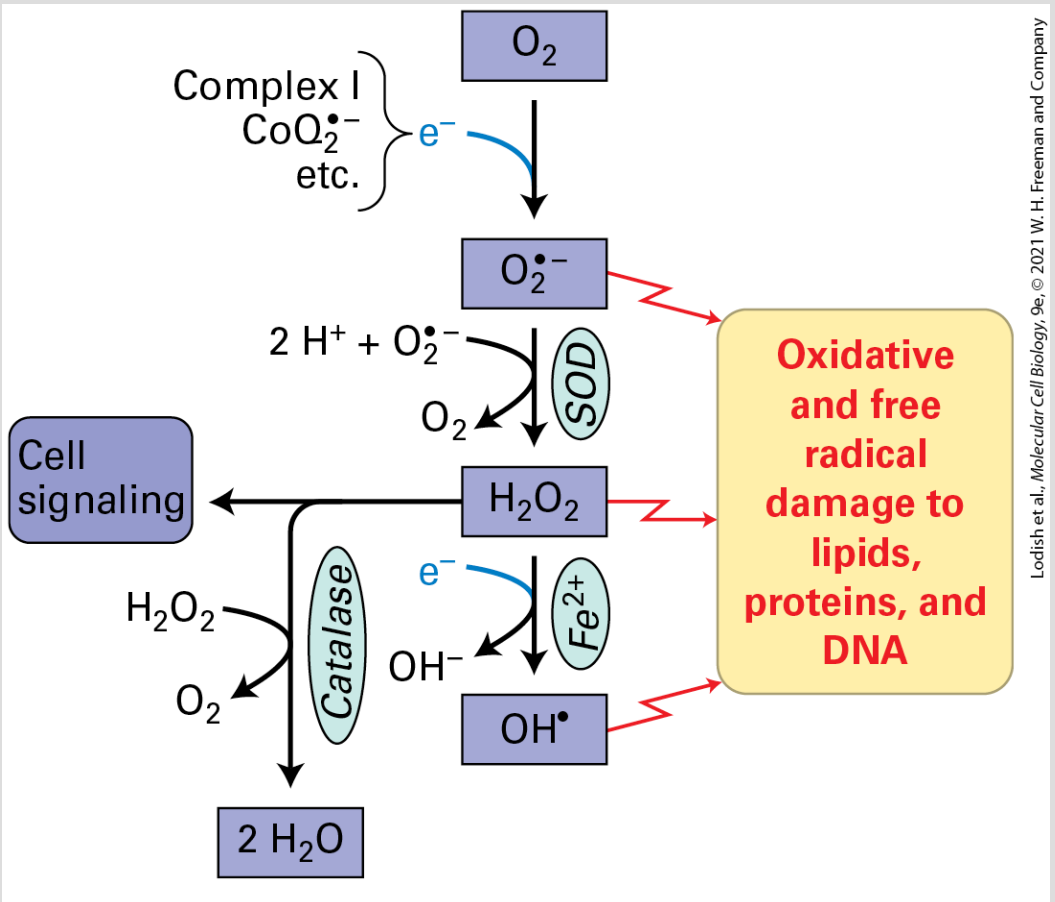


FIGURE 12-24 Generation and inactivation of toxic reactive oxygen species. Electrons from the electron-transport chains of mitochondria and chloroplasts, as well as some generated through other enzymatic reactions, reduce molecular oxygen (O_2), forming the highly reactive radical anion superoxide $(O_2{}^{\bullet -})$. Superoxide is rapidly converted by superoxide dismutase (SOD) to hydrogen peroxide (H_2O_2), which in turn can be converted by metal ions such as Fe^{2+} to hydroxyl radicals $(OH^\bullet)$ or inactivated to H_2O by enzymes such as catalase. Because of their high chemical reactivity, $O_2{}^{\bullet -}$, H_2O_2, $OH^\bullet$, and similar molecules are called *reactive oxygen species (ROS)*. They cause oxidative and free-radical damage to many biomolecules, including lipids, proteins, and DNA. This damage leads to cellular oxidative stress that can cause disease and, if sufficiently severe, can kill cells. In addition, ROS can function as intra- and intercellular signaling molecules.

The superoxide anion is an especially unstable, and thus reactive, ROS. Mitochondria have evolved several defense mechanisms that help protect against $O_2^{\bullet -}$ toxicity, including the use of enzymes that inactivate superoxide: first superoxide is converted to H_2O_2 by Mn-containing superoxide dismutase, called *SOD*, and then to H_2O by the enzyme catalase (see Figure 12-24). SOD and catalase are some of the fastest enzymes known so that they prevent the buildup of ROS. SOD is found within mitochondria and other cellular compartments. Hydrogen peroxide itself is an ROS that can diffuse readily across membranes and react with molecules throughout the cell. It can also be converted by certain metals, such as Fe^{2+}, into the even more dangerous hydroxyl radical ($OH^{\bullet}$). Thus cells depend on the inactivation of H_2O_2 by catalase and other enzymes, such as peroxiredoxin and glutathione peroxidase, which also detoxify the lipid hydroperoxide products formed when ROS react with unsaturated fatty acyl groups. Small-molecule antioxidant radical scavengers, such as vitamin E and α-lipoic acid, also protect against oxidative stress. Although in many cells catalase is located only in peroxisomes, in heart muscle cells it is found in mitochondria. This is not surprising because the heart is the most oxygen-consuming organ per gram in mammals.

As the rate of ROS production by mitochondria and chloroplasts reflects the metabolic state of these organelles (e.g., strength of proton-motive force, $NADH/NAD^+$ ratio), cells have developed ROS-sensing systems, such as ROS/redox-sensitive transcription

factors, to monitor the metabolic state of these organelles and respond accordingly — for example, by changing the rate of transcription of nuclear genes that encode organelle-specific proteins. Thus ROS such as H_2O_2 can function as a physiologically relevant intracellular, and even intercellular, signaling molecule. ROS have been reported to participate in cell processes as diverse as adaptation to low oxygen levels (hypoxia) and stress, growth factor and nutrient regulation of cell proliferation, cell differentiation, regulated cell death, and autophagy.

Experiments Using Purified Electron-Transport Chain Complexes Established the Stoichiometry of Proton Pumping

The multiprotein complexes of the electron-transport chain that are responsible for proton pumping have been identified by selectively extracting mitochondrial membranes with detergents, isolating each of the complexes in nearly pure form, and then preparing artificial phospholipid vesicles (liposomes) containing each complex. When an appropriate electron donor and electron acceptor are added to such liposomes, a change in the pH of the medium will occur if the embedded complex transports protons. Studies of this type as well as studies using membranes isolated from mitochondria indicate that NADH-CoQ reductase (complex I)

translocates four protons per pair of electrons transported, whereas cytochrome *c* oxidase (complex IV) translocates two protons per pair of electrons transported (see Figure 12-19).

Current evidence suggests that a total of 10 protons are transported from the matrix across the inner mitochondrial membrane for every electron pair that is transferred from NADH to O_2 (see Figure 12-19). Because succinate-CoQ reductase (complex II) does not transport protons, and because complex I is bypassed when the electrons come from succinate-derived $FADH_2$, only six protons are transported across the membrane for every electron pair that is transferred from this $FADH_2$ to O_2.

The Proton-Motive Force in Mitochondria Is Due Largely to a Voltage Gradient Across the Inner Membrane

The main result of the electron-transport chain is the generation of the proton-motive force, which is the sum of a transmembrane proton concentration (pH) gradient and an electric potential, or voltage gradient, across the inner mitochondrial membrane. The relative contributions of these two components to the total proton-motive force have been shown to depend on the permeability of the

membrane to ions other than H^+. A significant voltage gradient can develop only if the membrane is poorly permeable to other cations and to anions. Otherwise, anions would leak across the membrane from the matrix to the intermembrane space along with the protons and prevent a voltage gradient from forming. Similarly, if cations other than H^+ could leak across the membrane in a direction opposite to that of the H^+ (from the intermembrane space to the matrix), that leakage would counterbalance the positive charges delivered to the intermembrane space by the protons, short-circuiting voltage-gradient formation. Indeed, the inner mitochondrial membrane is poorly permeable to ions other than H^+. Thus proton pumping generates a voltage gradient that makes it energetically difficult for additional protons to move across the membrane because of charge repulsion. As a consequence, proton pumping by the electron-transport chain establishes a robust voltage gradient in the context of what turns out to be a rather small pH gradient.

Because mitochondria are much too small to be impaled with electrodes, the electric potential and pH gradient across the inner mitochondrial membrane cannot be directly measured. However, the electric potential can be measured indirectly by adding radioactive $^{42}K^+$ ions and a trace amount of valinomycin to a suspension of respiring mitochondria and measuring the amount of radioactivity that accumulates in the matrix. Although the inner membrane is normally impermeable to K^+, valinomycin is an

ionophore, a small lipid-soluble molecule that selectively binds a specific ion (in this case, K^+) and carries it across otherwise impermeable membranes. In the presence of valinomycin, $^{42}K^+$ equilibrates across the inner membrane of isolated mitochondria in accordance with the electric potential: the more negative the matrix side of the membrane, the more $^{42}K^+$ will be attracted to and accumulate in the matrix.

At equilibrium, the measured concentration of radioactive K^+ ions in the matrix, $[K^+{}_{\text{matrix}}]$, is about 500 times greater than that in the surrounding medium, $[K^+{}_{\text{out}}]$. Substitution of this value into the Nernst equation (see Chapter 11) shows that the electric potential E (in mV) across the inner membrane in respiring mitochondria is $-160\,\text{mV}$, with the matrix (inside) negative:

$$E = -59\ \log \frac{[K^+{}_{\text{matrix}}]}{[K^+{}_{\text{out}}]} = -59\ \log\ 500 = -160\,\text{mV}$$

Alternatively, the relative mitochondrial inner membrane potential can be measured using lipophilic cationic dyes such as tetramethylrhodamine methyl ester. The extent of accumulation of these fluorescent dyes in the mitochondrial matrix depends on the membrane potential and follows the Nernst equation. Various methods can be used to measure the fluorescence in the mitochondria, including flow cytometry, spectrofluorimetry, and quantitative fluorescence microscopy.

Researchers can measure the matrix (inside) pH by trapping pH-sensitive fluorescent dyes inside vesicles formed from the inner mitochondrial membrane, with the matrix side of the membrane facing inward. They can also measure the pH outside the vesicles (equivalent to the intermembrane space) and thus determine the pH gradient ($\Delta\mathrm{pH}$), which turns out to be about one pH unit. A difference of one pH unit represents a tenfold difference in H^+ concentration, so according to the Nernst equation, a pH gradient of one unit across a membrane is equivalent to an electric potential of 59 mV at 20 °C. Thus knowing the voltage and pH gradients, we can calculate the proton-motive force (pmf) as

$$\mathrm{pmf} = \Psi - \left(\frac{RT}{F} \times \Delta\mathrm{pH}\right) = \Psi - 59\,\Delta\mathrm{pH}$$

where R is the gas constant of $1.987\,\mathrm{cal/(degree \cdot mol)}$, T is the temperature (in degrees Kelvin), F is the Faraday constant $[23{,}062\,\mathrm{cal/(V \cdot mol)}]$, and Ψ is the transmembrane electric potential; Ψ and pmf are measured in millivolts. The electric potential Ψ across the inner membrane is $-160\,\mathrm{mV}$ (negative inside), and $\Delta\mathrm{pH}$ is equivalent to about 60 mV. Thus the total proton-motive force is $-220\,\mathrm{mV}$, with the transmembrane electric potential responsible for about 73 percent of the total.

KEY CONCEPTS OF SECTION 12.6

The Electron-Transport Chain and Generation of the Proton-Motive Force

- By the end of the citric acid cycle (stage II), much of the energy originally present in the covalent bonds of glucose and fatty acids has been converted into high-energy electrons in the reduced coenzymes NADH and $\mathrm{FADH_2}$. The energy from these electrons is used to generate the proton-motive force.
- In the mitochondrion, the proton-motive force is generated by coupling electron flow (from NADH and $\mathrm{FADH_2}$ to $\mathrm{O_2}$) to the energetically uphill transport of protons from the matrix across the inner membrane to the intermembrane space. This process, together with the synthesis of ATP from ADP and $\mathrm{P_i}$ driven by the proton-motive force, is called *oxidative phosphorylation*.
- As electrons flow from $\mathrm{FADH_2}$ and NADH to $\mathrm{O_2}$, they pass through multiprotein complexes. The four major complexes are NADH-CoQ reductase (complex I), succinate-CoQ reductase (complex II), $\mathrm{CoQH_2}$–cytochrome *c* reductase (complex III), and cytochrome *c* oxidase (complex IV) (see Figure 12-19).
- Each complex contains one or more electron-carrying prosthetic groups, which include iron-sulfur clusters, flavins, heme groups, and copper ions (see Table 12-3). Cytochrome *c*, which contains heme, and coenzyme Q (CoQ), a lipid-soluble small molecule, are mobile carriers that shuttle electrons between the complexes.
- Complexes I, III, and IV pump protons from the matrix into the intermembrane space. Complexes I and II reduce CoQ to $\mathrm{CoQH_2}$, which carries protons and high-energy electrons to complex III. The heme protein cytochrome *c* carries electrons from complex III to complex IV, which uses them to pump protons and reduce molecular oxygen to water.
- The high-energy electrons from NADH enter the electron-transport chain through complex I, whereas the high-energy electrons from $\mathrm{FADH_2}$ (derived from succinate in the citric acid cycle) enter the electron-transport chain through complex II. Additional electrons derived from $\mathrm{FADH_2}$ by the initial step of fatty acyl–CoA β-oxidation increase the supply of $\mathrm{CoQH_2}$ available for electron transport.
- The Q cycle allows four protons to be translocated per pair of electrons moving through complex III (see Figure 12-21).
- Each electron carrier in the chain accepts an electron or electron pair from a carrier with a less positive reduction potential and transfers the electron to a carrier with a more positive reduction potential. Thus the reduction potentials of electron carriers favor unidirectional, downhill, electron flow from NADH and $\mathrm{FADH_2}$ to $\mathrm{O_2}$ (see Figure 12-22).

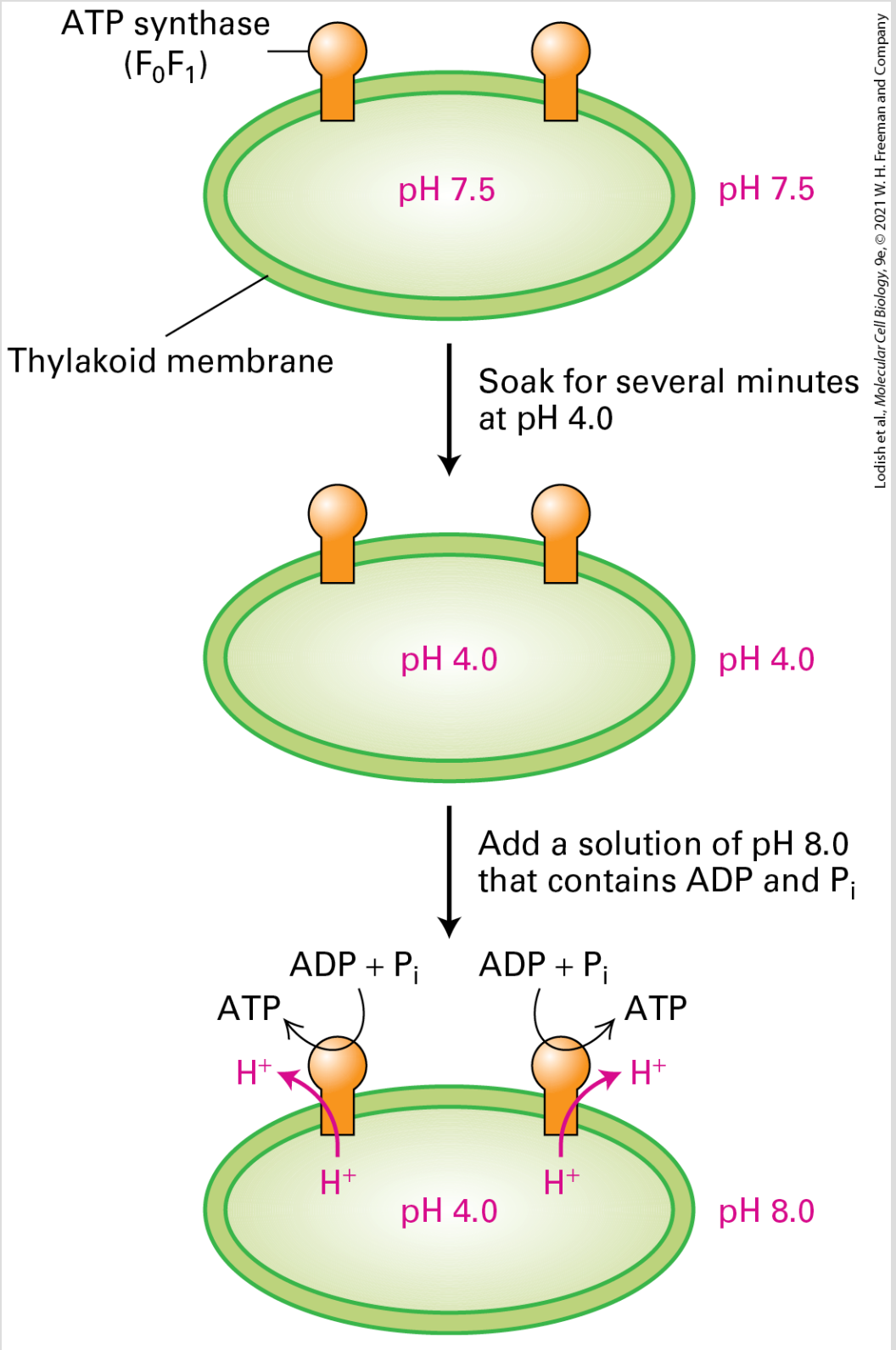
- Within the inner mitochondrial membrane, electron-transport complexes assemble into supercomplexes held together by cardiolipin, a specialized phospholipid. Supercomplex formation may enhance the speed and efficiency of generation of the proton-motive force or play other roles.
- Reactive oxygen species (ROS) are toxic by-products of the electron-transport chain that can modify and damage proteins, DNA, and lipids. Specific enzymes (e.g., glutathione peroxidase, catalase) and small-molecule antioxidants (e.g., vitamin E) help protect against ROS-induced damage (see Figure 12-24). ROS can also be used as intracellular signaling molecules.
- A total of $10\,\mathrm{H}^+$ ions are translocated from the matrix across the inner membrane per electron pair flowing from NADH to O_2 (see Figure 12-19), whereas $6\,\mathrm{H}^+$ ions are translocated per electron pair flowing from FADH_2 to O_2.
- The proton-motive force is largely due to a voltage gradient across the inner membrane produced by proton pumping; the pH gradient plays a quantitatively less important role.

12.7 Harnessing the Proton-Motive Force to Synthesize ATP

The hypothesis that a proton-motive force across the inner mitochondrial membrane is the immediate source of energy for ATP synthesis was proposed in 1961 by Peter Mitchell. Virtually all researchers studying oxidative phosphorylation and photosynthesis initially rejected his proposal (called the *chemiosmotic hypothesis*). They favored a mechanism similar to the then well-elucidated substrate-level phosphorylation in glycolysis, in which chemical transformation of a substrate molecule (like phosphoenolpyruvate in glycolysis) is directly coupled to ATP synthesis. Despite intense efforts by a large number of investigators, however, compelling evidence for such a substrate-level phosphorylation–mediated mechanism was never observed.

Definitive evidence supporting Mitchell's hypothesis depended on developing and employing techniques to purify and reconstitute organelle membranes and membrane proteins. In one such experiment, outlined in Figure 12-25, vesicles made from chloroplast thylakoid membranes (equivalent to the inner membranes of mitochondria, we will discuss these later in this chapter) that contain the enzyme ATP synthase, were studied. This and other experiments demonstrated that ATP synthase is an ATP-generating enzyme and that ATP generation is dependent on proton movement down an electrochemical gradient across the membrane

in which the enzyme resided. It turns out that the protons actually move *through* the ATP synthase protein as they traverse the membrane.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 12-25 Synthesis of ATP by ATP synthase depends on a pH gradient across the membrane. Isolated chloroplast thylakoid vesicles

containing ATP synthase (F_0F_1 complexes) were equilibrated with a buffered solution at pH 4.0. When the pH in the thylakoid lumen reached 4.0, the vesicles were rapidly mixed with a solution at pH 8.0 containing ADP and P_i. A burst of ATP synthesis accompanied the transmembrane movement of protons driven by the 10,000-fold H^+ concentration gradient (10^{-4} M versus 10^{-8} M). In similar experiments using "inside-out" preparations of mitochondrial membrane vesicles, an artificially generated membrane electric potential, also resulted in ATP synthesis.

As we shall see, ATP synthase is a multiprotein complex that can be subdivided into two subcomplexes, called F_0 (containing the transmembrane portions of the complex) and F_1 (containing the globular portions of the complex that sit above the membrane and point into the matrix in mitochondria). Thus ATP synthase is often called the **F_0F_1 complex**; we will use the two terms interchangeably.

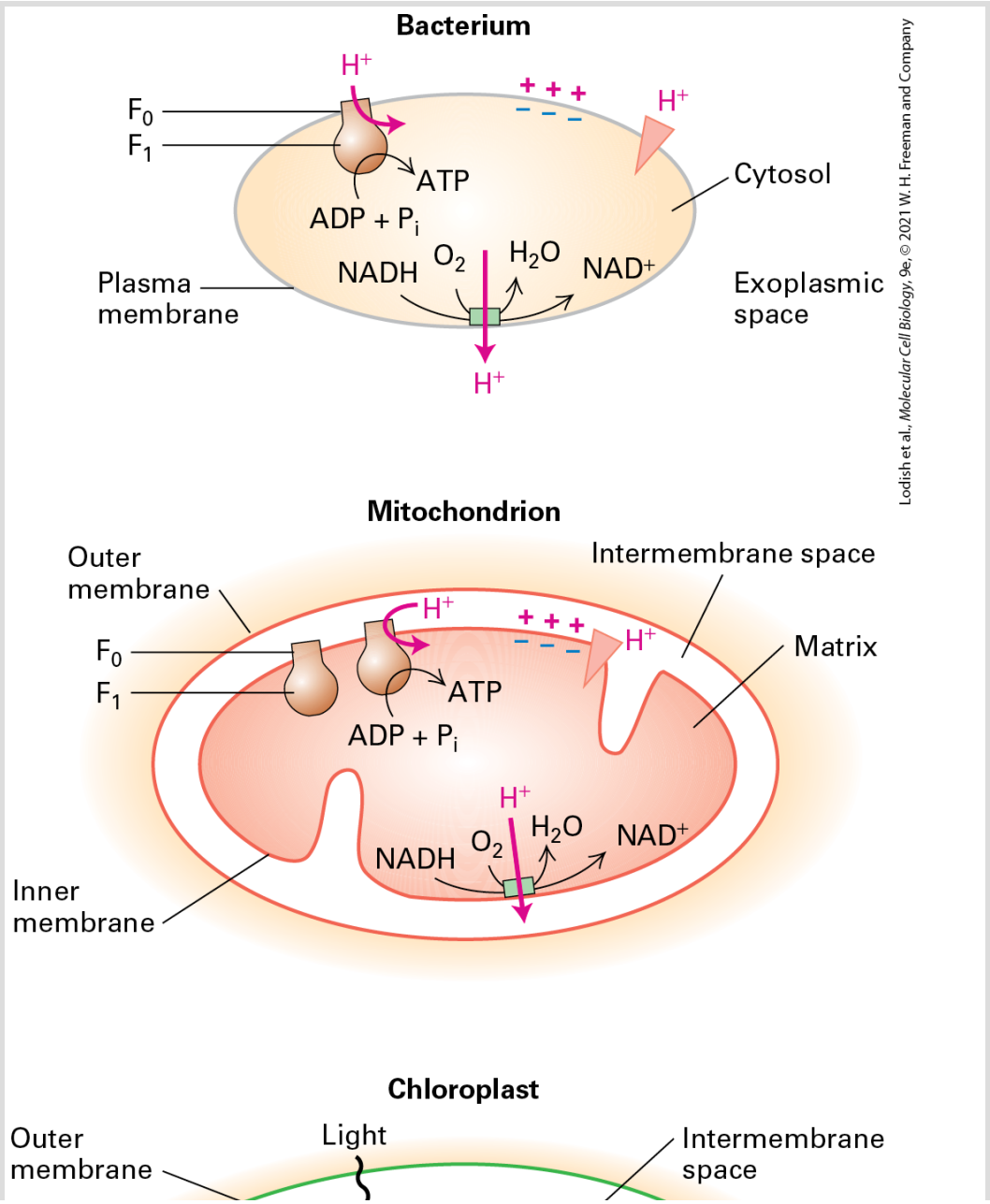
The Mechanism of ATP Synthesis Is Shared Among Bacteria, Mitochondria, and Chloroplasts

Eukaryotes use specialized membranes in intracellular organelles (mitochondria and chloroplasts) to make ATP. Of course, bacteria also need to make ATP, but they lack mitochondria or chloroplasts. Nevertheless, aerobic bacteria carry out oxidative phosphorylation by the same processes that occur in eukaryotic mitochondria and chloroplasts (Figure 12-26). Enzymes that catalyze the reactions of

both the glycolytic pathway and the citric acid cycle are present in the cytosol of bacteria; enzymes that oxidize NADH to NAD^+ and transfer the electrons to the ultimate acceptor O_2 reside in the bacterial plasma membrane. The movement of electrons through these membrane carriers is coupled to the pumping of protons out of the cell. The movement of protons back into the cell, down their concentration gradient through ATP synthase, releases energy that drives the synthesis of ATP. The bacterial ATP synthase (F_0F_1 complex) is essentially identical in structure and function to the mitochondrial and chloroplast ATP synthases, but is simpler to purify and study.

Why is the mechanism of ATP synthesis shared among both prokaryotic organisms and eukaryotic organelles? Primitive aerobic bacteria were probably the progenitors of both mitochondria and chloroplasts in eukaryotic cells (see Figure 12-7). According to this **endosymbiont hypothesis**, the inner mitochondrial membrane was derived from the bacterial plasma membrane, with its cytosolic face pointing toward what became the matrix of the mitochondrion. Similarly, in plants, a progenitor bacterium's plasma membrane became the chloroplast's thylakoid membrane, and its cytosolic face pointed toward what became the stromal space of the chloroplast (chloroplast structure and function will be described in Sections 12.7 and 12.8). In all cases, ATP synthase is positioned with the globular F_1 domain, which catalyzes ATP synthesis, on the cytosolic face of the membrane, so ATP is always formed on the cytosolic face (see

Figure 12-26). Protons always flow through ATP synthase from the membrane's exoplasmic face (facing the intermembrane space in mitochondria) to its cytosolic face (facing the matrix in mitochondria). This flow is driven by the proton-motive force. Invariably, the cytosolic face has a negative electric potential relative to the exoplasmic face.



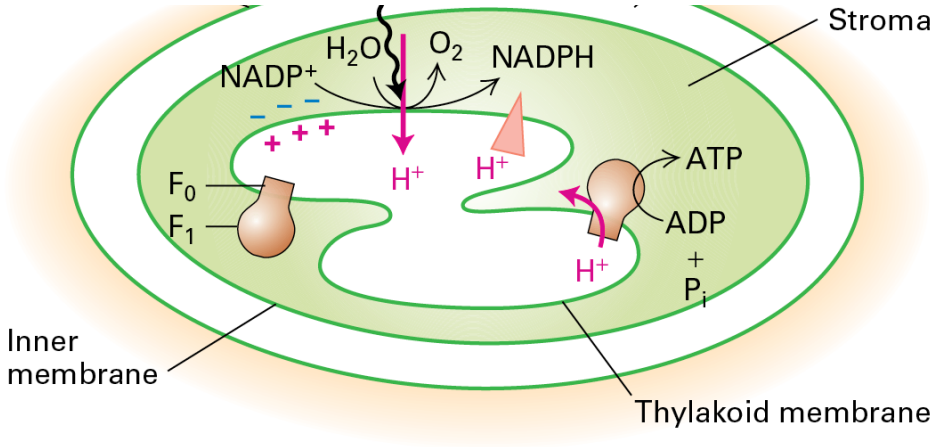


FIGURE 12-26 ATP synthesis by chemiosmosis is similar in bacteria, mitochondria, and chloroplasts. In chemiosmosis, a proton-motive force generated by proton pumping across a membrane is used to power ATP synthesis. The mechanism and membrane orientation of the process are similar in bacteria, mitochondria, and chloroplasts. In each illustration, the membrane surface facing a shaded area is a cytosolic face; the surface facing an unshaded, white area is an exoplasmic face. Note that the cytosolic face of the bacterial plasma membrane, the matrix face of the inner mitochondrial membrane, and the stromal face of the thylakoid membrane are all equivalent. During electron transport, protons are always pumped from the cytosolic face to the exoplasmic face, creating a proton concentration gradient (exoplasmic face > cytosolic face) and an electric potential (negative cytosolic face and positive exoplasmic face) across the membrane. During the synthesis of ATP, protons flow in the reverse direction (down their electrochemical gradient) through an ATP synthase (F_0F_1 complex), which protrudes in a knob at the cytosolic face in all cases.

In addition to ATP synthesis, the proton-motive force across the bacterial plasma membrane is used to power other processes, including the uptake of nutrients such as sugars (using proton/sugar symporters) and the rotation of bacterial flagella. Chemiosmotic coupling thus illustrates an important principle introduced in our

discussion of active transport in Chapter 11: the membrane potential, the concentration gradients of protons (and other ions) across a membrane, and the phosphoanhydride bonds in ATP are equivalent and interconvertible forms of potential energy (also see Chapter 2). Indeed, ATP synthesis through ATP synthase can be thought of as active transport in reverse.

ATP Synthase Comprises $\mathbf{F_0}$ and $\mathbf{F_1}$ Multiprotein Complexes

With general acceptance of Mitchell's chemiosmotic mechanism, researchers turned their attention to the structure and operation of ATP synthase. The complex has two principal subcomplexes, F_0 and F_1, both of which are multimeric proteins (Figure 12-27a). The F_0 subcomplex contains three types of integral membrane proteins, designated **a, b**, and **c**. In bacteria and in yeast mitochondria, the most common subunit stoichiometry is $\mathbf{a}_1\mathbf{b}_2\mathbf{c}_{10}$, but the number of **c** subunits per subcomplex varies among different eukaryotes from 8 to 15. In all cases, the **c** subunits form a doughnut-shaped ring ("**c** ring") in the plane of the membrane. The one **a** and two **b** subunits are rigidly linked to one another, but not to the **c** ring, a critical feature of the protein to which we will return shortly. As you can see in Figure 12-27a, when the ATP synthase is operating, portions of the structure do not move (orange) and the rest (green) rotate together about an axis perpendicular to the plane of the membrane.

As we shall see, the flow of protons through the ATP synthase powers the rotation of the complex of the c, γ, and ε subunits that in turn drives the synthesis of ATP from ADP and P_i on the β subunits.

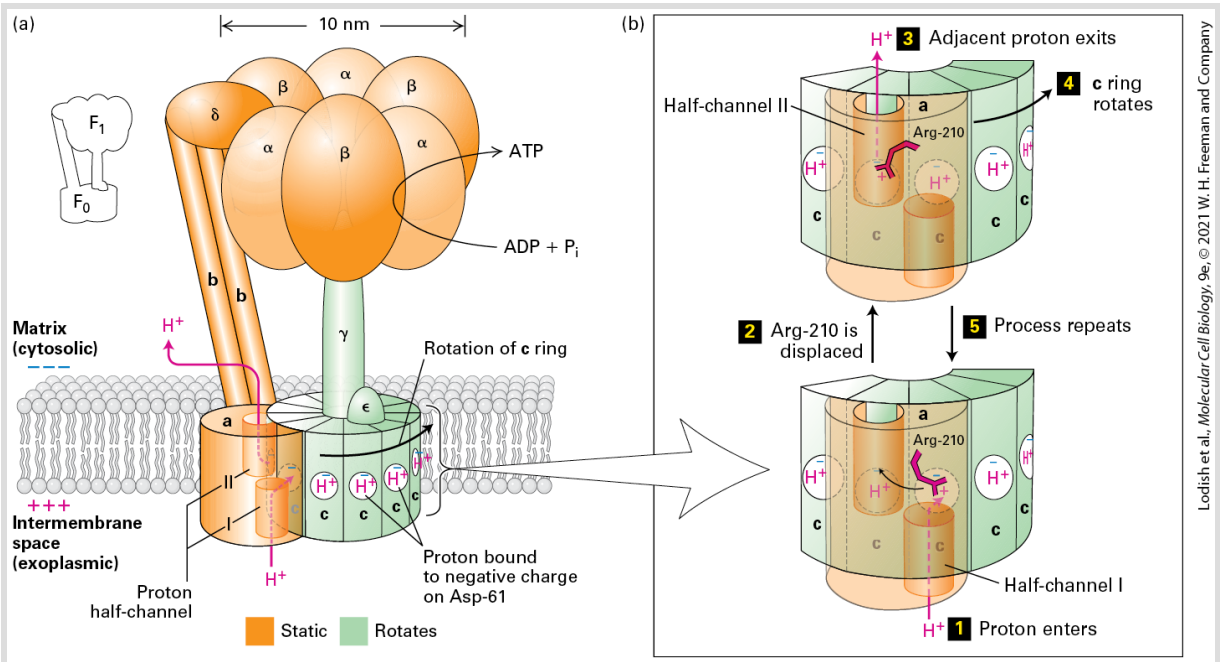


FIGURE 12-27 Structure of ATP synthase (the F_0F_1 complex) in the bacterial plasma membrane and mechanism of proton translocation across the membrane. (a) The F_0 membrane-embedded subcomplex of ATP synthase is built of three integral membrane proteins: one copy of **a**, two copies of **b**, and an average of ten copies of **c** arranged in a ring in the plane of the membrane. Two proton half-channels near the interfaces of subunit **a** with the **c** subunits mediate proton movement across the membrane (proton path is indicated by red arrows). Half-channel I allows protons to move one at a time from the exoplasmic medium (equivalent to intermembrane space in mitochondria) to the negatively charged side chain of Asp-61 in the center of a **c** subunit near the middle of the membrane. The proton-binding site in each **c** subunit is represented as a white circle with a blue "—" representing the negative charge on the side chain of Asp-61. Half-channel II permits protons to move from the Asp-61 of an adjacent **c** subunit into the cytosolic medium. In order for protons to move through the complex the c ring must rotate in the plane of the membrane, as shown by the detailed molecular structure of the **c** ring and a portion of the adjacent **a** subunit in Figure 12-30. The F_1 subcomplex of ATP synthase contains three copies each of subunits α and β, which form a hexamer resting atop the single rod-shaped γ

subunit, which is inserted into the **c** ring of F_0. The ε subunit is rigidly attached to the γ subunit and also to several of the **c** subunits. The δ subunit permanently links one of the α subunits in the F_1 subcomplex to the **b** subunit of F_0. Thus the F_0 **a** and **b** subunits and the F_1 δ subunit and $(\alpha\beta)_3$ hexamer form a rigid structure (orange) anchored in the membrane. During proton flow, the **c** ring and the attached F_1 ε and γ subunits rotate as a unit (green), causing conformational changes in the F_1 β subunits, leading to ATP synthesis. (b) Potential mechanism of proton translocation. Step **1**: A proton from the exoplasmic space enters half-channel I and moves toward the "empty" (unprotonated) Asp-61 proton-binding site. The negative charge (blue "−") on the unprotonated side chain Asp-61 is balanced, in part, by a positive charge on the side chain of Arg-210 (red "+"). Step **2**: The proton fills the empty proton-binding site and simultaneously displaces the positively charged Arg-210 side chain, which swings over to the filled proton-binding site on the adjacent **c** subunit (curved arrow). As a consequence, the proton bound at that adjacent site is displaced. Step **3**: The displaced adjacent proton moves through half-channel II and is released into the cytosolic space, leaving an empty proton-binding site on Asp-61. Step **4**: Counterclockwise rotation of the entire **c** ring moves the "empty" **c** subunit over half-channel I. Step **5**: The process is repeated. See M. J. Schnitzer, 2001, *Nature* **410**:878–881; P. D. Boyer, 1999, *Nature* **402**:247–249; and C. von Ballmoos, A. Wiedenmann, and P. Dimroth, 2009, *Annu. Rev. Biochem.* **78**:649–672.

The F_1 subcomplex is a water-soluble complex of five distinct polypeptides with the composition $\alpha_3\beta_3\gamma\delta\varepsilon$ that is normally firmly bound to the **a** subunit of the F_0 subcomplex at the surface of the membrane. The lower end of the rodlike γ subunit of the F_1 subcomplex is a coiled coil that fits into the center of the **c**-subunit ring of F_0 and appears rigidly attached to it. Thus when the **c**-subunit ring rotates, the rodlike γ subunit rotates with it. The F_1 ε subunit is rigidly attached to γ and also forms tight contacts with several of the **c** subunits of F_0. The α and β subunits are responsible for the overall globular shape of the F_1 subcomplex and associate in

alternating order to form a hexamer, αβαβαβ, or $(\alpha\beta)_3$, which rests atop, but is not rigidly attached to the single long γ subunit. The F_1 δ subunit is permanently linked to one of the F_1 α subunits and also binds to the **b** subunit of F_0. Thus the **a** and **b** subunits of the F_0 subcomplex and the δ subunit and $(\alpha\beta)_3$ hexamer of the F_1 subcomplex form a rigid structure anchored in the membrane. The rodlike **b** subunits form a "stator" that prevents the $(\alpha\beta)_3$ hexamer from moving while it rests on the γ subunit, whose rotation, together with that of the ε subunit and the **c**-subunit ring of F_0, plays an essential role in the ATP synthesis mechanism described below.

When ATP synthase is embedded in a membrane, the F_1 subcomplex forms a knob that protrudes from the cytosolic face (the matrix face in the mitochondrion). Because F_1 separated from membranes is capable of catalyzing ATP hydrolysis (ATP conversion to ADP plus P_i) in the absence of the F_0 subcomplex, it has been called the *F_1 ATPase*; however, its function in cells is the reverse, to synthesize ATP. ATP hydrolysis is a spontaneous process ($\Delta G < 0$); thus energy is required to drive the ATPase in reverse and generate ATP.

Rotation of the $F_1\gamma$ Subunit, Driven by Proton Movement Through F_0, Powers ATP Synthesis

Each of the three β subunits in the globular F_1 subcomplex of F_0F_1 can bind ADP and P_i and catalyze the endergonic synthesis of ATP when coupled to the flow of protons from the exoplasmic medium (the intermembrane space in the mitochondrion) to the cytosolic (matrix) medium. However, the energetic coupling of proton flow and ATP synthesis does not take place in the same portions of the protein, because the nucleotide-binding sites on the β subunits of F_1, where ATP synthesis occurs, are 9–10 nm from the surface of the membrane-embedded portion of F_0 through which the protons flow. The most widely accepted model for ATP synthesis by the F_0F_1 complex — the **binding-change mechanism** — posits an indirect, mechanical coupling (Figure 12-28).

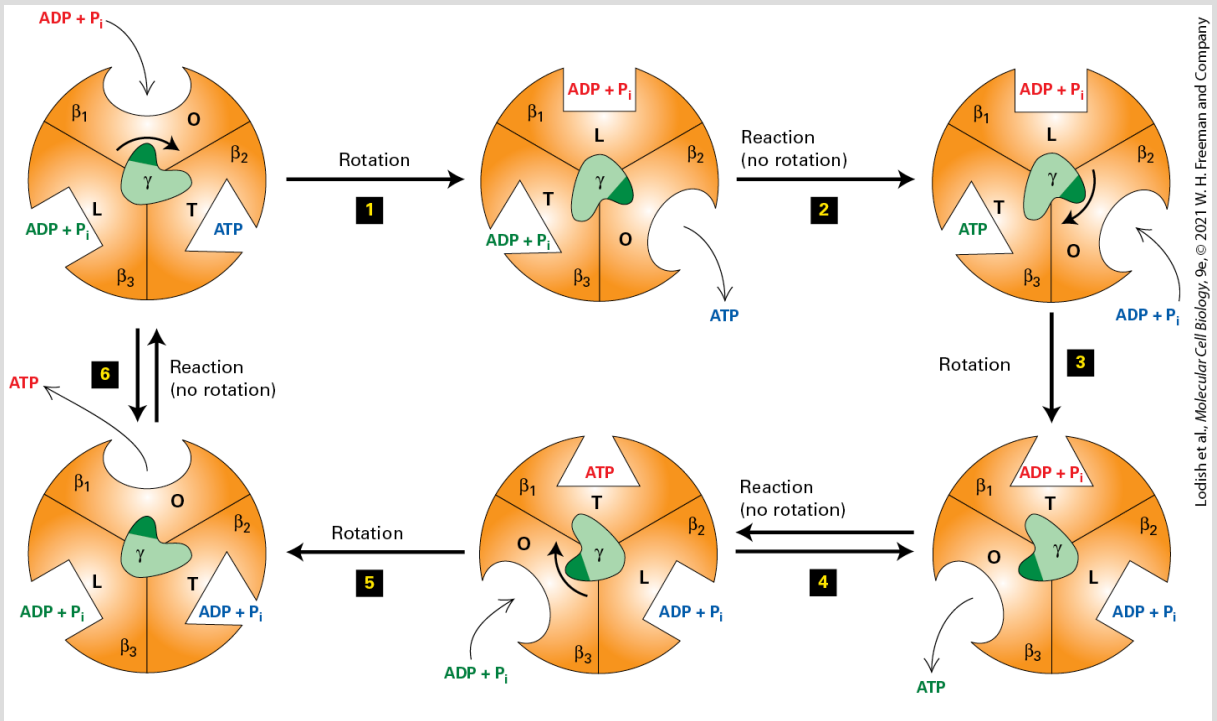


FIGURE 12-28 The binding-change mechanism of ATP synthesis from ADP and P_i. This view is looking up at F_1 from the membrane surface (see Figure 12-27). As the γ subunit

rotates by 120° in the center, each of the otherwise identical β subunits alternates among three conformational states (O, open, with oval representation of the nucleotide binding site; L, loose, with a rectangular binding site; T, tight, with a triangular site) that differ in their binding affinities for ATP, ADP, and P_i. To make it easier for you to follow the reactions, the three different sets of $ADP + P_i$ (and their product ATP) are colored red, blue, or green. The cycle begins (*upper left*) when ADP and P_i bind loosely to one of the three β subunits (here, arbitrarily designated β_1) whose nucleotide-binding site is in the O (open) conformation. Proton flux through the F_0 portion of the protein powers a 120° rotation of the γ subunit (relative to the fixed β subunits) (step **1**). This causes the rotating γ subunit, which is asymmetric, to push against each of the three β subunits in different ways, resulting in distinct conformational changes in each β subunit. In step **1** this results in an increase in the binding affinity of the β_1 subunit for ADP and $P_i(O \rightarrow L)$, an increase in the binding affinity of the β_3 subunit for ADP and P_i that were previously bound $(L \rightarrow T)$, and a decrease in the binding affinity of the β_2 subunit for a previously bound ATP $(T \rightarrow O)$, causing release of the bound ATP. Step **2**: Without additional rotation, the ADP and P_i in the T site (here, in the β_3 subunit) form ATP, a reaction that does not require an input of additional energy due to the special environment in the active site of the T state. At the same time, a new ADP and P_i bind loosely to the unoccupied O site on β_2. Step **3**: Proton flux powers another 120° rotation of the γ subunit, consequent conformational changes in the binding sites $(L \rightarrow T, O \rightarrow L, T \rightarrow O)$, and release of ATP from β_3. Step **4**: Without additional rotation, the ADP and P_i in the T site of β_1 form ATP, and additional ADP and P_i bind to the unoccupied O site on β_3. The process continues with rotation (step **5**) and ATP formation (step **6**) until the cycle is complete, with three ATPs having been produced for every 360° rotation of γ. See P. Boyer, 1989, *FASEB J.* **3**:2164–2178; Y. Zhou et al., 1997, *Proc. Natl. Acad. Sci. USA* **94**:10583–10587; and M. Yoshida, E. Muneyuki, and T. Hisabori, 2001, *Nat. Rev. Mol. Cell Biol.* **2**:669–677.

According to this mechanism, energy released by the downhill movement of protons through F_0 directly powers rotation of the **c**-subunit ring together with its rigidly attached γ and ε subunits (see Figure 12-27a). The γ subunit is a nonsymmetrical, rotating shaft, similar to a camshaft in a car engine, whose **c** ring–driven rotation

within the center of the static $(\alpha\beta)_3$ hexamer of F_1 causes it to push sequentially against each of the β subunits. The mechanical force exerted on the β subunits induces cyclical changes in their conformations and drives the conversion of ADP and P_i to ATP in the active site of each β subunit. As schematically depicted in a view of the bottom of the $(\alpha\beta)_3$ hexamer's globular structure in Figure 12-28, rotation of the γ subunit relative to the fixed $(\alpha\beta)_3$ hexamer causes the nucleotide-binding, active site of each β subunit to cycle through three conformational states in the following order:

1. An O (open) state that binds ATP very poorly and ADP and P_i weakly
2. An L (loose) state that binds ADP and P_i more strongly but cannot bind ATP
3. A T (tight) state that binds ADP and P_i so tightly that they spontaneously react and form ATP

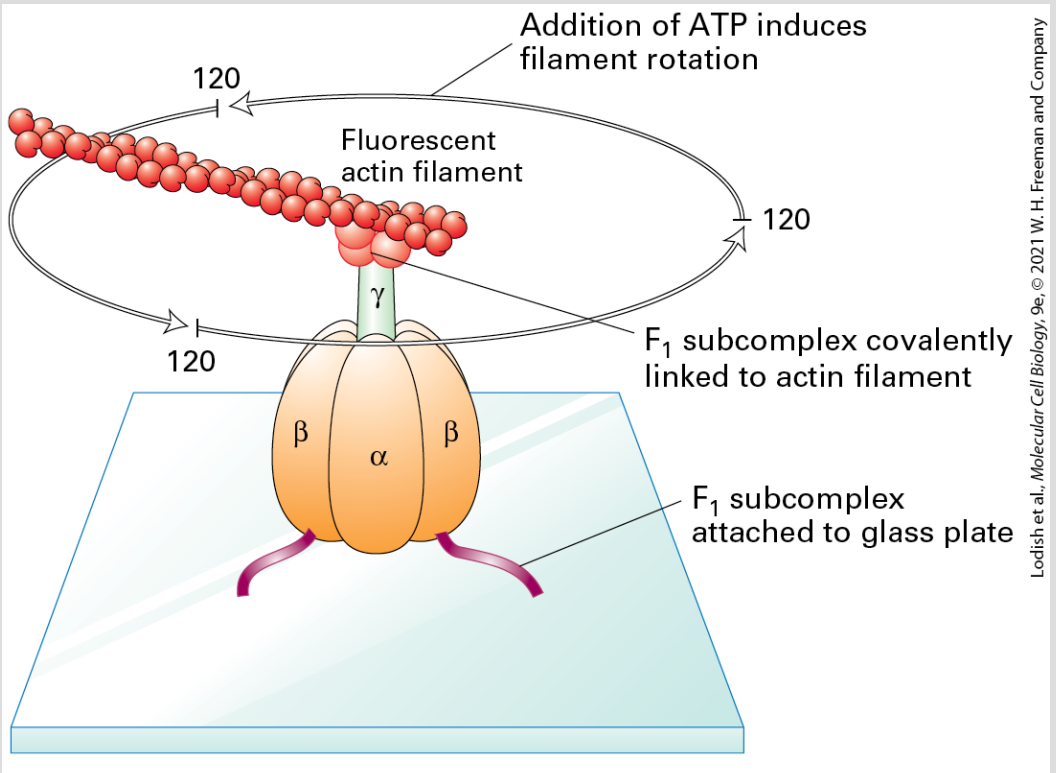
In the T state, the ATP produced is bound so tightly that it cannot readily dissociate from the site — it is trapped until further rotation of the γ subunit returns that β subunit to the O state, thereby releasing ATP and beginning the cycle again. ATP or ADP also binds to regulatory or allosteric sites on the three α subunits; this binding modifies the rate of ATP synthesis according to the level of ATP and ADP in the matrix, but is not directly involved in the catalytic step that synthesizes ATP from ADP and P_i.

Several types of evidence support the binding-change mechanism. First, biochemical studies showed that on isolated F_1 particles one of the three β subunits can tightly bind ADP and P_i and then form ATP, which remains tightly bound. The measured ΔG for this reaction is near zero, indicating that once ADP and P_i are bound to the T state of a β subunit, they spontaneously form ATP. Importantly, dissociation of the bound ATP from the β subunit on isolated F_1 particles occurs extremely slowly. This finding suggested that dissociation of ATP would have to be powered by a conformational change in the β subunit, which in turn would be due to **c**-ring rotation caused by proton movement.

A second line of evidence comes from x-ray crystallographic analysis of the $(\alpha\beta)_3$ hexamer, which yielded a striking conclusion: although the three β subunits are identical in sequence and overall structure, the ADP/ATP-binding sites have different conformations in each subunit. The most reasonable conclusion was that the three β subunits cycle in an energy-dependent reaction between three conformational states (O, L, T), in which the nucleotide-binding site has substantially different structures.

In other studies, intact F_0F_1 complexes were treated with chemical cross-linking agents that covalently linked the γ and ε subunits and the **c**-subunit ring. The observation that such treated complexes could synthesize ATP or use ATP to power proton pumping indicates that the cross-linked proteins normally rotate together.

Finally, rotation of the γ subunit relative to the fixed $(\alpha\beta)_3$ hexamer, as proposed in the binding-change mechanism, was observed directly by microscopy in the clever experiment depicted in Figure 12-29. A fluorescently labeled actin filament was attached to the γ subunit of F_1; the F_1-mediated ATP hydrolysis then caused the filament to rotate like a propeller in 120° steps. In one modification of this experiment in which tiny gold particles, rather than an actin filament, were attached to the γ subunit, rotation rates of 134 revolutions per second were observed. Hydrolysis of three ATPs, which you recall is the reverse reaction catalyzed by the same enzyme, is thought to power one revolution; this result is close to the experimentally determined rate of ATP hydrolysis by F_0F_1 complexes: about 400 ATPs per second. These and similar experiments established that the γ subunit, along with the attached **c** ring and ε subunit, does indeed rotate, thereby driving the conformational changes in the β subunits that are required for the binding of ADP and P_i, followed by synthesis and subsequent release of ATP.



EXPERIMENTAL FIGURE 12-29 The γ subunit of the F_1 subcomplex rotates relative to the $(\alpha\beta)_3$ hexamer. F_1 subcomplexes were engineered to contain β subunits with an additional His-6 sequence, which causes them to adhere to a glass plate coated with a metal reagent that binds polyhistidine. The γ subunit in the engineered F_1 subcomplexes was linked covalently to a fluorescently labeled actin filament. When viewed in a fluorescence microscope, the actin filament was seen to rotate counterclockwise in discrete 120° steps in the presence of ATP due to ATP hydrolysis by the β subunits. See H. Noji et al., 1997, *Nature* **386**: 299–302; and R. Yasuda et al., 1998, *Cell* **93**:1117–1124.

Multiple Protons Must Pass Through ATP Synthase to Synthesize One ATP

A simple calculation indicates that the passage of more than one proton across the membrane and down the electrochemical

gradient is required to synthesize one molecule of ATP from ADP and P_i. Although the ΔG for this reaction under standard conditions is $+7.3\,\text{kcal/mol}$, at the concentrations of reactants in the mitochondrion, ΔG is probably higher ($+10$ to $+12\,\text{kcal/mol}$). We can calculate the amount of free energy released by the passage of 1 mol of protons down an electrochemical gradient of 220 mV (0.22 V) from the Nernst equation, setting $n = 1$ and measuring ΔE in volts:

$$\begin{aligned}\Delta G(\text{cal/mol}) &= -nF\Delta E = -(23{,}062\,\text{cal}\cdot V^{-1}\cdot \text{mol}^{-1})\Delta E \\ &= (23{,}062\,\text{cal}\cdot V^{-1}\cdot \text{mol}^{-1})(0.22\,\text{V}) \\ &= 5074\,\text{cal/mol, or } -5.1\,\text{kcal/mol}\end{aligned}$$

Because the downhill movement of 1 mol of protons releases just over 5 kcal of free energy, the passage of at least two protons is required for synthesis of each molecule of ATP from ADP and P_i.

F_0 c Ring Rotation Is Driven by Protons Flowing Through Transmembrane Channels

Each copy of the **c** subunit contains two membrane-spanning α helices that form a hairpin-like structure (Figure 12-30a, *right*). An aspartate residue, Asp-61 (*E. coli* ATP synthase numbering), in the center of one of these helices in each **c** subunit is thought to play a key role in proton movement by binding and releasing protons as

they traverse the membrane. Chemical modification of this aspartate by the poison dicyclohexylcarbodiimide, or its mutation to alanine, specifically blocks proton movement through F_0. According to one current model, the protons traverse the membrane via two staggered half-channels, I and II (see Figure 12-27a and b). They are called *half-channels* because each extends only halfway across the membrane; the intramembrane termini of the channels are at the level of Asp-61 in the middle of the membrane. Half-channel I is open only to the exoplasmic face, and half-channel II is open only to the cytosolic face. Prior to rotation, each of the Asp-61 carboxylate side chains in the **c** subunits is bound to a proton, except for that on the **c** subunit in contact with half-channel I. The negative charge on that unprotonated carboxylate (the "empty" proton-binding site; see Figure 12-27b, *bottom*) is neutralized by interaction with the positively charged side chain of Arg-210 from the **a** subunit. Proton translocation across the membrane begins when a proton from the exoplasmic medium (intermembrane space in mitochondria) moves upward through half-channel I (see Figure 12-27b, step **1**). As that proton moves into the empty proton-binding site, it displaces the Arg-210 side chain, which swings toward the filled proton-binding site of the adjacent **c** subunit in contact with half-channel II (step **2**). As a consequence, the positive side chain of Arg-210 displaces the proton bound to Asp-61 of the adjacent **c** subunit. This displaced proton is now free to travel up half-channel II and out into the cytosolic medium (matrix in mitochondria) (step **3**). Thus when one proton entering from half-channel I binds to the **c** ring, a

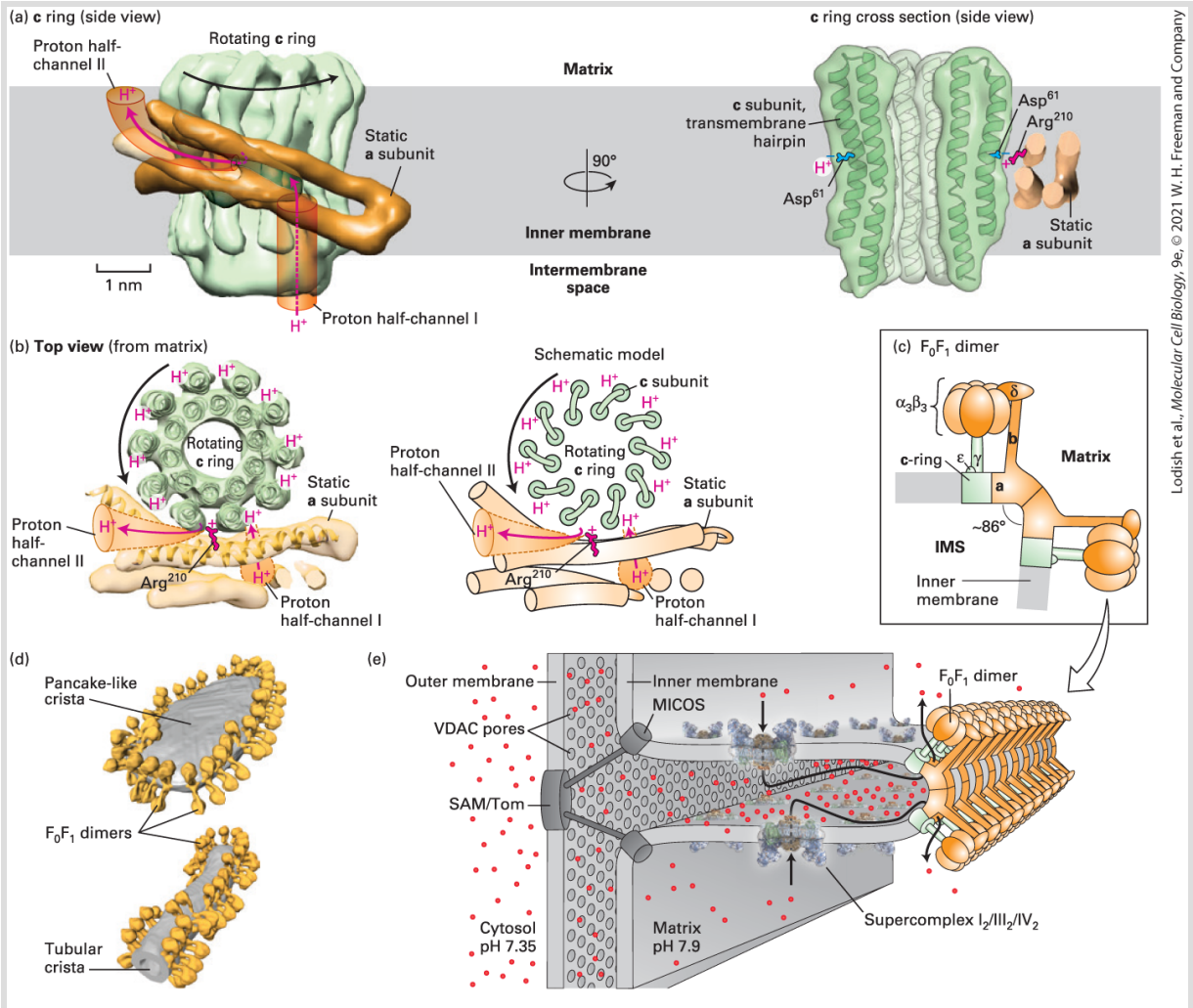
different proton on an adjacent **c** subunit is released to the opposite side of the membrane via half-channel II. Rotation of the entire **c** ring due to thermal/Brownian motion (step **4**) then allows the newly unprotonated **c** subunit to move into alignment above half-channel I as an adjacent, protonated **c** subunit rotates in to take its place under half-channel II. The entire cycle is then repeated (step **5**) as additional protons move down their electrochemical gradient from the exoplasmic medium to the cytosolic medium. During each partial rotation (360° divided by the number of **c** subunits in the ring), the **c** ring rotation is ratcheted, in that net movement of the ring occurs in only one direction. The energy driving the protons across the membrane, and thus the rotation of the **c** ring, comes from the electrochemical gradient across the membrane. If the direction of proton flow is reversed, which can be done by experimentally reversing the direction of the proton gradient and the proton-motive force, the direction of **c** ring rotation is reversed.

Because the γ subunit of F_1 is tightly attached to the **c** ring of F_0, rotation of the **c** ring associated with proton movement causes rotation of the γ subunit. According to the binding-change mechanism, a 120° rotation of γ powers synthesis of one ATP (see Figure 12-28). Thus complete rotation of the **c** ring by 360° would generate three ATPs. In *E. coli*, where the F_0 composition is $\mathbf{a}_1\mathbf{b}_2\mathbf{c}_{10}$, movement of 10 protons drives one complete rotation and thus synthesis of three ATPs. This value is consistent with experimental data on proton flux during ATP synthesis, providing indirect support

for the model coupling proton movement to **c** ring rotation depicted in Figure 12-27. The F_0 from mammals contains 8 **c** subunits per ring. The F_0 from chloroplasts contains 14 **c** subunits per ring, and movement of 14 protons would be needed for synthesis of three ATPs. Why these otherwise similar F_0F_1 complexes have evolved to have different H^+:ATP ratios is not clear.

High-resolution electron microscopic tomography (Figure 12-30) has provided additional insights into the structure of the **c** ring/**a** subunit interface and other features of F_0F_1 structure and function. The experiments were performed using F_0F_1 either dissolved in detergent, then incorporated into artificial phospholipid bilayers, or in isolated mitochondrial membranes. Figures 12-30a and b show two views of the two membrane-spanning α helices in each copy of the **c** subunit (green) that together form the **c** ring. In a portion of the **a** subunit (orange), a bundle of four α helices that are almost parallel to and embedded within the inner mitochondrial membrane forms the interface with the **c** ring and positions the side chain of the **a** subunit's Arg 210 adjacent to the **c** ring so that it can mediate proton displacement from Asp 61 as shown in Figure 12-27. The **c** ring/**a** subunit interface also forms the two proton half-channels through which protons flow out of the intermembrane space (red arrow), around the **c** ring (black arrows in Figure 12-30b), and then out into the matrix (red arrow). Each F_0F_1 monomer bends the membrane by approximately 43° (Figure 12-30c). The monomers dimerize to impart high membrane curvature ($\sim 86^\circ$) and then align

in long rows, contributing to the formation of the edges and tips of the pancake-like (flat) and tubular cristae (Figure 12-30d and e, also see Figure 12-6a).



EXPERIMENTAL FIGURE 12-30 High-resolution electron microscopy-based mechanism of proton translocation and bending of cristae membranes by ATP synthase. (a) and (b) The interface between the **c** ring (green) and **a** subunit (orange) of detergent-solubilized mitochondrial ATP synthase from the alga *Polytomella* sp., imaged by single-particle cryoelectron microscopy ($\sim 0.62\ \mathrm{nm}$ resolution), is shown (a) from within the plane of the inner mitochondrial membrane (side view) and (b) after a 90° rotation (top view). The movement of protons through half-channels I and II and the rotation of the **c** ring are described in detail in Figure 12-27. (a) Cross section through the **c** ring (right) shows that

each **c** subunit is a transmembrane helical hairpin — two adjacent transmembrane α helices connected by a short nonhelical linker on the matrix side of the membrane. The negative side chain of the **c** subunit's Asp61 in the middle of the membrane is thought to both serve as a binding site for translocating protons and interact with the side chain of the **a** subunit's Arg210. (c) A model of the bovine heart mitochondrial ATP synthase dimer is based on cryoelectron tomography and electron crystallographic image processing from crystalline ATP synthase in artificial membranes. Each $\mathrm{F_0F_1}$ monomer bends the membrane by ~ 43° toward the intermembrane space (IMS), resulting in dimers bending the membrane by ~ 86°. The rotating **c** ring and γ and ε subunits are colored green, and the remaining static portions of the enzyme are shown in orange. (d) Cryoelectron tomographic image of frozen membranes from purified *Saccharomyces cerevisiae* (yeast) mitochondria. The surfaces of the ATP synthase complexes (yellow) and the membrane (gray) show that the enzymes dimerize as in (c) and align into long rows that bend the membranes into characteristic tubular and flat, pancake-like cristae. (e) Cartoon illustrating the topological relationships of the rows of ATP synthase dimers (orange and green) on the curved edge of a crista to other components of the inner membrane (electron transport supercomplexes $(\mathrm{I_2/III_2/IV_2})$ and the crista junction complex MICOS, gray) and the outer membrane (SAM and Tom complexes, gray) that link MICOS to the outer membrane. The cycling of protons (red spheres) from the matrix through the electron-transport chain to the inner membrane space and then back into the matrix via ATP synthase is indicated by black arrows. The relative sizes of the proteins and membrane bilayers are not drawn to scale.

[Parts (a) and (b) Reprinted with permission of Nature Publishing Group, from Allegretti, M., et al., "Horizontal membrane-intrinsic α-helices in the stator a-subunit of an F-type ATP synthase," *Nature*, 2015, **521**, pp. 237–240, 2015; permission conveyed through Copyright Clearance Center, Inc.; Part (c) From *Proc Natl Acad Sci USA* 2012. **109**(34):13602–13607, Fig. 4C and D. *Structure of the yeast F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae* by Davies et al. Part (e) Data from W. Kühlbrandt, 2015, *BMC Biol.* **13**:89.]

ATP-ADP Exchange and Phosphate Transport Across the Inner

Mitochondrial Membrane Are Required to Supply ADP and Phosphate for ATP Synthesis

It is estimated that at any one time an adult human contains about 50 g of ATP, yet that same individual must hydrolyze 65 kg of ATP per day. Thus mitochondria must generate 65 kg of ATP/day per person from ADP and P_i and transport most of that ATP out of the matrix where it is generated (Figures 12-11 and 12-26) so that it can be used elsewhere in the cells. Clearly large amounts of ADP and P_i must be transported into the matrix to sustain ATP synthesis. This transport is mediated by two proteins in the inner membrane (Figure 12-31): a *phosphate* (P_i, HPO_4^{2-}) *transporter*, which mediates the import of one HPO_4^{2-} coupled to either the import of one proton or the export of one OH^- (export of OH^- shown in the figure and we will assume phosphate/hydroxyl exchange below), and an *ATP/ADP exchanger* (see Figure 12-31). The 30,000-Da ATP/ADP exchanger makes up 10–15 percent of the protein in the inner mitochondrial membrane, so it is one of the more abundant mitochondrial proteins. Both of these transporters are members of the SLC25 family of mitochondrial inner membrane transporters. The SLC25 transporters each have six transmembrane α-helices and use the alternating access mechanism (see Figure 11-6) to move their substrates across the inner mitochondrial membrane. The P_i and ATP/ADP transporters together produce an influx into the matrix

of one ADP^{3-} and one P_i^{2-} and an efflux of one ATP^{4-} together with one OH^-. Each OH^- transported outward combines with a proton, translocated during electron transport to the intermembrane space, to form H_2O. Thus proton translocation via electron transport together with the relatively high matrix concentration of ATP under normal circumstances drive the overall reaction in the direction of ATP export and ADP and P_i import. In adult humans, each ATP molecule is recycled across the inner membrane about 1300 times a day.

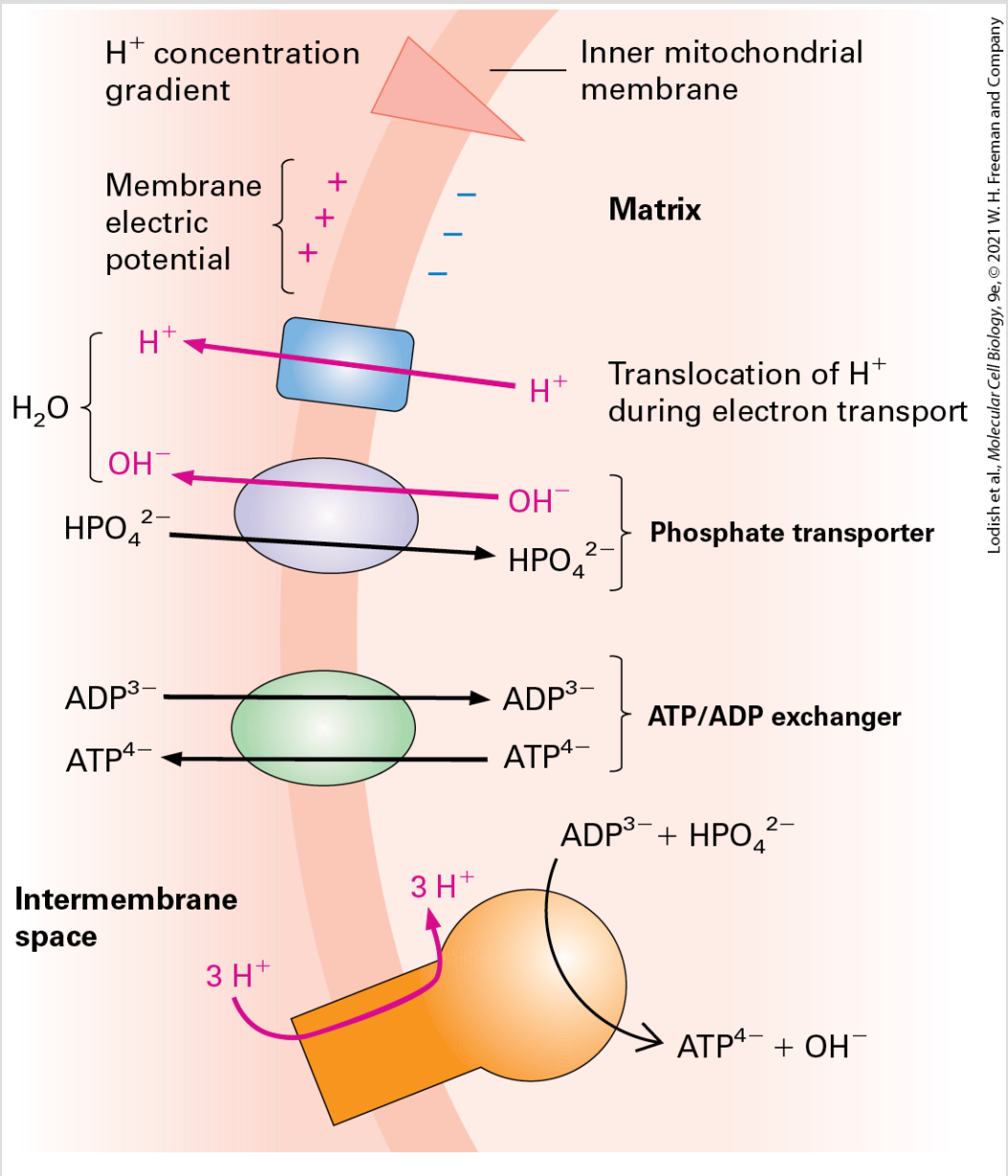


FIGURE 12-31 The phosphate and $\mathbf{ATP/ADP}$ transport system in the inner mitochondrial membrane. The coordinated action of two transporters (purple and green ovals), which results in the uptake into the matrix of one ADP^{3-} and one $HPO_4{}^{2-}$ in exchange for one ATP^{4-} and one hydroxyl, is powered by the outward translocation of one proton (mediated by the proteins of the electron-transport chain, blue rectangle) during electron transport. Red arrows indicate that movement of a proton and hydroxyl out of the matrix results in export of H_2O. The outer membrane is not shown here because it is permeable to molecules smaller than 5000 Da.

Because some of the protons translocated out of the mitochondrial matrix during electron transport provide the power (by combining with the exported OH^-) for the phosphate transport and help neutralize the extra negative charge carried into the matrix on ATP relative to ADP, fewer protons are available for ATP synthesis. It is estimated that for every four protons translocated across the inner membrane during electron transport, about three are used to synthesize one ATP molecule and one is used to power the export of ATP from the mitochondrion in exchange for ADP and P_i. This expenditure of energy from the proton concentration gradient to export ATP from the mitochondrion in exchange for ADP and P_i ensures a high ratio of ATP to ADP in the cytosol, where hydrolysis of the high-energy phosphoanhydride bond of ATP is used to power many energy-requiring reactions.

Studies of what turned out to be ATP/ADP exchanger activity were first recorded about 2000 years ago, when Dioscorides (~AD 40–90) described the effects of a poisonous herb from the thistle *Atractylis gummifera*. The same agent is found in the traditional Zulu multipurpose herbal remedy *impila* (*Callilepis laureola*), although this herb has caused numerous poisonings. The active agent, atractyloside, inhibits the ATP/ADP transporter and thus inhibits oxidative phosphorylation of extramitochondrial ADP, but not intramitochondrial (matrix) ADP. Atractyloside, which shows the importance of the ATP/ADP transporter, continues to be used to study the mechanism by which this transporter functions.

Dioscorides lived near Tarsus, at the time a province of Rome in southeastern Asia Minor, in what is now Turkey. His five-volume *De Materia Medica* (*The Materials of Medicine*) "on the preparation, properties, and testing of drugs" described the medicinal properties of about 1000 natural products. For about 1600 years, it was the basic reference in medicine for using drugs from northern Europe to the Indian Ocean, comparable to today's *Physicians' Desk Reference*. ■

The Rate of Mitochondrial Oxidation Normally Depends on ADP Levels

If intact isolated mitochondria are provided with NADH (or a source of $FADH_2$ such as succinate) plus O_2 and P_i, but not ADP, the oxidation of NADH and the reduction of O_2 rapidly cease as the amount of endogenous ADP is depleted by ATP formation. If ADP is then added, the oxidation of NADH is rapidly restored. Thus mitochondria can oxidize $FADH_2$ and NADH only as long as there is a source of ADP and P_i to generate ATP. This phenomenon, termed **respiratory control**, occurs because oxidation of NADH and succinate ($FADH_2$) is obligatorily coupled to proton transport across the inner mitochondrial membrane. If the resulting proton-motive force is not dissipated during the synthesis of ATP from ADP and P_i (or during other energy-requiring processes), both the transmembrane proton concentration gradient and the membrane electric potential will increase to very high levels. At this point,

pumping of additional protons from the matrix across the inner membrane requires so much energy that it eventually ceases, blocking the coupled oxidation of NADH and other substrates.

Mitochondria in Brown Fat Use the Proton-Motive Force to Generate Heat

Brown-fat tissue, whose color is due to the presence of abundant mitochondria, is specialized for the generation of heat. In contrast, *white-fat tissue* is specialized for the storage of fat and contains relatively few mitochondria.

The inner membranes of brown-fat mitochondria contain a protein called **uncoupling protein 1 (UCP1)**, a protein that functions as an **uncoupler** of oxidative phosphorylation and ATP generation. UCP1 is one of several uncoupling proteins (UCPs) found in most eukaryotes (but not in fermentative yeasts). As with the ATP/ADP exchanger and the phosphate transporter, UCPs are members of the SLC25 family of transporters. UCP1, which requires a bound fatty acid molecule as a cofactor, functions as an intermembrane space-to-matrix proton transporter. Fatty-acid-containing UCP1 dissipates the proton-motive force by rendering the inner mitochondrial membrane permeable to protons. As a consequence, the energy released by NADH oxidation in the electron-transport chain and used to transport protons out of the matrix is not available to be

used to synthesize ATP via ATP synthase. Instead, when protons move back into the matrix down their concentration gradient via UCP1, the energy is released as heat. UCP1 is a proton transporter, not a proton channel, and shuttles protons across the membrane at a rate that is 1-million-fold slower than that of typical ion channels (see Figure 11-3). Certain small-molecule poisons also function as uncouplers by rendering the inner mitochondrial membrane permeable to protons. One example is the lipid-soluble chemical 2,4-dinitrophenol (DNP), which can reversibly bind to and release protons and shuttle them across the inner membrane from the intermembrane space into the matrix. Remarkably, when a fatty acid binds to the ATP/ADP exchanger, this transporter can also transport protons from the matrix to the intermembrane space and thus, just like UCP1, uncouple oxidative phosphorylation that generates a proton-motive force and ATP synthesis.

Environmental conditions regulate the amount of UCP1 in brown-fat mitochondria. For instance, when rats adapt to cold temperatures, the ability of their tissues to generate heat is increased by the induction of UCP1 synthesis. In cold-adapted animals, UCP1 may constitute up to 15 percent of the total protein in the inner membranes of brown-fat mitochondria.

In the newborn human, UCP1 expression by brown-fat mitochondria is vital to survival, as it is in hibernating mammals. In adult humans there is usually little brown fat, which is located in the

neck, clavicle, and other sites. The amounts of human brown fat are significantly increased upon exposure to cold. In fur seals and other animals naturally acclimated to the cold, muscle-cell mitochondria contain UCP1; as a result, much of the proton-motive force is used for generating heat, thereby maintaining body temperature. Detailed analyses of the biochemical properties and developmental origins of thermogenic fat cells have uncovered the existence of at least two subtypes of such cells: classic brown-fat cells that develop from precursor cells also used to generate skeletal muscle cells, and *beige-fat cells* whose detailed properties (gene expression pattern, responses to hormonal signals) differ from those of brown-fat cells and are formed from white-fat cells or their progenitors. Thus there are three types of fat: brown, white, and beige. Further characterization of beige-fat cells and their influence on normal metabolism and disease may lead to new approaches to treat or prevent some metabolic disorders.

KEY CONCEPTS OF SECTION 12.7

Harnessing the Proton-Motive Force to Synthesize ATP

- The chemiosmotic hypothesis, proposed by Peter Mitchell, states that a proton-motive force across the inner mitochondrial membrane is the immediate source of energy for ATP synthesis.
- Bacteria, mitochondria, and chloroplasts all use the same chemiosmotic mechanism and a similar ATP synthase to generate ATP (see Figure 12-26).
- ATP synthase (also called the F_0F_1 *complex*) catalyzes ATP synthesis as protons flow through the inner mitochondrial membrane (the plasma membrane in bacteria) down their electrochemical proton gradient.

- F_0 contains a ring of 8–14 **c** subunits, depending on the organism, that is rigidly linked to the rod-shaped γ subunit and the ε subunit of F_1. These subunits rotate during ATP synthesis. Resting atop the γ subunit is the hexameric knob of F_1 $[(\alpha\beta)_3]$, which protrudes into the mitochondrial matrix (cytosol in bacteria). The three β subunits are the sites of ATP synthesis (see Figure 12-27 and 12-30a and b).
- Rotation of the asymmetic F_1 γ subunit, which is inserted in the center of the nonrotating $(\alpha\beta)_3$ hexamer, leads to sequential changes in the conformation of the nucleotide-binding sites in the three F_1 β subunits (see Figure 12-28). By means of this binding-change mechanism, the β subunits bind ADP and P_i, condense them to form ATP, and then release the ATP. Three ATPs are made for each revolution of the rigid complex of the **c** ring and γ and ε subunits.
- Movement of protons across the membrane via two half-channels at the interface of the F_0 **a** subunit and the **c** ring powers rotation of the **c** ring with its attached F_1 ε and γ subunits.
- The F_0F_1 complex bends the inner mitochondrial membrane, contributing to its characteristic high curvature and to the tubular and pancake-like structures of the cristae (see Figure 12-30c, d, and e).
- The proton-motive force also powers the uptake of P_i and ADP from the cytosol in exchange for mitochondrial ATP and OH^-, thus reducing the energy available for ATP synthesis. The ATP/ADP antiporter that participates in this exchange is one of the most abundant proteins in the inner mitochondrial membrane (see Figure 12-31).
- Continued mitochondrial oxidation of NADH and reduction of O_2 are dependent on sufficient ADP being present in the matrix. This phenomenon, termed *respiratory control*, is an important mechanism for coordinating oxidation and ATP synthesis in mitochondria.
- In brown fat, the inner mitochondrial membrane contains the uncoupler protein UCP1, a proton transporter that dissipates the proton-motive force into heat. Certain chemicals also function as uncouplers (e.g., DNP) and have the same effect, uncoupling oxidative phosphorylation from electron transport. There are two distinct types of thermogenic fat cells: brown-fat and beige-fat cells.

12.8 Chloroplasts and Photosynthesis

We now shift our attention to photosynthesis, the second key process for synthesizing ATP. In plants, photosynthesis occurs in chloroplasts, large organelles found mainly in leaf cells. During photosynthesis, chloroplasts capture the energy of sunlight, convert it into chemical energy in the form of ATP and NADPH, and then use this energy to synthesize complex carbohydrates out of carbon dioxide and water. It is estimated that annually photosynthesis on earth results in the conversion of 104.9×10^{15} g of carbon from CO_2 into organic molecules. In cyanobacteria, algae, and plants photosynthesis also generates molecular oxygen from water. The principal carbohydrates produced are polymers of hexose (six-carbon) sugars: sucrose, a glucose-fructose disaccharide (see Figure 2-19), and **starch**, a mixture of two types of large, insoluble glucose polymers called *amylose* and *amylopectin*. Starch is the primary storage carbohydrate in plants (Figure 12-32). Starch is synthesized and stored in the chloroplast. Sucrose is synthesized in the leaf cytosol from three-carbon precursors generated in the chloroplast; it is transported to non-photosynthetic (nongreen) plant tissues (e.g., roots and seeds), which metabolize it for energy by the pathways described in the previous sections.

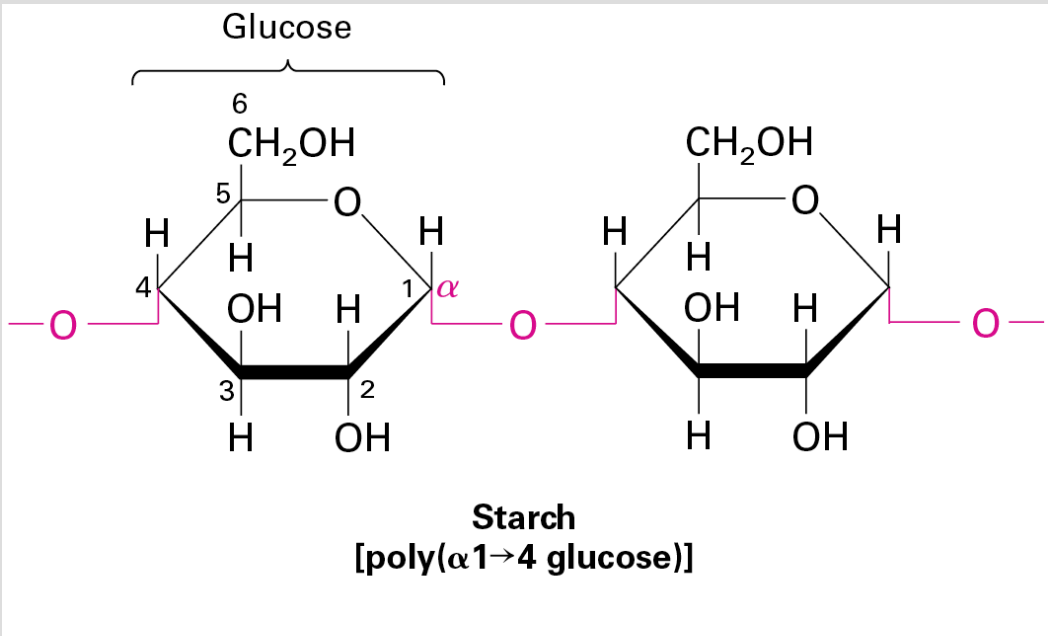


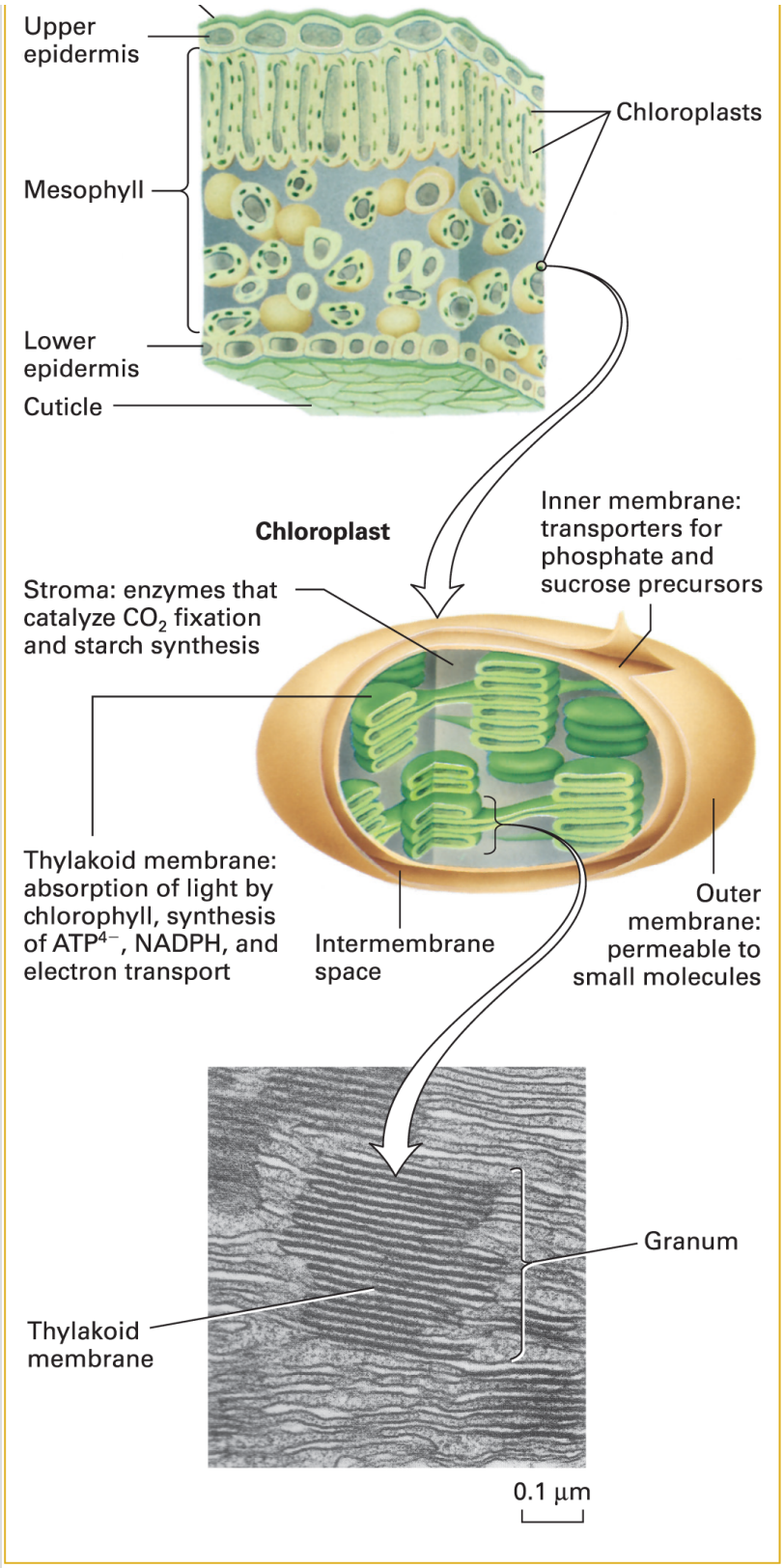
FIGURE 12-32 Structure of starch. The large glucose polymer and the disaccharide sucrose (see Figure 12-15) are the principal end products of photosynthesis. Both are built of six-carbon sugars (hexoses).

Thylakoid Membranes in Chloroplasts Are the Sites of Photosynthesis in Plants

Chloroplasts are lens-shaped organelles with a diameter of approximately 5 μm and a width of approximately 2.5 μm (Figure 12-33 and 1-20). They contain >2500 different proteins, 95 percent of which are encoded in the nucleus, made in the cytosol, imported into the organelle, and then transported to their appropriate membrane or space (see Chapter 13). Chloroplasts are bounded by two membranes, which do not contain chlorophyll and do not

participate directly in the generation of ATP and NADPH driven by light (Figures 12-33 and 12-34). Like that of mitochondria, the outer membrane of chloroplasts contains porins and thus is permeable to metabolites of small molecular weight. The inner membrane forms a permeability barrier that contains transport proteins for regulating the movement of metabolites into and out of the organelle.

Unlike mitochondria, chloroplasts contain a third membrane — the **thylakoid membrane** — on which the light-driven generation of ATP and NADPH occurs. The chloroplast thylakoid membrane is believed to constitute a single sheet that forms numerous small, interconnected flattened structures, the **thylakoids**, which are commonly arranged in stacks termed *grana* (see Figure 12-33). The spaces within all the thylakoids constitute a single continuous compartment, the **thylakoid lumen** (see Figure 12-34). The thylakoid membrane contains a number of integral membrane proteins; bound to these are several important prosthetic groups and light-absorbing pigments, most notably chlorophylls. Starch synthesis and storage occurs in the **stroma**, the aqueous compartment between the thylakoid membrane and the inner membrane. In photosynthetic bacteria, extensive invaginations of the plasma membrane form a set of internal membranes, also termed *thylakoid membranes*, where photosynthesis occurs.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freem

FIGURE 12-33 Structure of leaf and chloroplast. Like mitochondria, plant chloroplasts are bounded by two membranes separated by an intermembrane space. Photosynthesis occurs on a third membrane, the thylakoid membrane, which is surrounded by the inner membrane and which forms a series of flattened vesicles (thylakoids) that enclose a luminal space. The green color of plants is due to the color of chlorophyll and carotenoids, all of which is located within the thylakoid membrane. A granum is a stack of adjacent thylakoids. The stroma is the space between the inner membrane and the thylakoids.

[From Katherine Esau, D-120, Special Collections, University of California Library, Davis.]

Chloroplasts Contain Large DNAs Often Encoding More Than a Hundred Proteins

Like mitochondria, chloroplasts are thought to have evolved from an ancestral endosymbiotic photosynthetic bacterium (see Figure 12-7). However, the endosymbiotic event that gave rise to chloroplasts occurred more recently (1.2 billion–1.5 billion years ago) than the event that led to the evolution of mitochondria (1.5 billion–2.2 billion years ago) (see Figure 1-1). Consequently, contemporary chloroplast DNAs show less sequence diversity than do mtDNAs. As with mtDNA, chloroplast DNA is inherited through the female parent (egg).

In higher plants, chloroplast DNA molecules are 120–160 kb long, depending on the species. Plant chloroplast DNAs are long head-to-

tail linear concatemers plus recombination intermediates between these long linear molecules. They contain 120–135 genes, 130 in the important model plant *Arabidopsis thaliana* (see Figure 1-23). *A. thaliana* chloroplast DNA encodes 76 protein-coding genes and 54 genes with RNA products such as rRNAs and tRNAs. Chloroplast DNAs encode the subunits of a bacteria-like RNA polymerase, and they express many of their genes from polycistronic operons, in which multiple proteins are translated from a single messenger RNA, as in bacteria. Some chloroplast genes contain introns, but these introns are similar to the specialized introns found in some bacterial genes and in mitochondrial genes from fungi and protozoans, rather than the introns of nuclear genes. Many genes essential for chloroplast function have been transferred to the nuclear genome of plants over evolutionary time. Chloroplast proteins that are encoded by nuclear genes are synthesized on cytosolic ribosomes and then incorporated into the organelle (see Chapter 13). Recent estimates from sequence analysis of the *A. thaliana* and cyanobacterial genomes indicate that somewhat less than 4500 genes have been transferred from the original endosymbiont to the nuclear genome.

Chloroplast DNA can be genetically manipulated by a process similar to that used for the transformation of yeast cells (see Chapter 6) or by techniques such as CRISPR/Cas9 to engineer plants that are resistant to bacterial and fungal infections, drought, and herbicides as well as engineered so that they can be used to make

human pharmaceutical drugs (called *pharming*). The first such pharming drug, approved in the United States for use in adults in 2012 and children in 2014, is an enzyme to treat Gaucher's disease, a genetic disorder. This approach might also be used for the engineering of food crops containing high levels of all the amino acids essential to humans. ■

Light Absorption by Photosystems in Chloroplasts Provides the Energy That Drives the Synthesis of NADPH and ATP and the Generation of O_2 from H_2O

In the 1950s, biophysicist Robert Emerson and colleagues discovered that the rate of plant photosynthesis generated by light of wavelength 700 nm could be greatly enhanced by adding light of shorter wavelengths (higher energy). They found that a combination of light at, say, 600 and 700 nm supports a rate of photosynthesis higher than the sum of the rates for the two separate wavelengths. This so-called Emerson effect led researchers to conclude that photosynthesis in plants, as well as in eukaryotic single-celled algae and in several photosynthetic bacteria (e.g., the cyanobacteria and prochlorophytes), involves the interaction of two separate multiprotein complexes called **photosystems**, each of which

absorbs light at defined wavelengths. Photosystems are found in all photosynthetic organisms, both eukaryotic and prokaryotic. We will focus on the two photosystems found in plants, **photosystem I (PSI)** that is primarily in unstacked regions of the thylakoid membranes and **photosystem II (PSII)** that is primarily in regions of stacked thylakoids (grana, see Figure 12-33).

Figure 12-34 shows an overview of the four stages of photosynthesis in plants. PSI (Figure 12-34, stage 2) is driven by light of 700 nm or less; PSII (Figure 12-34, stage 1), only by shorter wavelength light ($< 680\,\mathrm{nm}$). In stage 1 in cyanobacteria, algae, and plants PSII absorbs light, generates high energy electrons, and transfers those electrons to a quinone electron carrier, similar in structure to Coenzyme Q, called QH_2. PSII also forms O_2 from H_2O. In stage 2, the energy from the electrons in QH_2 drives a **proton pump** in the multiprotein cytochrome *bf* complex that generates a proton-motive force across the thylakoid membrane, which, as in mitochondria, is subsequently used for ATP production (stage 3). The energy-depleted electrons in the cytochrome *bf* complex are transferred by a small protein carrier called *plastocyanin* to PSI which uses light to reenergize the electrons that are then used by PSI to reduce $NADP^+$ to NADPH. The ATP and NADPH are used in stage 4 to incorporate carbon from CO_2 into small molecules (commonly referred to as **carbon fixation**) and then into six carbon sugars and starch.

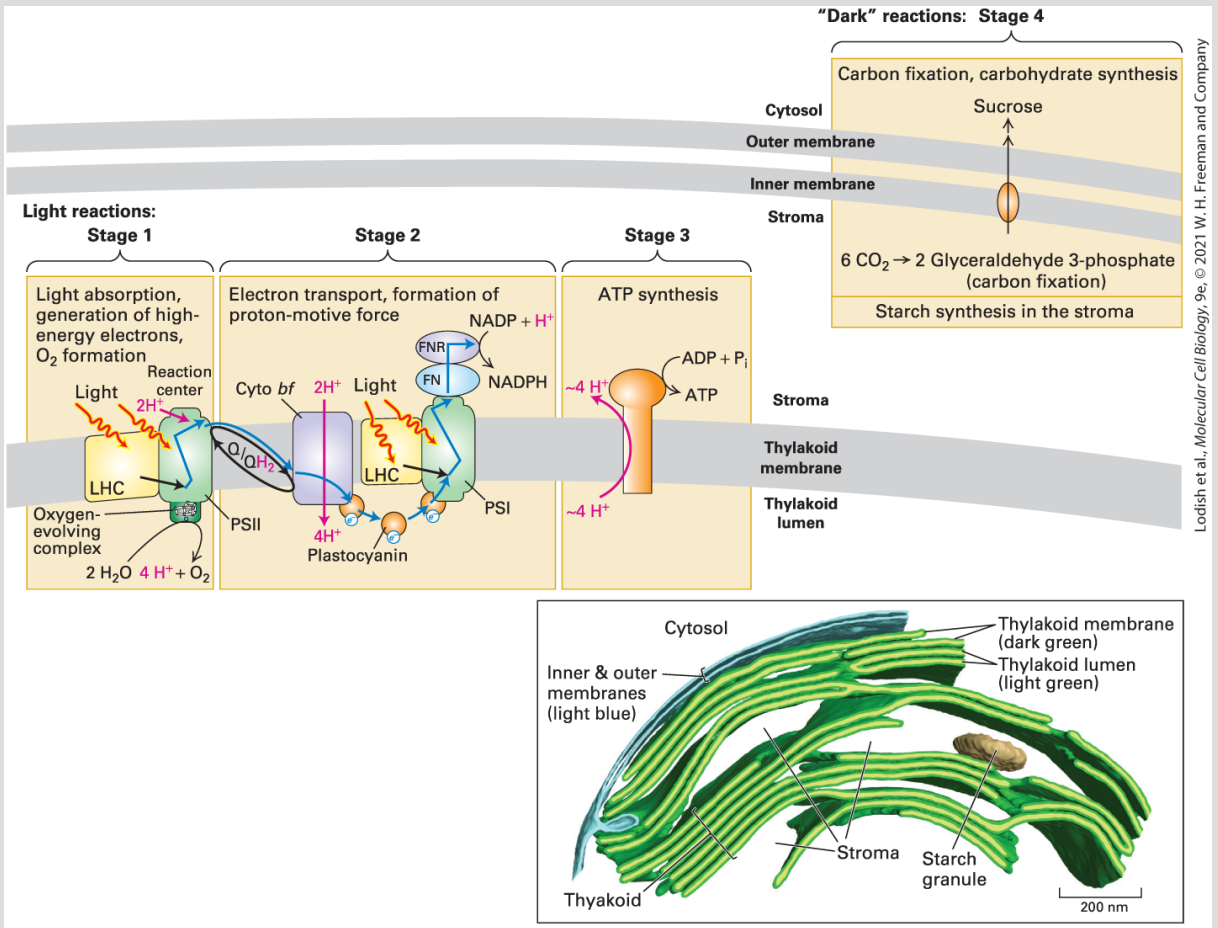


FIGURE 12-34 Overview of the four stages of photosynthesis. In **stage 1**, light is absorbed by light-harvesting complexes (LHCs) and the reaction center of photosystem II (PSII). LHCs (yellow) transfer the absorbed energy to PSII's reaction center (black arrow), which use it or the energy absorbed directly from photons to (i) generate high-energy electrons (electron paths shown by blue arrows) and (ii) oxidize water to molecular oxygen via PSII's oxygen-evolving complex. In PSII, oxygen is formed on the lumenal side of the membrane, and the high-energy electrons are transferred to a quinone molecule, Q, bound near the stromal (cytosolic) face of the thylakoid membrane. The Q then combines with two protons to generate the lipid-soluble reduced quinone QH_2 that is released within the thylakoid membrane for diffusion to the cytochrome *bf* complex (Cyto *bf*) in **stage 2**. The oxygen-evolving complex of PSII contains a manganese (Mn, violet), calcium (green), and oxygen (red) cluster (Mn_4CaO_5, represented by the ball-and-stick model) that directly mediates the oxidation of water, as well as a bound Cl^- ion (teal) that influences the reaction rate.

In **stage 2**, QH_2 from **stage 1** binds to a site near the lumenal surface on the cytochrome *bf* complex (Cyto *bf*), a proton pump. The energy stored in QH_2 is released by a Q cycle similar

to that in mitochondrial complex III (see Figure 12-21). This energy drives protons from QH_2 and from the stroma across the thylakoid membrane to generate a proton-motive force. Then the two energy-depleted electrons originally from QH_2 are transferred from Cyto *bf*, via a water-soluble protein called *plastocyanin*, to photosystem I (PSI). After additional energy is transferred to those electrons by absorption of light in PSI and its associated LHC, they move to ferredoxin (FN) and then to ferredoxin-$NADP^+$ reductase (FNR), which uses their energy to reduce NAD and form the high-energy electron carrier NADPH. In **stage 3**, as in mitochondria, flow of protons down their concentration and voltage gradient through the F_0F_1 ATP synthase drives ATP synthesis. The chloroplast F_0F_1 ATP synthase generates one ATP for every 4–5 protons that pass through it. Stages 1–3 in plants take place in the thylakoid membrane of the chloroplast. In **stage 4**, in the chloroplast stroma, the energy stored in NADPH and ATP is used to incorporate CO_2 into the three-carbon molecule glyceraldehyde 3-phosphate, the first step in a process known as *carbon fixation*. These molecules are then transported to the cytosol of the cell for conversion to hexose sugars in the form of sucrose. Glyceraldehyde 3-phosphate is also used to make starch within the chloroplast. *Inset:* Three-dimensional reconstruction from cryoelectron tomography of a chloroplast in the unicellular green alga *Chlamydomonas reinhardtii*, showing thylakoid membranes (dark green), thylakoid lumen (light green), inner and outer membranes (blue), and one small starch granule (tan).

[Inset from Engel, B. D., et al., "Native Architecture of the Chlamydomonas Chloroplast Revealed by in Situ Cryo-Electron Tomography," *eLIFE*, 2015;4: e04889]

The overall reaction of oxygen-generating photosynthesis in plants,

$$6\,CO_2 + 6\,H_2O \rightarrow 6\,O_2 + C_6H_{12}O_6$$

is the reverse of the overall reaction by which carbohydrates are oxidized to CO_2 and H_2O. In effect, photosynthesis in chloroplasts produces energy-rich sugars that are broken down and harvested for energy by mitochondria using oxidative phosphorylation.

Photosynthesis in green and purple bacteria employs a PSI-like system, but not PSII, and does not generate oxygen.

In this section, we provide an overview of the stages in oxygen-generating photosynthesis and introduce the main molecular components of the process, including the chlorophylls, the principal light-absorbing pigments.

Three of the Four Stages in Photosynthesis Occur on the Thylakoid Membrane and Only During Illumination

Each of the four stages of photosynthesis is localized to a defined area of the chloroplast. All the reactions in stages 1–3 are catalyzed by multiprotein complexes embedded in the thylakoid membrane. The generation of a proton-motive force and the use of that proton-motive force to synthesize ATP resemble stages III and IV of mitochondrial oxidative phosphorylation. The enzymes of carbon fixation that incorporate CO_2 into chemical intermediates and then convert them to starch are soluble constituents of the chloroplast stroma; the enzymes that form sucrose from three-carbon intermediates are in the cytosol. All four stages of photosynthesis are tightly coupled and controlled so as to produce the amount of

carbohydrate required by the plant and to respond appropriately to environmental conditions.

The reactions that generate the ATP and NADPH used in carbon fixation are directly dependent on light energy; thus stages 1–3 are called the *light reactions* of photosynthesis. The reactions in stage 4 are indirectly dependent on light energy; they are sometimes called the *dark reactions* of photosynthesis because they can occur in the dark, using the supplies of ATP and NADPH generated by light energy (see Figure 12-34). However, the reactions in stage 4 are not confined to the dark; in fact, they occur primarily during illumination.

Stages 1 and 2 of Photosynthesis Convert Sunlight into High Energy Electrons That Generate a Proton-Motive Force and NADPH

The absorption of light energy and its conversion into chemical energy in plants occurs in photosystem II (PSII in stage 1 and photosystem I (PSI) in stage 2 (Figure 12-34). Photosystems — sometimes referred to as *core photosystems* — consist of two closely linked components: a *reaction center*, where the primary events of photosynthesis — light absorption and generation of high-energy

electrons — occur; and core *antenna complexes*. Each photosystem is also associated with external antenna complexes termed **light-harvesting complexes (LHCs)**, made up of specialized proteins that capture light energy and efficiently transmit it to the reaction center to generate high-energy electrons (see Figure 12-34).

The proteins of the reaction centers and the antenna complexes contain tightly bound light-absorbing pigment molecules. Chlorophyll *a*, the principal pigment involved in photosynthesis, is present in both reaction centers and antenna complexes. The initial step in photosynthesis is the absorption of light by chlorophyll *a* and other pigments. Like the heme component of cytochromes, chlorophylls consist of a porphyrin ring attached to a long hydrocarbon side chain (Figure 12-35). In contrast to the hemes (see Figure 12-17), chlorophylls contain a central Mg^{2+} ion (rather than Fe^{2+}) and have an additional five-member ring.

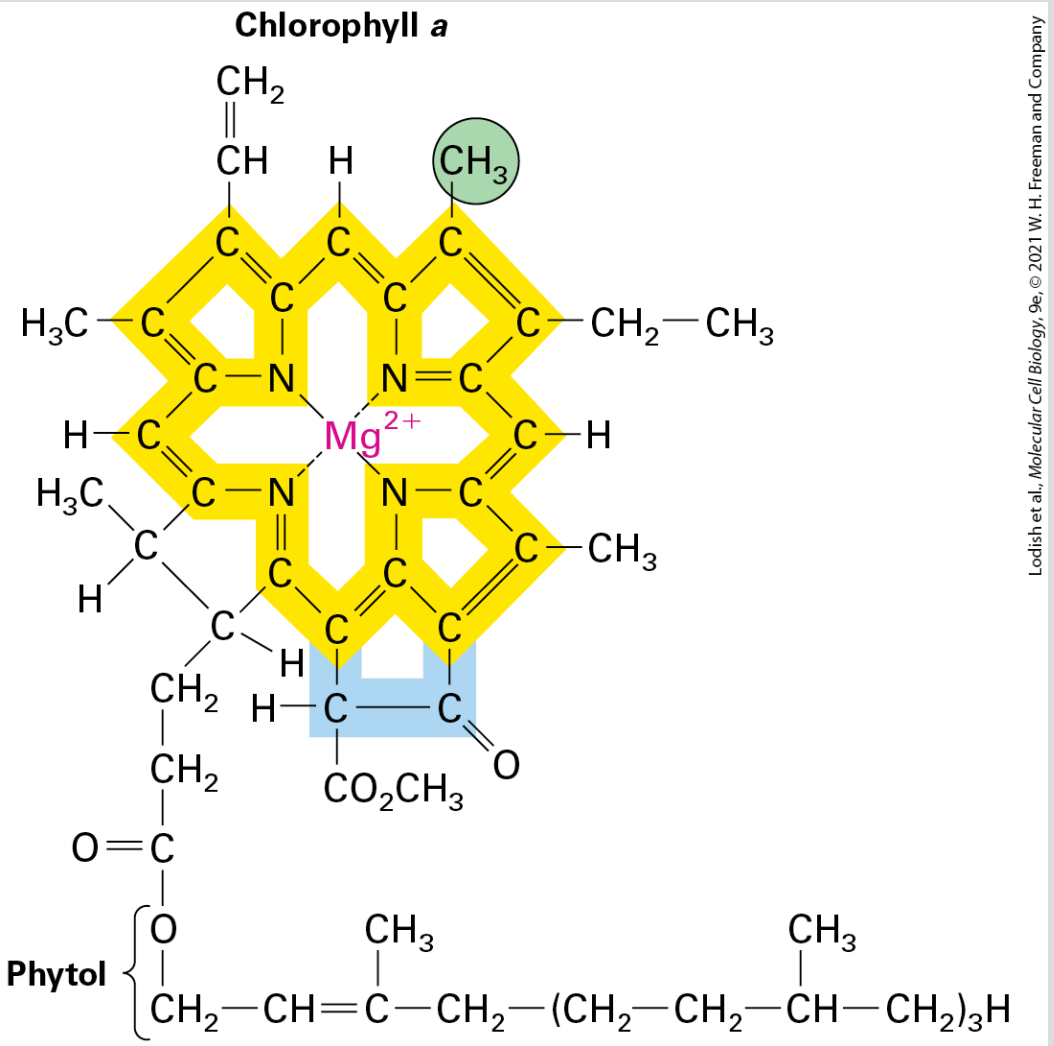
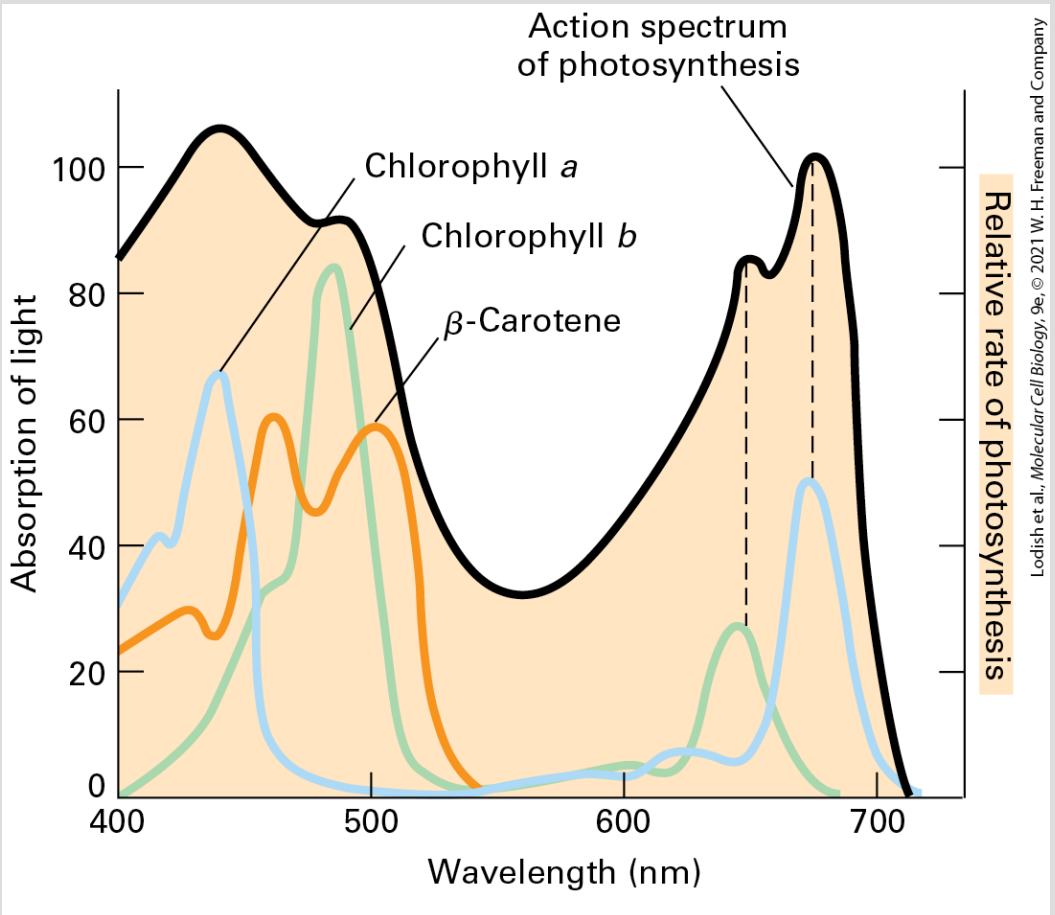


FIGURE 12-35 Structure of chlorophyll *a*, the principal pigment that traps light energy. Electrons are delocalized among three of chlorophyll *a*'s four central rings (yellow) and the atoms that interconnect them. In chlorophyll, a Mg^{2+} ion, rather than the Fe^{2+} ion found in heme, sits at the center of the porphyrin ring, and an additional five-member ring (blue) is present; otherwise, the structure of chlorophyll is similar to that of heme, found in molecules such as hemoglobin and cytochromes (see Figure 12-17a). The hydrocarbon phytol "tail" facilitates the binding of chlorophyll to hydrophobic regions of chlorophyll-binding proteins. The CH_3 group (green) is replaced by a formaldehyde (CHO) group in chlorophyll *b*.

The antenna complexes contain both chlorophyll *a* and other light-absorbing pigments, including other forms of chlorophyll (e.g., *chlorophyll* b in vascular plants) and *carotenoids* in both plants, algae, and photosynthetic bacteria. Carotenoids, of which there are several different molecular types, consist of long branched hydrocarbon chains with alternating single and double bonds; they are similar in structure to the visual pigment retinal (see Figure 15-34b), which absorbs light in the eye. The presence of various antenna pigments, which absorb light at different wavelengths, greatly extends the range of light that can be absorbed and used for photosynthesis. One of the strongest pieces of evidence for the involvement of chlorophylls and carotenoids in photosynthesis is that the absorption spectrum of these pigments is similar to the action spectrum of photosynthesis (Figure 12-36). The latter is a measure of the relative ability of light of different wavelengths to support photosynthesis.



EXPERIMENTAL FIGURE 12-36 The rate of photosynthesis is greatest at the wavelengths of light absorbed by three plant pigments. The action spectrum of photosynthesis in plants (the relative ability of light of different wavelengths to support photosynthesis) is shown in black. The energy from light can be converted into ATP only if it can be absorbed by pigments in the chloroplast. Absorption spectra (showing how well light of different wavelengths is absorbed) for three photosynthetic pigments (chlorophyll *a*, chlorophyll *b*, and β-carotene) present in the antennas of plant photosystems are shown in color. This figure shows as a function of the wavelength of light in nanometers (horizontal axis) a comparison of the action spectrum of photosynthesis (*right* vertical axis) with the individual absorption spectra of these pigments (*left* vertical axis). The data suggest that photosynthesis at 680 nm is primarily due to light absorbed by chlorophyll *a*; at 650 nm, to light absorbed by chlorophyll *b*; and at shorter wavelengths, to light absorbed by chlorophylls *a* and *b* and by carotenoid pigments, including β-carotene.

Quantum mechanics established that light, a form of electromagnetic radiation, has properties of both waves and particles. When light interacts with matter, it behaves as discrete packets of energy (quanta) called *photons*. The energy of a photon is proportional to the frequency of the light wave and thus inversely proportional to its wavelength. Thus photons of *shorter* wavelengths have *higher* energies. The energy of visible light is considerable. Light with a wavelength of $550\,\mathrm{nm}$ $(550 \times 10^{-7}\,\mathrm{cm})$, typical of sunlight, has about 52 kcal of energy per mole of photons. This is enough energy to synthesize several moles of ATP from ADP and P_i if all the energy were used for this purpose. However, in general in the electron-transport chain shown in stages 1–3 in Figure 12-34, the energy carried in 14 electrons derived from absorption of 28 photons (14 via photosystem II and 14 via photosystem I) is used to generate about 9 molecules of ATP and convert about 7 molecules of $NADP^+$ to NADPH.

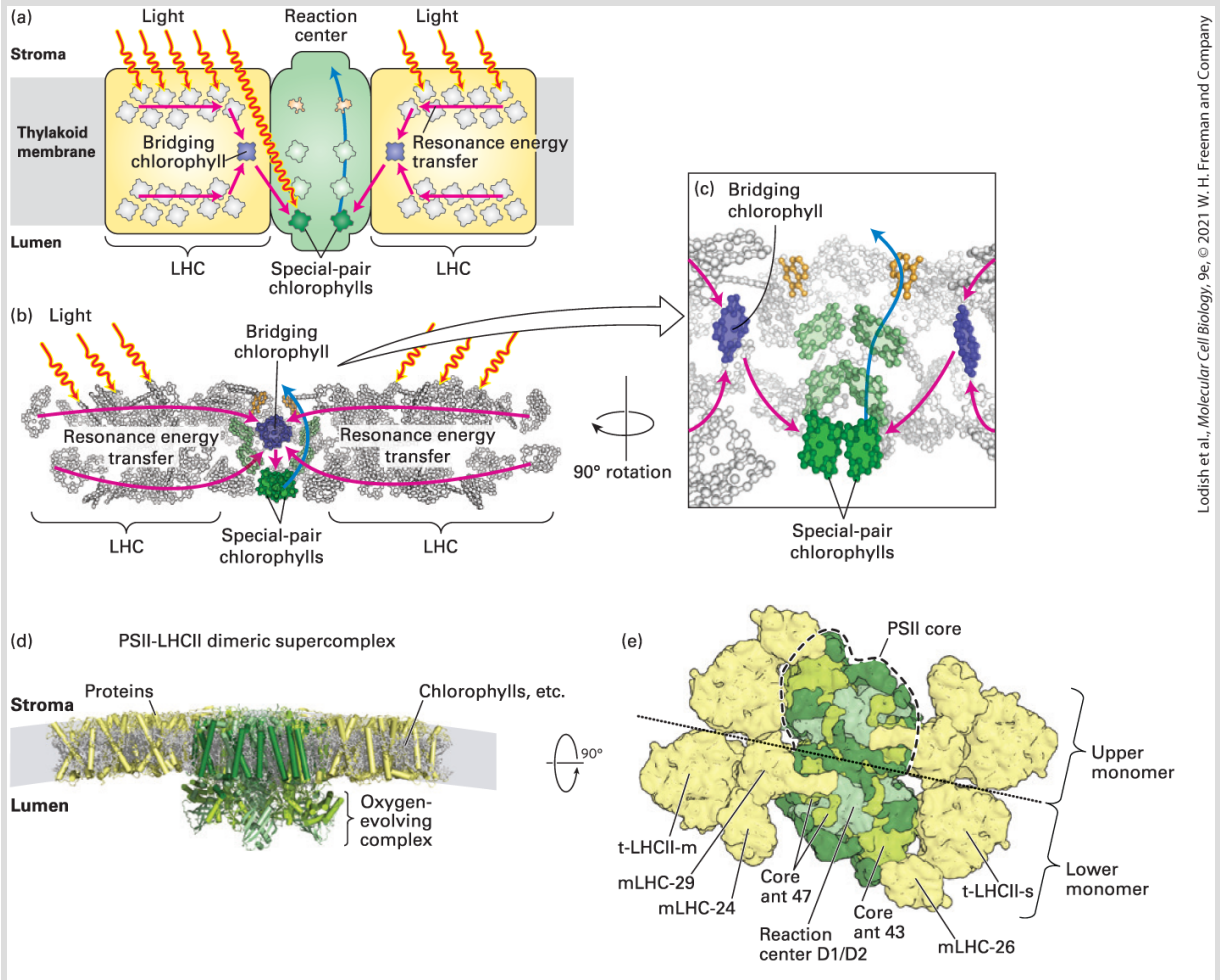
When chlorophyll *a* (or any other molecule) absorbs visible light, the absorbed light energy raises electrons in the chlorophyll *a* to a higher energy (excited) state. This state differs from the ground (unexcited) state largely in the distribution of the electrons around the C and N atoms of the porphyrin ring. Excited states are unstable, and the electrons return to the ground state by one of several competing processes. For chlorophyll *a* molecules dissolved in organic solvents such as ethanol, the principal processes that dissipate the excited-state energy are: (1) the emission of light

(fluorescence and phosphorescence), (2) a process called *intersystem crossing* that leads to the transfer of energy to reactive oxygen species, and (3) thermal emission (heat). However, when the same chlorophyll *a* is bound in the unique protein environment of the reaction center, dissipation of excited-state energy occurs by a different process: the release of a high energy electron. As detailed in the next section, this is the key to photosynthesis.

Core Antenna Complexes and Light-Harvesting Complexes Increase the Efficiency of Photosynthesis

Two chlorophyll *a* molecules of the six chlorophylls within the reaction center are capable of directly absorbing light and initiating photosynthesis by a mechanism we will discuss shortly. In photosynthetic bacteria, the two bacteriochlorophylls that initiate photosynthesis are called the *special-pair bacteriochlorophylls*. Plants also have **special-pair chlorophylls** whose relative positions in the reaction center are similar to those in bacteria. Although in plants the specific pair of reaction-center chlorophylls responsible for initiating photosynthesis is not clear, for simplicity we will refer to the initiating chlorophylls in both plants and bacteria as the special-pair chlorophylls.

While the special-pair chlorophylls can directly absorb light to initiate photosynthesis, they are most commonly energized indirectly by energy transferred to them from other light-absorbing and energy-transferring pigments. These other pigments, which include many other chlorophylls, absorb photons and pass the energy to the special-pair chlorophylls (Figure 12-37). Some of these pigments are bound to protein subunits that are considered to be intrinsic components of the core photosystem, which is made up of several distinct protein chains. These pigment-containing proteins are called *core antenna complexes*. Other pigments are incorporated into protein complexes that bind to, but are distinct from, the core photosystems and are called *light-harvesting complexes (LHCs)*. The topological relationships of the reaction centers, core antennas and LHCs in the dimeric PSII-LHC supercomplex from the garden pea are illustrated in Figure 12-27d and e. Even at the maximum light intensity encountered by photosynthetic organisms (tropical noontime sunlight), each reaction-center chlorophyll *a* molecule directly absorbs only about one photon per second, which is not enough to support photosynthesis sufficient for the needs of the plant. The involvement of core antenna complexes and LHCs greatly increases the efficiency of photosynthesis, especially at more typical light intensities, by increasing absorption of 680-nm light and by extending the range of wavelengths of light that can be absorbed using other antenna pigments. For example, there are approximately 150 antenna chlorophyll molecules for each reaction center in plant PSII.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 12-37 Light-harvesting complexes and photosystems in cyanobacteria and plants. (a) Diagram of the membrane of a cyanobacterium, in which each multiprotein light-harvesting complex (LHC, light yellow) contains 90 chlorophyll molecules (gray) and 31 other small molecules, all held in a specific geometric arrangement for optimal light absorption and energy transfer. Of the six chlorophyll molecules in the reaction center (dark and light green), two constitute the special-pair chlorophylls that initiate photoelectron transport (light blue arrow) when excited. The portions of the two reaction center quinone molecules are orange. Resonance energy transfer (pink arrows) rapidly funnels energy from absorbed light (wavy red arrows) to one of two "bridging" chlorophylls (blue) and thence to the special-pair chlorophylls (dark green) in the reaction center. (b) Three-dimensional organization of photosystem I (PSI) and its associated LHCs from *Pisum sativum* (garden pea), as determined by x-ray crystallography, seen from the plane of the membrane. Only the LHC chlorophylls (white with gray outlines) and the reaction-center electron carriers (same colors as in panel a) are shown. For simplicity, the special-pair chlorophylls are shown to initiate photoelectron transport (light blue arrow), although in plants the specific pair of

chlorophylls responsible is not clear. (c) Expanded view of the reaction center from (b), rotated 90° about a vertical axis. (d and e) Three-dimensional organization of the dimeric supercomplex of photosystem II (PSII) and its associated LHCs from *Pisum sativum* (garden pea), as determined by cryo-electron microscopy, seen from the plane of the membrane (panel d) or rotated 90° and viewed from the stroma (panel e). The 28 polypeptide chains in each monomeric supercomplex are colored based on their functions and locations: reaction center proteins D1 and D2 (pale green), internal antenna complexes - Core ant 47 and Core ant 43 - (lime green), and various LHC antenna proteins (light yellow). In addition to the reaction center and internal antennas, the PSII core contains 12 small proteins (dark green) and the oxygen-evolving complex on the luminal surface (shades of green, panel d). The 157 chlorophylls, 2 pheophytins (chlorophylls without the central magnesium), 44 carotenoids and other attached small molecules are white and gray. Two of the six electron-transporting cholorphylls present in the reaction center of PSI (panel a) are replaced in PSII by two pheophytins (not shown). The dimeric supercomplex comprises almost 100,000 atoms of which 73 percent compose the proteins. In panel e, the upper and lower monomers in the dimeric supercomplexe are separated by a black dotted line. The core PSII complex (reaction center, internal antennas and small associated proteins) is outlined in the upper monomer by a curved dashed line. In the lower monomer individual proteins are labeled: Reaction center D1/D2; core antenna complexes Core ant47 and Core ant43; and external antenna complexes, including the trimeric LHCII complexes t-LHCII-m and tLHCII-s and the monomeric antenna complexes mLHC-24, mLHC-26 and mLHC-29. See W. Kühlbrandt, 2001, *Nature* **411**:896, and P. Jordan et al., 2001, *Nature* **411**:909.

[Parts (b) and (c) Data from A. Ben-Sham et al., 2003, *Nature* **426**:630, PDB ID 1qvz; and Y. Mazor, A. Borovikova, and N. Nelson, 2015, *eLife* **4**:e07433, PDB ID 4y28; part (d) Data from: X. Su et al., 2017, *Science* **357**:815–820 and PDB ID 5XNL.]

Photosystem core proteins and LHC proteins orient the pigment molecules in the precise orientations and positions that are optimal for light absorption and rapid (~ 0.2–200×10^{-12} seconds) energy transfer, called *resonance energy transfer*, to one of the special-pair chlorophyll *a* molecules in the associated reaction center. Resonance energy transfer does not involve the transfer of an

electron. Studies on one of the two photosystems in cyanobacteria, which are similar to those in multicellular, seed-bearing plants, suggest that energy from absorbed light is funneled first to a "bridging" chlorophyll in each LHC and then to the special-pair chlorophylls in the reaction center (see Figure 12-37a). Strikingly, the molecular structures of LHCs from plants and cyanobacteria are completely different from those from green and purple bacteria, even though both types contain carotenoids and chlorophylls. Figure 12-37b shows the distribution of the chlorophyll pigments in photosystem I from *Pisum sativum* (garden pea) together with those from peripheral LHC antennas. The large number of core and LHC antenna chlorophylls that surround the reaction center permit efficient transfer of absorbed light energy to the special-pair chlorophylls in the reaction center. Each chloroplast reaction center is associated with multiple internal antennas and light-harvesting complexes; the LHCs associated with PSII (e.g., LHCII) and with PSI (e.g., LHCI) contain different proteins. Cryoelectron microscopy has been used to determine the structure of a dimeric supercomplex of photosystem II and LHCII from the garden pea (Figure 12-37d and e). The structure reveals the relative positions of the proteins in its reaction center (pale green), internal antenna complexes (lime green), LHC antennas (light yellow) and other PSII core proteins (dark green), viewed both in the plane of the membrane (panel d) and from the stroma (panel e). Depending on light conditions, LHCIIs can dissociate from PSII and bind to PSI. Although LHC antenna chlorophylls can transfer light energy absorbed from a

photon, they cannot release an electron. As we've seen already, this function resides in the reaction-center chlorophylls.

Multiple Mechanisms Protect Cells Against Damage from Reactive Oxygen Species During Photoelectron Transport

As we saw earlier in the case of mitochondria, reactive oxygen species (ROS) generated during electron transport (see Figure 12-24) can both serve as signals to regulate organelle function and cause damage to a variety of biomolecules. The same is true of chloroplasts. For example, hydrogen peroxide (H_2O_2) regulates gas-exchange pores (stomata) in plants subjected to drought stress to prevent dehydration and controls cyclic electron flow.

Even though the PSI and PSII photosystems, with their associated light-harvesting complexes, are remarkably efficient at converting radiant energy to useful chemical energy in the form of ATP and NADPH — especially under low intensity light conditions — they are not perfect. Depending on the intensity of the light and the physiological conditions of the cells, a relatively small — but significant — amount of the energy absorbed by chlorophyll in the light-harvesting antennas and reaction centers results in the chlorophyll being converted to an activated state called *triplet*

chlorophyll. In this state, the chlorophyll can transfer some of its energy to molecular oxygen (O_2), converting it from its normal, *relatively* unreactive ground state, called *triplet oxygen* ($^3O_2$), to a very highly reactive (ROS) singlet state, $^1O_2$. Some of this $^1O_2$ can be used for signaling to the nucleus to communicate the metabolic state of the chloroplast to the rest of the cell. However if the majority of the $^1O_2$ is not quickly quenched (inactivated), it will react with, and usually damage, nearby molecules. This damage, called *photoinhibition*, can suppress the efficiency of thylakoid activity.

Chloroplasts use two major strategies for what is called *non-photochemical quenching* to avoid excess-light-induced damage. The first strategy for non-photochemical quenching is the use of specialized small organic *scavenger molecules* that quench the $^1O_2$ and are incorporated within photosystem/LHC supercomplexes. Examples include carotenoids (polymers of unsaturated isoprene groups, including β-carotene, which gives carrots their orange color) and α-tocopherol (a form of vitamin E). The second strategy involves dissipating the initially absorbed light energy as heat before it can generate damaging $^1O_2$. The extent of energy dissipation is regulated, depending on the light conditions. When exposed to intense light and thus the greatest risk of $^1O_2$-induced damage, the non-photochemical quenching is rapidly induced, and when light intensity falls, quenching is suppressed. The mechanisms underlying the second strategy include enzymatic changes in the structures of certain carotenoids and rapid responses to PSII-

dependent increases in the proton concentration in the thylakoid lumen.

KEY CONCEPTS OF SECTION 12.8

Chloroplasts and Photosynthesis

- Chloroplasts capture the energy of sunlight, convert it into chemical energy in the form of ATP and NADPH, and then use this energy to make complex carbohydrates out of carbon dioxide and water.
- The principal end products of photosynthesis in plants are molecular oxygen and polymers of six-carbon sugars (starch and sucrose).
- In chloroplasts, the light-capturing and ATP- and NADPH- generating reactions of photosynthesis occur in the thylakoid membrane. The permeable outer membrane and the inner membrane surrounding the thylakoid membrane in chloroplasts do not participate directly in photosynthesis (see Figures 12-33 and 12-34).
- Like mitochondria, chloroplasts contain their own DNA and are thought to have evolved from an ancestral endosymbiotic photosynthetic bacterium.
- Plants contain two light-absorbing photosystems, PSI and PSII, which have different functions and are physically separated in the thylakoid membrane. PSII generates high-energy electrons that convert the quinone Q to QH_2 and convert H_2O into O_2, and PSI reduces $NADP^+$ to NADPH. Cyanobacteria have two analogous photosystems.
- There are four stages in photosynthesis: (1) absorption of light, generation of high-energy electrons, and formation of O_2 from H_2O; (2) electron transport leading to generation of a proton-motive force, absorption of light to reenergize electrons, and reduction of $NADP^+$ to NADPH; (3) synthesis of ATP; and (4) conversion of CO_2 into carbohydrates (carbon fixation). The overall reaction of oxygen-generating photosynthesis in plants is $6\,CO_2 + 6\,H_2O \rightarrow 6\,O_2 + C_6H_{12}O_6$ (see Figure 12-34).
- In photosystems (PSI and PSII), light energy is transferred directly or indirectly to one of two special-pair chlorophyll *a* molecules bound to reaction-center proteins in the thylakoid membrane. This energy is used to generate a high-energy electron that separates from a positively charged chlorophyll *a* by a process called *photoelectron transport*. The high-energy electrons drive the generation of NADPH and ATP.
- In photosystems, the reaction centers are bound to multiple internal antennas and external antennas that contain chlorophylls *a* and *b*, carotenoids, and other pigments that absorb light at multiple wavelengths. The external antennas are called *light-*

harvesting complexes (LHCs). Depending on light conditions, LHCIIs can dissociate from PSII to which they often bind and bind to PSI. Energy, but not electrons, is transferred from the pigments in the internal antennas and the LHCs to the special-pair chlorophylls in the reaction center by resonance energy transfer to increase the capacity to convert light to high-energy electrons for photosynthesis (see Figure 12-37).

- The absorption of light by pigments in the chloroplast can generate reactive oxygen species (ROS), including singlet oxygen, $^1O_2$, and hydrogen peroxide, H_2O_2. In small amounts, they are used as intracellular signaling molecules; in larger amounts, they can be toxic.
- Chloroplasts employ multiple strategies for preventing the buildup of damaging $^1O_2$. For example, small-molecule scavengers, such as carotenoids, and antioxidant enzymes help to protect against ROS-induced damage. Chloroplasts also use regulated non-photochemical quenching mechanisms to protect against light-induced damage.

12.9 Use of Light Energy to Generate Molecular Oxygen, NADPH, and ATP in Stages 1–3 of Photosynthesis

In the previous section, we presented an overview of the four stages of photosynthesis. In this section, we will describe some of the molecular details of how in stages 1 and 2 light energy is used to generate O_2 from H_2O, generate NADPH from $NADP^+$, and produce a proton-motive force, and how in stage 3 that proton-motive force drives the synthesis of ATP from ADP and P_i (see Figure 12-34).

The First Three Stages of Photosynthesis

Stage 1: Absorption of Light Energy by PSII, Generation of High-Energy Electrons, and Formation of O_2 from Water

Within the reaction center of PSII, two adjacent special-pair chlorophylls lie close to the luminal face of the thylakoid membrane (Figure 12-38). When a photon of light with a wavelength of about 680 nm is absorbed by one of these two chlorophyll *a* molecules in

purple bacteria, the energy of that chlorophyll *a* molecule increases by 42 kcal/mol (the first excited state). This energized molecule rapidly donates an electron to the adjacent chlorophyll, which passes it on to a series of intermediate small-molecule acceptors bound to the PSII. These transfers are represented by the zigzag blue lines in Figure 12-34 and in Figure 12-38 (*left*). The electron is rapidly passed on to the primary electron acceptor, which is the oxidized form of a molecule called *quinone (Q)*, near the stromal surface of the thylakoid membrane (Figures 12-34 and 12-38). Q is similar in structure and function to CoQ in mitochondria, in that it is a small, lipid-soluble, organic molecule that serves as an electron carrier that undergoes oxidation to Q and reduction to QH_2. This light-driven electron transfer, called **photoelectron transport**, is very efficient and depends on the environment of both the chlorophylls and the acceptor within the reaction center. Photoelectron transport, which occurs nearly every time a photon is absorbed, leaves a positive charge on the chlorophyll *a* (chlorophyll a^+) close to the luminal surface of the thylakoid membrane (on the opposite side from the stroma) and generates a high-energy electron that moves to Q to form the negatively charged acceptor (Q^-) near the stromal surface.

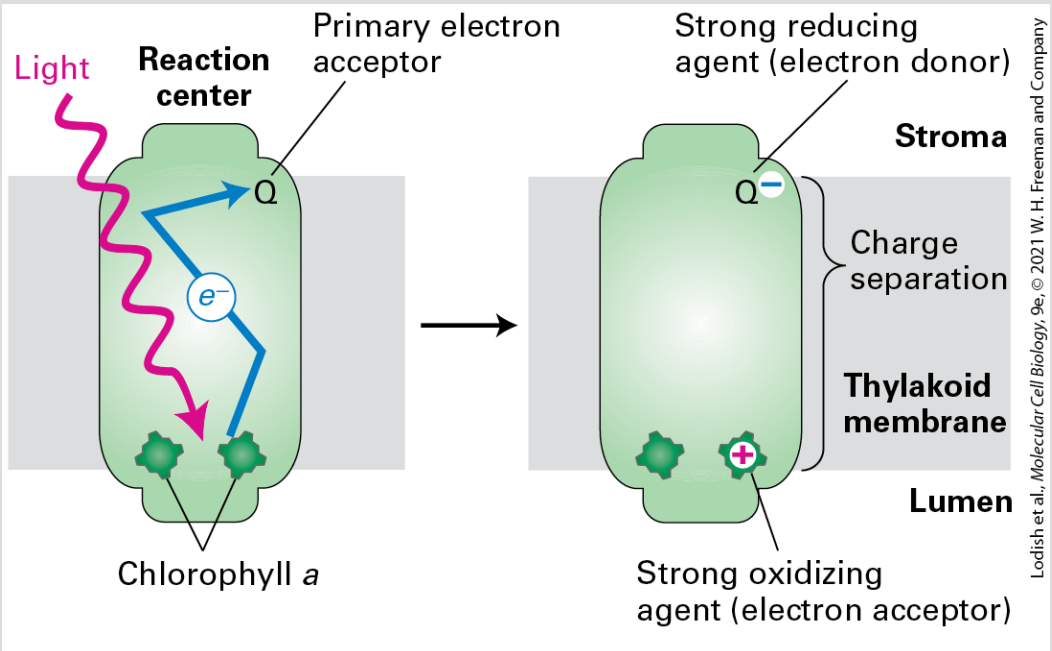


FIGURE 12-38 Photoelectron transport in a photosystem, the primary event in photosynthesis. After absorption of a photon of light, one of the excited special-pair chlorophyll *a* molecules in the reaction center (*left*) donates, via several intermediates (not shown), an electron to a loosely bound acceptor molecule, the quinone Q, on the stromal surface of the thylakoid membrane, creating an essentially irreversible charge separation across the membrane (*right*). Subsequent transfers of this electron release energy that is used to generate a proton gradient and thus ATP and NADPH (see Figures 12-34). The positively charged chlorophyll a^+ generated when the light-excited electron moves to Q is eventually neutralized by the transfer to the chlorophyll a^+ of another electron. In plants, the oxidation of H_2O to O_2 provides this neutralizing electron and takes place in PSII (see Figure 12-34). Photosystem I uses a similar photoelectron-transport pathway, but as we will see below, instead of oxidizing water, it receives an electron from a protein carrier called *plastocyanin* to neutralize the positive charge on chlorophyll a^+ (see Figure 12-34).

Absorption of a second photon and transfer of a second electron to Q^- forms a powerful reducing agent that combines with two protons

to form QH_2, the reduced form of the quinone. QH_2 is a source of high-energy electrons and protons for stage 2 described below. The herbicide *atrazine* is one of the most commonly used weed killers in US agriculture. Atrazine binds to PSII, blocks the binding of Q, and thus blocks formation of QH_2 and downstream electron transport.

The positively charged special-pair reaction-center chlorophyll a^+ is the *strongest* biological oxidant known. In plants, the oxidizing power of four chlorophyll a^+ molecules is used by a subcomplex in PSII called the *oxygen-evolving complex* (Mn_4CaO_5, see Figure 12-34) to remove four electrons from two H_2O molecules bound to a site on the luminal surface to form O_2:

$$2\,H_2O + 4\,\text{chlorophyll}\,a^+ \rightarrow 4\,H^+ + O_2 + 4\,\text{chlorophyll}\,a$$

However, absorption of each photon by PSII results in the transfer of just one electron. Thus it is necessary for a single PSII to lose an electron and then partially oxidize the oxygen-evolving complex, and then sequentially repeat this process three more times in order for its oxygen-evolving complex to extract four electrons from two water molecules and generate one O_2 molecule. These potent biological reductants and oxidants along with PSI provide all the energy needed to drive all subsequent reactions of photosynthesis: electron and proton transport and NADPH synthesis (stage 2), ATP synthesis (stage 3), and CO_2 fixation (stage 4).

Chlorophyll *a* can absorb light at discrete wavelengths shorter (and therefore of higher energy) than 680 nm (see Figure 12-36). Such absorption raises the molecule into one of several excited states whose energies are higher than that of the first excited state described above. These higher excited states decay to the lower energy first excited state by releasing energy in $\sim$100–200 $\times$ 10^{-15} seconds (~100–200 femtoseconds, fs) as heat. Because photoelectron transport and the resulting charge separation occur only from the first excited state of the reaction-center chlorophyll *a*, the quantum yield — the amount of photosynthesis per absorbed photon — is the same for all wavelengths of visible light shorter than 680 nm. How closely the wavelength of light matches the absorption spectrum of the pigment determines how likely it is that the photon will be absorbed. Once absorbed, the photon's exact wavelength is not critical, provided it is at least energetic enough (has a short enough wavelength) to push the chlorophyll *a* into the first excited state.

In chloroplasts, the special-pair reaction-center chlorophylls that initiate photoelectron transport in PSII and in PSI differ in their light-absorption maxima (680 and 700 nm, respectively) because of differences in their protein environments. For this reason, these chlorophylls are often denoted P_{680} (PSII) and P_{700} (PSI). Furthermore, the two photosystems are distributed differently in thylakoid membranes: PSII primarily in regions of stacked thylakoids (grana, see Figure 12-33) and PSI primarily in unstacked

regions. The stacking of thylakoid membranes may be due to the binding properties of the proteins associated with PSII, especially light-harvesting complexes.

Stage 2: Generation of a Proton-Motive Force and NADPH

Electrons move from the quinone primary electron acceptor QH_2 generated in stage 1 to a proton pump, called the *cytochrome* bf *complex*, in stage 2 (Figure 12-34). The high energy electrons in QH_2 provide the energy for the cytochrome *bf* complex to drive across the thylakoid membrane protons from QH_2 as well as protons from the stroma to generate a proton-motive force, forming a pH gradient across the membrane ($pH_{lumen} < pH_{stroma}$). This process is analogous to the generation of a proton-motive force across the inner mitochondrial membrane and in bacterial membranes during electron transport (see Figure 12-25). A small, water-soluble, copper-containing, electron carrier protein called *plastocyanin* transfers the now lower energy electrons from the cytochrome *bf* complex to a complex called *photosystem I (PSI)* (see Figure 12-34). The electrons from QH_2 lose considerable energy by driving proton pumping and thus have relatively low energy when they are transferred to PSI from plastocyanin. After additional energy is transferred to those electrons by absorption of light in photosystem I (PSI) by special-pair chlorophylls, they move to two protein components of PSI, first to ferredoxin (FN) and then to ferredoxin-$NADP^+$ reductase (FNR),

which use the energy in these electrons to drive the synthesis of the high-energy electron carrier NADPH from its oxidized form, **nicotinamide adenine dinucleotide phosphate ($NADP^+$)**. The structure of $NADP^+$ is identical to that of NAD^+ except for the presence of an additional phosphate group. Both molecules gain and lose electrons in the same way (see Figure 2-33).

Thus the overall reaction of stages 1 and 2 can be summarized as

$$2H_2O + 2NADP^+ \xrightarrow{\text{Light}} 2H^+ + 2NADPH + O_2$$

Generation of a Proton-Motive Force, but No NADPH or O_2, in Stage 2 by Cyclic Electron Flow

Stage 2 generates a proton gradient (used to make ATP in stage 3) and NADPH, and the pathway of stage 2 shown in Figure 12-34 has been called the *linear electron flow* pathway. In some circumstances, however, especially when plants are stressed by conditions such as drought, high light intensity, or low carbon dioxide levels, cells must generate greater amounts of ATP relative to NADPH than they can produce by linear electron flow. To do this, they photosynthetically produce ATP from PSI without concomitant NADPH production. This is accomplished by a PSI-dependent and PSII-independent process called *cyclic photophosphorylation*, or *cyclic electron flow* (Figure 12-39). In this process, electrons cycle among PSI, ferredoxin, plastoquinone (Q), and the cytochrome *bf* complex,

bypassing the ferredoxin-$NADP^+$ reductase at PSI that normally generates NADPH. Thus during cyclic electron flow, proton pumping permits additional ATP synthesis, but no *net* NADPH is generated, and there is no oxidization of H_2O to produce O_2.

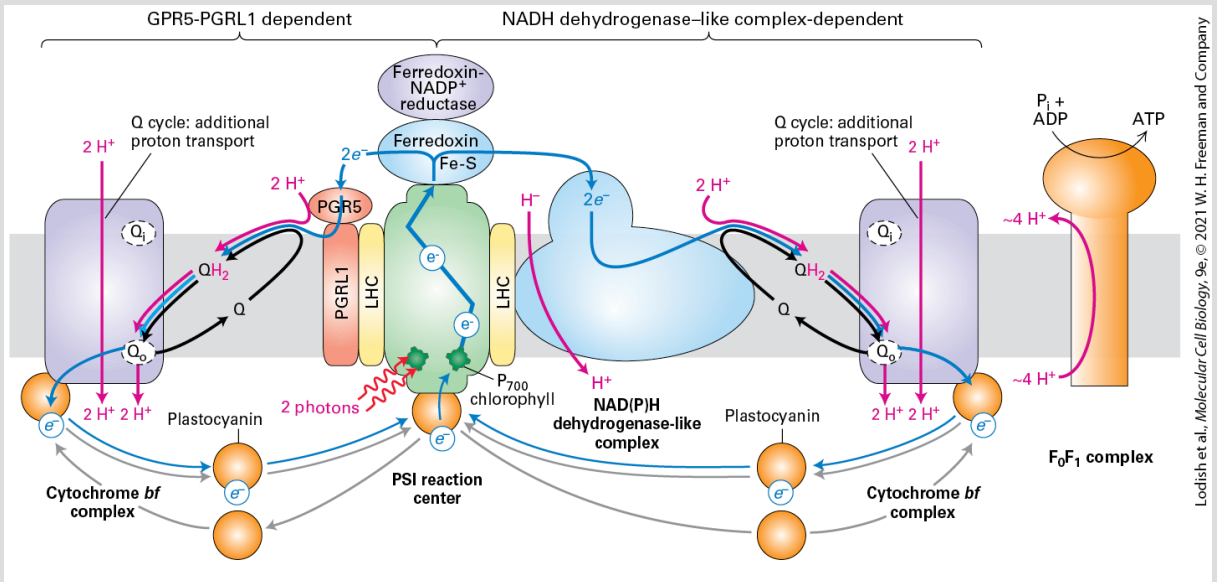


FIGURE 12-39 Cyclic electron flow in plants, which generates a proton-motive force and ATP but no oxygen or NADPH. In cyclic electron flow, light energy is used by PSI to transport electrons in a cycle to generate a proton-motive force and subsequently in stage 3 ATP, without oxidizing water or generating NADPH. High-energy electrons are transferred via the ferredoxin of PSI either to a PGR5-PGRL1 heterodimer (red, *left* pathway) or to the NADH dehydrogenase–like complex (blue, *right* pathway), where they then reduce plastoquinone (Q) to QH_2. Each of these two electron acceptors forms independent supercomplexes with PSI via light-harvesting complex (LHC) subunits (yellow). (The PGR5-PGRL1 heterodimer and NADH dehydrogenase–like complex are not found together in the same supercomplex, but here are drawn together with only one PSI to emphasize the similarities of the two mechanisms of cyclic electron flow.) QH_2 then transfers the electrons to the cytochrome *bf* complex, then to plastocyanin, and finally back to PSI, as is the case for the linear electron flow pathway (see Figure 12-34).

In higher plants, there are two cyclic electron flow pathways that control the ATP:NADPH ratio (see Figure 12-39). The major pathway is the PGR5-PGRL1-dependent pathway, which we shall describe shortly; this pathway ensures efficient photosynthesis and protects against stress. The minor pathway is the *NADH dehydrogenase–like complex-dependent pathway*, which appears to respond to stress and to be a target of H_2O_2-mediated regulation. The NADH dehydrogenase–like complex is a very large multiprotein complex that is very similar in shape and composition to mitochondrial complex I (see Figure 12-19), which oxidizes NADPH or NADH while reducing CoQ to $CoQH_2$. The NADH dehydrogenase–like complex, however, appears to lack the subunit necessary for NADPH oxidation.

During cyclic electron flow, high-energy electrons generated by light absorption and photoelectron transport in PSI are transferred either to the PGR5-PGRL1 heterodimer or to the NADH dehydrogenase–like complex from the ferredoxin subunit of PSI. Indeed, there is evidence that each of these two electron acceptors independently associates with PSI in supercomplexes mediated by LHC subunits. Both of these electron acceptors then reduce Q to QH_2, which then delivers protons and electrons to the cytochrome *bf* complex via a Q cycle, as we described earlier for linear electron flow (see Figure 12-34). Protons are transported across the thylakoid membrane into the lumen by the cytochrome *bf* complex and possibly by the NADH dehydrogenase–like complex. Finally, plastocyanin returns the

electrons from the cytochrome *bf* complex to PSI to complete the cycle. The proton-motive force generated by cyclic electron flow drives ATP synthesis by the F_0F_1 complex (ATP synthase) and thus increases the ATP:NADPH ratio.

Stage 3: Synthesis of ATP

The light-driven proton gradient generated across the thylakoid membrane in stages 1 and 2 can then be used to drive ATP synthesis (see Figure 12-34). The protons move down their concentration gradient from the thylakoid lumen to the stroma through the chloroplast F_0F_1 complex (ATP synthase), which couples proton movement to the synthesis of ATP from ADP and P_i, as we have seen for the ATP synthases in mitochondria and bacteria (see Figures 12-27, 12-28, and 12-33).

Relative Activities of Photosystems I and II Are Regulated

Chloroplasts respond to changes in the wavelengths and intensities of ambient light (as a consequence of the time of day, cloudiness, etc.) by altering the relative outputs of PSI and PSII to maintain the appropriate balance of ATP and NADPH production.

One mechanism underlying the regulation of the relative activities of PSI (preferentially located in the unstacked thylakoid membranes) and PSII (preferentially located in the stacked grana) is phosphorylation and dephosphorylation of thylakoid membrane proteins, including PSII and LHCII. Changes in phosphorylation, particularly for LHCII, can alter the intra–thylakoid membrane distribution (grana vs. unstacked membranes) of this antenna complex and thus its differential interactions with PSI and PSII. The more LHCII is associated with a particular photosystem, the more efficiently that system will be activated by light, and the greater its contribution to electron flow. It appears that the efficient operation of electron-transport chains has involved the evolution of functional complexes of increasing size and complexity, from individual proteins to complexes to supercomplexes to super-supercomplexes. Regulating the supramolecular organization of the photosystems in plants has the effect of directing them toward primarily ATP production alone or toward the generation of reducing equivalents (NADPH) and ATP, depending on ambient light conditions and the metabolic needs of the plant. Both NADPH and ATP are required to convert CO_2 to sucrose or starch, the fourth stage in photosynthesis, which we cover in the next section.

KEY CONCEPTS OF SECTION 12.9

Use of Light Energy to Generate Molecular Oxygen, NADPH, and ATP in Stages 1–3 of Photosynthesis

- In stage 1 of photosynthesis (see Figure 12-34), light energy is converted by photosystem II (PSII) into high energy electrons that are transferred by photoelectron transport from the special-pair chlorophylls, via intermediates, to a quinone (Q) on the opposite side of the membrane, generating a charge separation (see Figure 12-38) and the small lipid-soluble electron transport molecule QH_2.
- In green plants, the positively charged chlorophyll created by photoelectron transport in PSII is regenerated to the neutral, reduced form by the oxygen-evolving complex subunit of PSII. Electrons from water are transferred by the oxygen-evolving complex to the positively charged chlorophyll of PSII, resulting in the formation of molecular oxygen (O_2) in the thylakoid lumen.
- In stage 2 (see Figure 12-34), high-energy electrons are transported from the reduced quinone QH_2 from stage 1 to a proton pump (cytochrome *bf* complex) that moves across the membrane protons from both QH_2 and the stroma to the thylakoid lumen, forming a pH gradient (proton-motive force) across the thylakoid membrane.
- The electrons depleted of energy by the cytochrome *bf* complex during proton pumping are transferred by the water-soluble protein plastocyanin to PSI, where they reenergize when light energy is absorbed by its special-pair chlorophylls.
- Photoelectron transport within PSI transfers the reenergized electrons to the ultimate electron acceptor, usually $NADP^+$, reducing it to NADPH.
- The relative activities of PSI and PSII are regulated to maintain the appropriate balance of ATP and NADPH production.
- Reversible phosphorylation and dephosphorylation of the light-harvesting complex II (LHCII) control the functional organization of the photosynthetic apparatus in thylakoid membranes to regulate the relative activities of PSI and PSII and thus the relative amounts of ATP and NADPH generated. Very large super-supercomplexes appear to be involved in this regulation.
- In stage 3, movement of protons down their electrochemical gradient generated in stage 2 (proton-motive force) through F_0F_1 complexes (ATP synthase) powers the synthesis of ATP from ADP and P_i.

12.10 ATP and NADPH Drive Carbon Fixation in the Calvin Cycle and Carbohydrate Synthesis in Stage 4 of Photosynthesis

In stage 4, the NADPH and ATP generated in stages 2 and 3 drive the synthesis from CO_2 and H_2O of six-carbon sugars (and their polymers). The overall chemical equation is

$$\begin{aligned} &6\,CO_2 + 18\,ATP^{4-} + 12\,NADPH + 12\,H_2O \rightarrow \\ &\qquad C_6H_{12}O_6 + 18\,ADP^{3-} + 18\,P_i^{2-} + 12\,NADP^+ + 6\,H^+ \end{aligned}$$

In addition to CO_2 fixation — incorporation of gaseous CO_2 into small organic molecules and then sugars — the synthesis of almost all amino acids, all fatty acids and carotenes, all pyrimidines, and probably all purines occurs in chloroplasts. The synthesis of sugars from CO_2 is the most extensively studied biosynthetic pathway in plant cells. We first consider the unique pathway, known as the **Calvin cycle** (after discoverer Melvin Calvin), that fixes CO_2 into three-carbon compounds, powered by energy released during ATP hydrolysis and oxidation of NADPH. Because CO_2 is initially incorporated into a three-carbon compound, the Calvin cycle is also called the C_3 *pathway* of carbon fixation.

Rubisco Fixes CO_2 in the Chloroplast Stroma

In the first step of the Calvin cycle (Figure 12-40, *top*), the enzyme **ribulose 1,5-bisphosphate carboxylase**, or **rubisco**, incorporates CO_2 into precursor molecules that are subsequently converted into carbohydrates. Rubisco is located in the stroma of the chloroplast. This enzyme adds CO_2 to the five-carbon sugar ribulose 1,5-bisphosphate to form two molecules of the three-carbon compound 3-phosphoglycerate (Figure 12-41). Rubisco is a large enzyme ($\sim 520\,\mathrm{kDa}$) whose most common form is composed of eight identical large and eight identical small subunits. The large subunit, which contains the catalytic site, is encoded in chloroplast DNA and the small subunit is encoded by a family of genes in the nucleus. Because the catalytic rate of rubisco is quite low, many copies of the enzyme are needed to fix sufficient CO_2. Indeed, this enzyme makes up almost 50 percent of the soluble protein in a chloroplast and is believed to be the most abundant protein on Earth. It is estimated that rubisco fixes more than 10^{11} tons of atmospheric CO_2 each year.

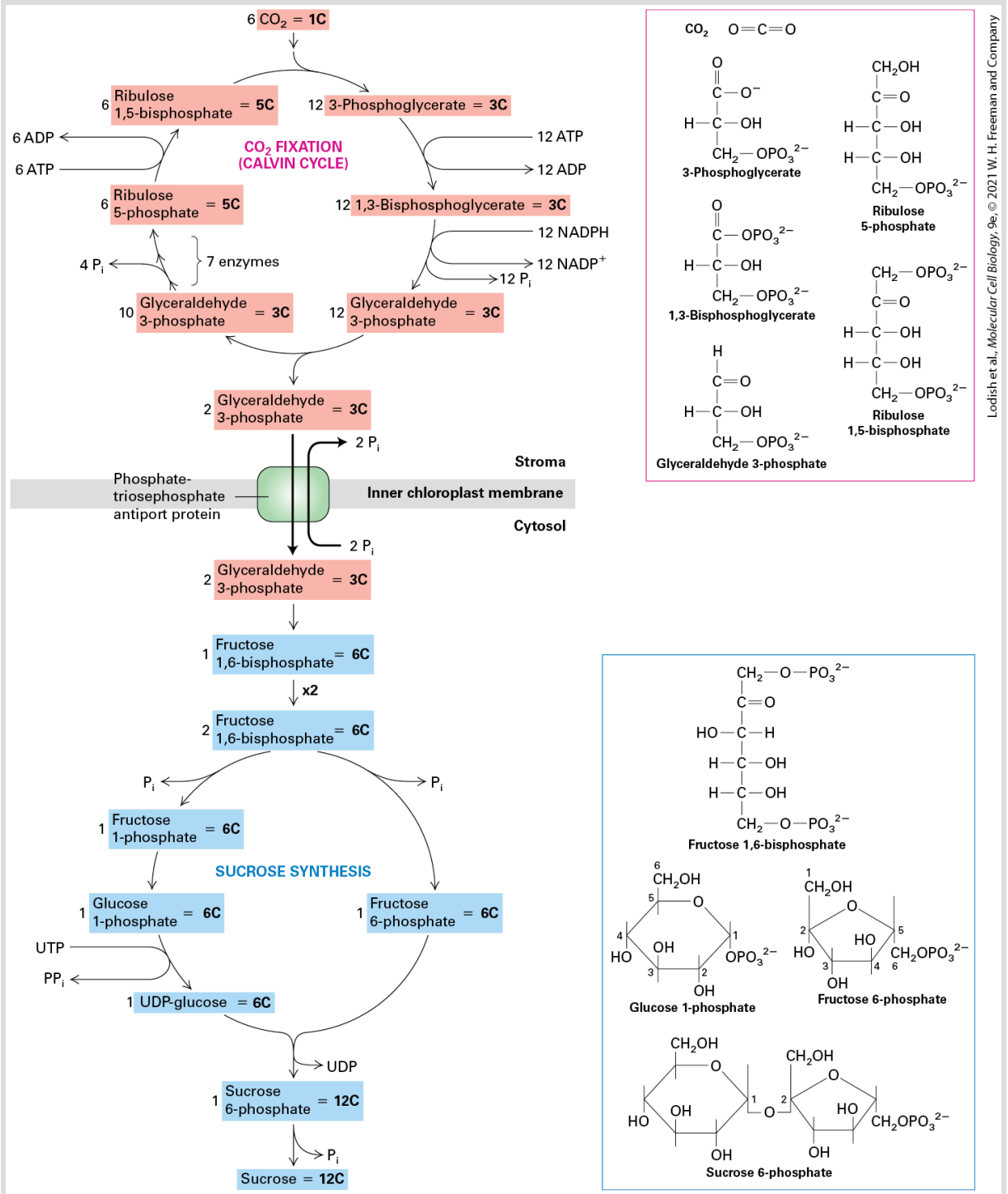


FIGURE 12-40 The pathway of carbon during photosynthesis. *Top*: Six molecules of CO_2 are converted into two molecules of glyceraldehyde 3-phosphate. These reactions, which constitute the Calvin cycle, occur in the stroma of the chloroplast. Via the phosphate/triosephosphate antiporter, some glyceraldehyde 3-phosphate is transported

to the cytosol in exchange for phosphate. No fixed CO_2 leaves the chloroplast unless phosphate is fed into it to replace the phosphate carried out of the stroma in the form of glyceraldehyde 3-phosphate. *Bottom*: In the cytosol, an exergonic series of reactions converts glyceraldehyde 3-phosphate to fructose 1,6-bisphosphate. Two molecules of fructose 1,6-bisphosphate are used to synthesize one molecule of the disaccharide sucrose and release inorganic phosphate that can be transported back into the stroma. The mechanism of the conversion of glyceraldehyde 3-phosphate to hexoses is almost the reverse of that in glycolysis (see Figure 12-3). Some glyceraldehyde 3-phosphate (not shown here) is also converted to amino acids and fats, compounds essential for plant growth.

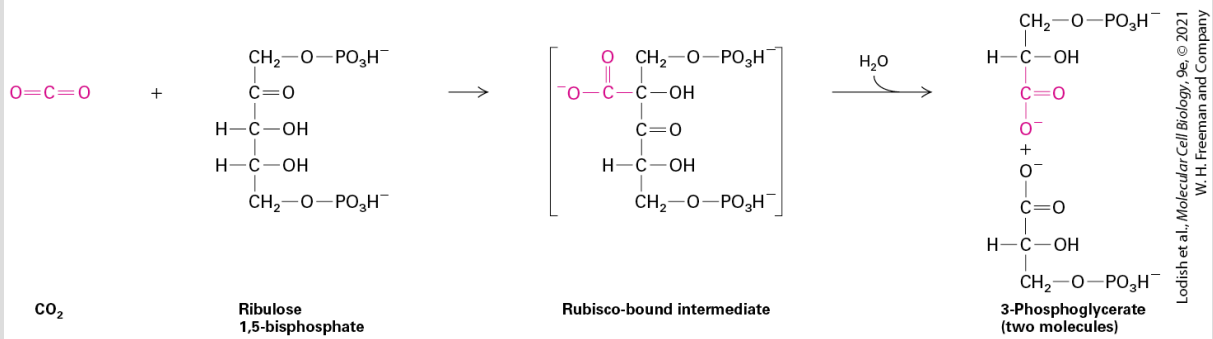
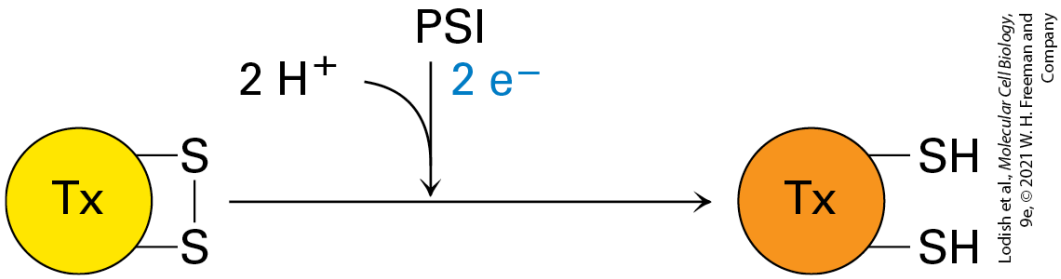


FIGURE 12-41 The initial reaction of rubisco that fixes CO_2 into organic compounds. In this reaction, catalyzed by ribulose 1,5-bisphosphate carboxylase (rubisco), CO_2 (red) condenses with the five-carbon sugar ribulose 1,5-bisphosphate. The products are two molecules of 3-phosphoglycerate.

Rubisco is a light- and redox-sensitive enzyme, although its regulation is very complex and not yet fully understood. Rubisco is spontaneously activated in the presence of high CO_2 and Mg^{2+} concentrations. The activating reaction entails the covalent addition of CO_2 to the side-chain amino group of a lysine in the enzyme's active site, forming a carbamate group that then binds a Mg^{2+} ion, which is required for enzymatic activity. Under normal conditions,

however, with ambient levels of CO_2, this reaction is slow and usually requires catalysis by *rubisco activase*, a member of the AAA^+ family of ATPases. Rubisco activase hydrolyzes ATP and uses the energy released to clear the active site of rubisco of small molecule inhibitors so that CO_2 can be added to its active site lysine. Rubisco activase also accelerates an activating conformational change in rubisco (from an inactive-closed to an active-opened state). The regulation of rubisco activase by the protein *thioredoxin (Tx)* is, at least in part in some species, responsible for rubisco's light/redox sensitivity. In the dark, thioredoxin contains a disulfide bond; in the light, electrons are transferred from PSI, via ferredoxin, to thioredoxin, reducing its disulfide bond:



In the dark, when thioredoxin becomes reoxidized, rubisco activase and other enzymes are reoxidized and so inactivated. Thus the activities of these enzymes are sensitive to the redox state of the stroma, which in turn is light sensitive — an elegant mechanism for the regulation of enzymatic activity by light.

Furthermore, rubisco activase's activity is sensitive to the ratio of ATP to ADP. If that ratio is low (relatively high ADP), then the

activase will not activate rubisco (and so the cell will expend less of its scarce ATP to fix carbon). Photosynthesis is sensitive to a variety of typical plant stresses — moderate heat, cool temperatures, drought (limited water), high salt, high light intensity, and UV radiation. At least some of these stresses influence CO_2 fixation by reducing the activity of rubisco activase and thus rubisco. Inhibition of CO_2 fixation reduces consumption of NADPH. Under strong light conditions, the high $NADPH/NADP^+$ ratio can reduce electron flow to $NADP^+$ and increase leakage to O_2, resulting in increased ROS formation, which can both initiate cellular-signaling pathways and interfere with a variety of cellular processes. Given the key role of rubisco in controlling energy utilization and carbon flux — both in individual chloroplasts and, in a sense, throughout the entire biosphere — it is not surprising that its activity is carefully regulated.

The fate of 3-phosphoglycerate formed by rubisco is complex: some is converted to hexoses incorporated into starch or sucrose, but some is used to regenerate ribulose 1,5-bisphosphate. The fixation of 6 CO_2 molecules and the net formation of 2 glyceraldehyde 3-phosphate molecules (see Figure 12-40, *top*) require the consumption of 18 ATPs and 12 NADPHs generated by the light-requiring processes of photosynthesis.

Synthesis of Sucrose Using Fixed CO_2 Is Completed in the Cytosol

After its formation in the chloroplast stroma, glyceraldehyde 3-phosphate is transported to the cytosol in exchange for phosphate. The final steps of sucrose synthesis (Figure 12-40, *bottom*) occur in the cytosol of leaf cells.

Synthesis of Starch in the Stroma

The key monomer substrate used to build large starch polymers is ADP-glucose. This polymerization takes place in the stroma, and starch polymers are stored there in densely packed crystalline aggregates called *granules* (see Figures 12-33 and 12-34). The enzymes that generate ADP-glucose from glucose 1-phosphate and ATP are found in both the stroma and the cytosol, indicating that hexoses of various structures are imported from the cytosol into the stroma for starch synthesis.

Photorespiration Competes with Carbon Fixation and Is Reduced in C_4 Plants

As noted above, rubisco catalyzes the incorporation of CO_2 into ribulose 1,5-bisphosphate as part of photosynthesis. It can catalyze a second, distinct, and *competing* reaction with the same substrate — ribulose 1,5-bisphosphate — but with O_2 in place of CO_2 as a second substrate, in a process known as **photorespiration** (Figure 12-42).

The products of this second reaction are one molecule of 3-phosphoglycerate and one molecule of the two-carbon compound phosphoglycolate. The carbon-fixing reaction is favored when the ambient CO_2 concentration is relatively high, whereas photorespiration is favored when CO_2 is low and O_2 is relatively high. Photorespiration takes place in light, consumes O_2, and converts ribulose 1,5-bisphosphate in part to CO_2. As Figure 12-42 shows, photorespiration is wasteful to the energy economy of the plant: it consumes ATP and O_2, and it generates CO_2 without fixing carbon. Indeed, when CO_2 is low and O_2 is high, much of the CO_2 fixed by the Calvin cycle is lost as the result of photorespiration. This surprising, wasteful alternative reaction catalyzed by rubisco may be a consequence of the inherent difficulty the enzyme has in specifically binding the relatively featureless CO_2 molecule and of the ability of both CO_2 and O_2 to react and form distinct products with the same initial enzyme/ribulose 1,5-bisphosphate intermediate.

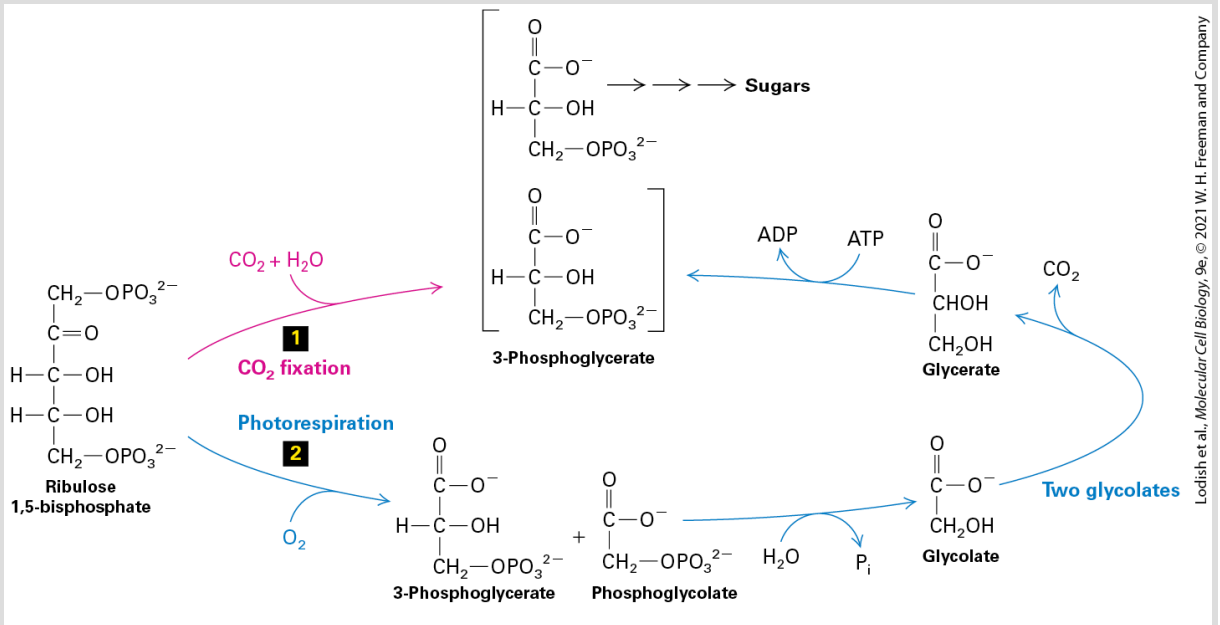


FIGURE 12-42 CO_2 fixation and photorespiration. These competing pathways are both initiated by ribulose 1,5-bisphosphate carboxylase (rubisco), and both use ribulose 1,5-bisphosphate. CO_2 fixation (pathway **1**) is favored by high CO_2 and low O_2 concentrations; photorespiration (pathway **2**) occurs at low CO_2 and high O_2 concentrations (i.e., under normal atmospheric conditions). The phosphoglycolate produced by photorespiration is recycled via a complex set of reactions that take place in peroxisomes and in mitochondria, as well as in chloroplasts. The net result: for every two molecules of phosphoglycolate formed by photorespiration (four C atoms), one molecule of 3-phosphoglycerate is ultimately formed and recycled, and one molecule of CO_2 is lost.

Excessive photorespiration could become a problem for plants in a hot, dry environment because they must keep the gas-exchange pores (stomata) in their leaves closed much of the time to prevent excessive loss of moisture. As a consequence, the CO_2 level inside the leaf can fall below the K_m of rubisco for CO_2. Under these conditions, the rate of photosynthesis is slowed, photorespiration is greatly favored, and the plant might be in danger of fixing inadequate amounts of CO_2. Corn, sugarcane, crabgrass, and other

plants that can grow in hot, dry environments have evolved a way to avoid this problem by using a two-step pathway of CO_2 fixation in which a CO_2-hoarding step in cells called *mesophyll cells* precedes the Calvin cycle in cells called *bundle sheath cells* (Figure 12-43). This pathway has been named the **C_4 pathway** because the first molecules formed in mesophyll cells during carbon fixation in this pathway are four-carbon compounds, such as oxaloacetate and malate, rather than the three-carbon molecules that initiate the Calvin cycle (C_3 pathway). These four-carbon molecules then move into bundle sheath cells where they release CO_2 that serves as a substrate for the Calvin cycle. The relatively high concentration of CO_2 and low concentration of O_2 in bundle sheath cells favor the fixation of CO_2 by rubisco to form 3-phosphoglycerate and suppress photorespiration.

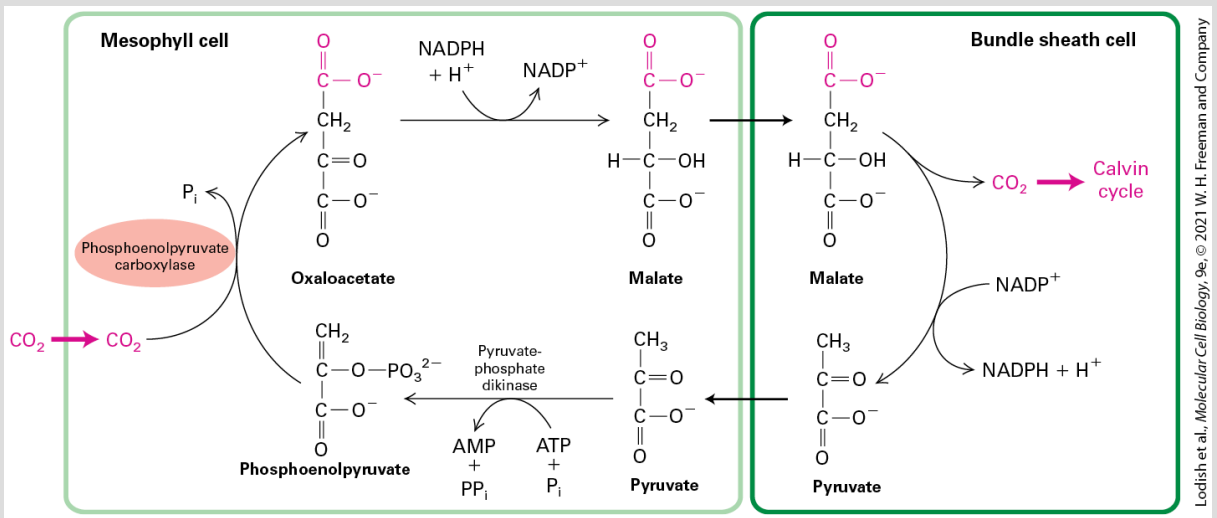


FIGURE 12-43 The C_4 pathway. The key enzyme in the C_4 pathway is phosphoenolpyruvate carboxylase, which assimilates CO_2 (red) to form oxaloacetate in mesophyll cells.

Decarboxylation of malate or other C_4 intermediates in bundle sheath cells releases CO_2, which enters the standard Calvin cycle via rubisco (see Figure 12-40, *top* and 12-41).

C_4 plants are superior to C_3 plants in their ability to use the available CO_2 because the C_4 enzyme phosphoenolpyruvate carboxylase has a higher affinity for CO_2 than does rubisco. However, one ATP is converted to one AMP in the cyclical C_4 process (to generate phosphoenolpyruvate from pyruvate); thus the overall efficiency of the photosynthetic production of sugars from NADPH and ATP is lower than it is in C_3 plants, which use only the Calvin cycle for CO_2 fixation. Nonetheless, the net rates of photosynthesis for C_4 grasses, such as corn or sugarcane, can be two to three times the rates for otherwise similar C_3 grasses, such as wheat, rice, or oats, owing to the elimination of losses from photorespiration.

KEY CONCEPTS OF SECTION 12.10

ATP and NADPH Drive Carbon Fixation in the Calvin Cycle and Carbohydrate Synthesis in Stage 4 of Photosynthesis

- In stage 4, the NADPH and ATP generated in stages 2 and 3 provide the energy and the electrons to drive the fixation of CO_2, which results in the synthesis of carbohydrates. These reactions occur in the thylakoid stroma and cytosol.
- Stages 1–3 are called the *light reactions* of photosynthesis because they are directly dependent on light energy. The reactions in stage 4 are indirectly dependent on light energy; they are sometimes called the *dark reactions* of photosynthesis because they can occur in the dark, although they often occur during illumination.
- In the Calvin cycle, also called the C_3 *pathway* of carbon fixation, atmospheric CO_2 is incorporated, or fixed, into organic molecules in a series of reactions that occur in the chloroplast stroma.

- The initial reaction in the Calvin cycle, catalyzed by the enzyme rubisco, forms a three-carbon intermediate, 3-phosphoglycerate. Some of the glyceraldehyde 3-phosphate generated in the cycle is transported to the cytosol and converted to sucrose (see Figure 12-40).
- Rubisco is light- and redox-sensitive. Its activity is regulated by the enzyme rubisco activase.
- ADP-glucose is polymerized to starch polymers in the stroma, where they are stored in densely packed aggregates called *granules*.
- In C_3 plants, a substantial fraction of the CO_2 fixed by the Calvin cycle can be lost as the result of photorespiration, a wasteful reaction catalyzed by rubisco that is favored at low CO_2 and high O_2 levels (see Figure 12-42).
- In C_4 plants, CO_2 is fixed initially in the mesophyll cells by a rubisco-independent reaction with phosphoenolpyruvate. The four-carbon molecules so generated are shuttled to the bundle sheath cells, where the CO_2 is released and then fixed by rubisco in the Calvin cycle. The rate of photorespiration in C_4 plants is much lower than it is in C_3 plants.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the concepts in this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional tools including videos, animations, and quizzes

Key Terms

acetyl-coenzyme A (acetyl CoA)
adenosine triphosphate (ATP)
aerobic oxidation (aerobic respiration)
anaerobic respiration
ATP synthase
binding-change mechanism
Calvin cycle
carbon fixation
catabolism
C_4 pathway
chemiosmosis
chlorophylls

chloroplast
citric acid cycle
cytochrome
electron-transport chain
endosymbiont
endosymbiont hypothesis
energy charge
fermentation
fission (mitochondrial)
flavin adenine dinucleotide (FAD)
F_0F_1 complex
fusion (mitochondrial)
glycolysis
light-harvesting complexes (LHCs)
membrane contact sites (MCSs)
mitochondria-associated membranes (MAMs)
mitochondrial cristae
mitochondrial DNA
mitochondrial inner membrane
mitochondrial intermembrane space
mitochondrial matrix
mitochondrial outer membrane
mitochondrion
mitophagy
multiprotein complexes
nicotinamide adenine dinucleotide (NAD^+)
nicotinamide adenine dinucleotide phosphate ($NADP^+$)
oxidative phosphorylation

peroxisomes
photoelectron transport
photorespiration
photosynthesis
photosystem
photosystem I (PSI)
photosystem II (PSII)
porin
prosthetic group
proton-motive force
proton pump
proton pumping
pyruvate
Q cycle
reactive oxygen species (ROS)
reduction potential (*E*)
respiration
respiratory chain
respiratory control
rubisco (ribulose 1,5-bisphosphate carboxylase)
special-pair chlorophylls
starch
stroma
substrate-level phosphorylation
supercomplexes
superoxide anions
thylakoid lumen
thylakoid membrane

thylakoids
triacylglycerol (triglyceride)
ubiquinone [coenzyme Q (CoQ)]
uncoupler
uncoupling protein 1 (UCP1)

Review the Concepts

1. The proton-motive force is essential for both mitochondrial and chloroplast function. What produces the proton-motive force, and what is its relationship to ATP? The compound 2,4-dinitrophenol (DNP), which was used in diet pills in the 1930s but later shown to have dangerous side effects, allows protons to diffuse across membranes. Why is it dangerous to consume DNP?
2. Mitochondria and chloroplasts are thought to have evolved from symbiotic bacteria. What is the experimental evidence from this chapter that supports this hypothesis?
3. The inner mitochondrial membrane exhibits all of the fundamental characteristics of a typical cell membrane, but it also has several unique characteristics that are closely associated with its role in oxidative phosphorylation. What are these unique characteristics? How does each contribute to the function of the inner membrane?
4. Maximal production of ATP from glucose involves the reactions of glycolysis, the citric acid cycle, and the

electron-transport chain. Which of these reactions requires O_2, and why? Which, in certain organisms or physiological conditions, can proceed in the absence of O_2?

5. Fermentation permits the continued extraction of energy from glucose in the absence of oxygen. If glucose catabolism is anaerobic, why is fermentation necessary for glycolysis to continue?
6. Describe the step-by-step process by which electrons from glucose catabolism in the cytoplasm are transferred to the electron-transport chain in the inner mitochondrial membrane. In your answer, note whether the electron transfer at each step is direct or indirect.
7. Mitochondrial oxidation of fatty acids is a major source of ATP, yet fatty acids can be oxidized elsewhere. What organelle, besides the mitochondrion, can oxidize fatty acids? What is the fundamental difference between oxidation occurring in this organelle and mitochondrial oxidation?
8. Each of the cytochromes in the mitochondrion contains prosthetic groups. What is a prosthetic group? Which type of prosthetic group is associated with the cytochromes? What property of the various cytochromes ensures unidirectional electron flow along the electron-transport chain?
9. The electron-transport chain consists of a number of multiprotein complexes, which work in conjunction to pass electrons from an electron carrier, such as NADH, to O_2. What is the role of these complexes in ATP synthesis? It has

been demonstrated that respiration supercomplexes contain all the protein components necessary for respiration. Why is this beneficial for ATP synthesis, and what is one way that the existence of supercomplexes has been demonstrated experimentally? Coenzyme Q (CoQ) is not a protein, but a small, hydrophobic molecule. Why is it important for the functioning of the electron-transport chain that CoQ is a hydrophobic molecule?

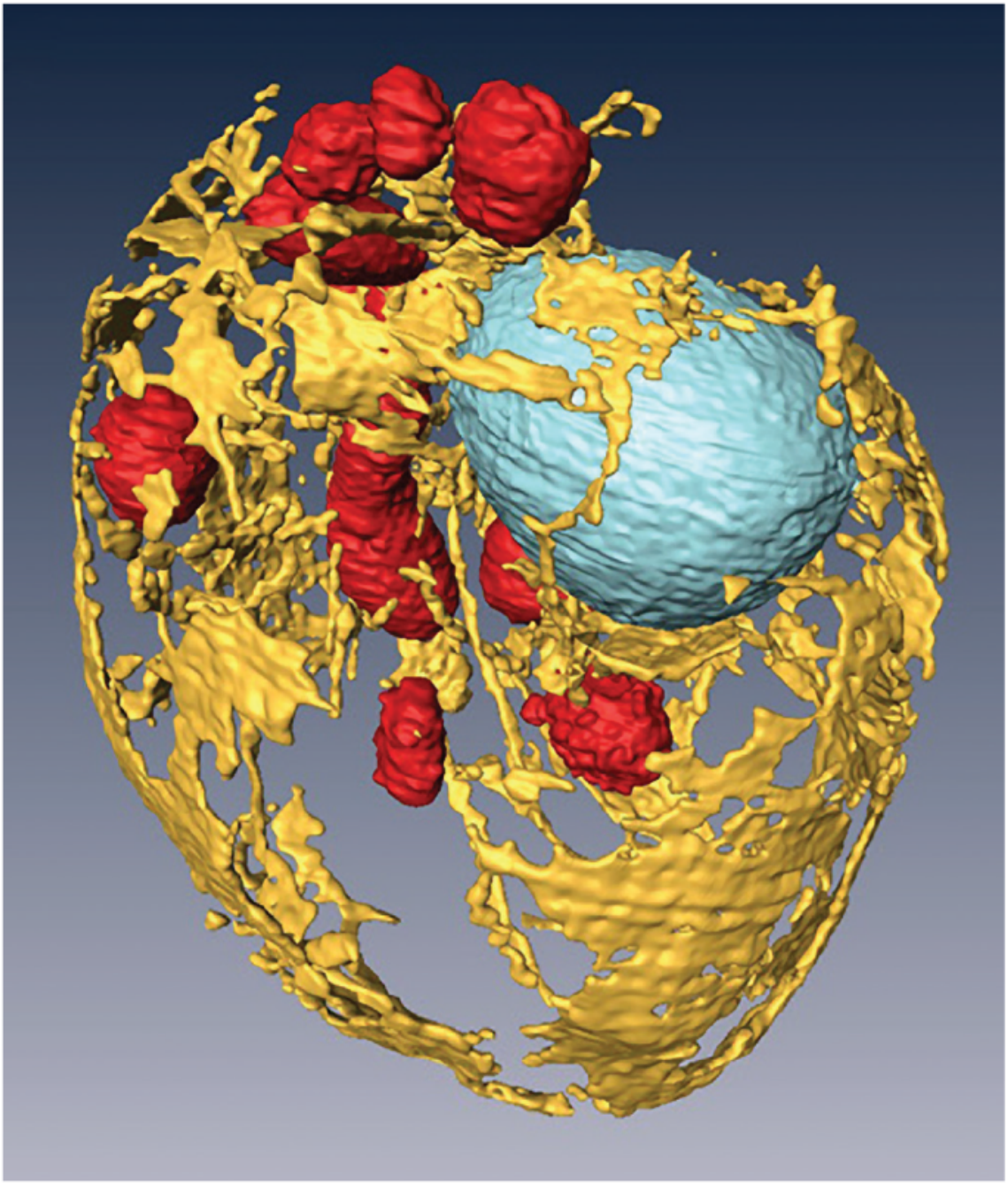
10. It is estimated that each electron pair donated by NADH leads to the synthesis of approximately three ATP molecules, whereas each electron pair donated by $FADH_2$ leads to the synthesis of approximately two ATP molecules. What is the underlying reason for the difference in yield for electrons donated by $FADH_2$ versus NADH?
11. Describe the main functions of the different components of the ATP synthase enzyme in the mitochondrion. A structurally similar enzyme is responsible for the acidification of lysosomes and endosomes. Given what you know about the mechanism of ATP synthesis, explain how this acidification might occur.
12. Much of our understanding of ATP synthase is derived from research on aerobic bacteria. What makes these organisms useful for this research? Where do the reactions of glycolysis, the citric acid cycle, and the electron-transport chain occur in these organisms? Where is the proton-motive force generated in aerobic bacteria? What other cellular processes depend on the proton-motive force in these organisms?

13. An important function of the inner mitochondrial membrane is to provide a selectively permeable barrier to the movement of water-soluble molecules and thus to generate different chemical environments on either side of the membrane. However, many of the substrates and products of oxidative phosphorylation are water soluble and must cross the inner membrane. How does this transport occur?
14. The Q cycle plays a major role in the electron-transport chain of mitochondria, chloroplasts, and bacteria. What is the function of the Q cycle, and how does it carry out this function? What electron-transport components participate in the Q cycle in mitochondria, in purple bacteria, and in chloroplasts?
15. True or false: Since ATP is generated in chloroplasts, cells capable of undergoing photosynthesis do not require mitochondria. Explain. Name and describe the idea that explains how mitochondria and chloroplasts are thought to have originated in eukaryotic cells.
16. Write the overall reaction of oxygen-generating photosynthesis. Explain the following statement: The O_2 generated by photosynthesis is simply a by-product of the pathway's generation of carbohydrates and ATP.
17. Photosynthesis can be divided into multiple stages. What are the stages of photosynthesis, and where does each occur within the chloroplast? Where is the sucrose produced by photosynthesis generated?

18. The photosystems responsible for absorption of light energy are each composed of two linked components, the reaction center and an antenna complex. What is the pigment composition and role of each component in the process of light absorption? What evidence exists that the pigments found in these components are involved in photosynthesis?
19. Photosynthesis in green and purple bacteria does not produce O_2. Why? How can these organisms still use photosynthesis to produce ATP? What molecules serve as electron donors in these organisms?
20. Chloroplasts contain two photosystems. What is the function of each? For linear electron flow, diagram the flow of electrons from photon absorption to NADPH formation. What does the energy stored in the form of NADPH synthesize?
21. The Calvin cycle reactions that fix CO_2 do not function in the dark. What are the likely reasons for this? How are these reactions regulated by light?
22. Rubisco, which may be the most abundant protein on Earth, plays a key role in the synthesis of carbohydrates in organisms that use photosynthesis. What is rubisco, where is it located, and what function does it serve?

CHAPTER 13

Moving Proteins into Membranes and Organelles



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

A three-dimensional reconstruction of the internal membranes of a yeast cell using scanning electron microscopy. The cell wall has been removed and the organelles highlighted with false color to reveal the endoplasmic reticulum (yellow), mitochondria (red), and nucleus (blue). Cell diameter is $3.5\ \mu m$.

[From D. Wei et al., 2012, "High-Resolution Three-Dimensional Reconstruction of a Whole Yeast Cell Using Focused-Ion Beam Scanning Electron Microscopy," *Biotechniques* **53**(1):41–48.]

OUTLINE

13.1 Targeting Proteins to and Across the ER Membrane
13.2 Insertion of Membrane Proteins into the ER
13.3 Protein Modifications, Folding, and Quality Control in the ER
13.4 Targeting of Proteins to Mitochondria and Chloroplasts
13.5 Targeting of Peroxisomal Proteins
13.6 Transport into and out of the Nucleus

A typical mammalian cell contains up to 10,000 different kinds of proteins; a yeast cell, about 5000. The vast majority of these proteins are synthesized by cytosolic ribosomes, and although many newly synthesized proteins remain within the cytosol (see Chapter 5), as many as half of the different kinds of proteins produced in a typical cell are delivered to one or another of the various membrane-bounded compartments of the cell. For example, many receptor proteins and transport proteins must be delivered to the plasma membrane, digestive enzymes and polypeptide signaling molecules must be directed to the cell surface for secretion from the cell, and enzymes such as RNA and DNA polymerases must be targeted to the nucleus. These and all the other proteins produced by a cell must reach their correct locations for the cell to function properly.

The delivery of newly synthesized proteins to their proper cellular destinations, usually referred to as *protein targeting* or *protein sorting*, encompasses two very different kinds of processes: **signal-based targeting** to a variety of organelles and **vesicle-based trafficking** in the secretory pathway. The first kind of process involves the targeting of a newly synthesized protein from the cytoplasm to an intracellular organelle. Targeting can occur during translation or soon after synthesis

of the protein is complete. For membrane proteins, targeting leads to insertion of the protein into the lipid bilayer of the membrane, whereas for water-soluble proteins, targeting leads to translocation of the entire protein across the membrane into the aqueous interior of the organelle. Proteins are sorted to the endoplasmic reticulum (ER), mitochondria, chloroplasts, peroxisomes, and nucleus by this general process (Figure 13-1).

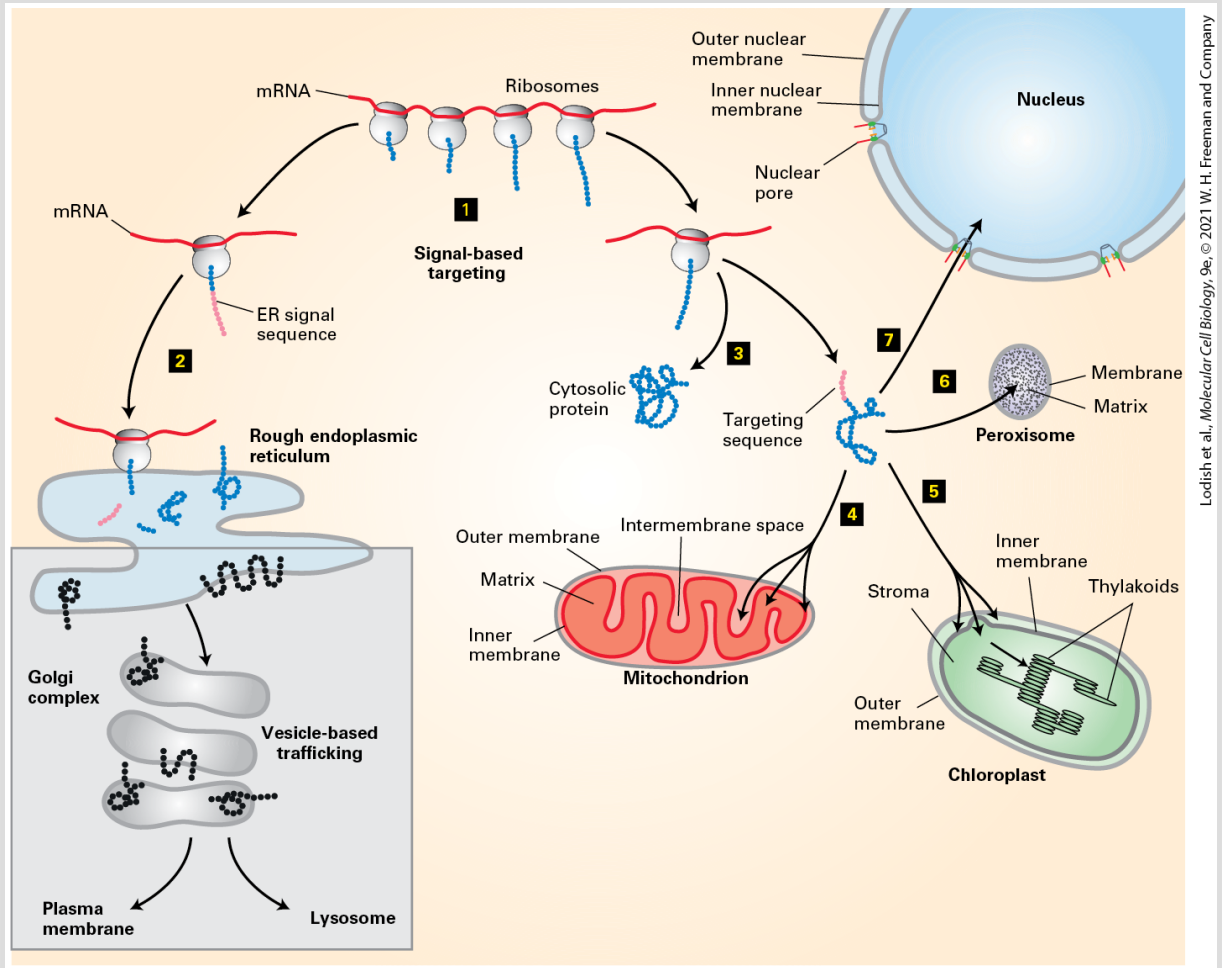


FIGURE 13-1 Overview of major protein-sorting pathways in eukaryotes. All nuclear DNA–encoded mRNAs are translated on cytosolic ribosomes (step 1), but proteins can be directed to different intracellular locations according to targeting signals within the polypeptide sequence. *Left* (secretory pathway): Ribosomes synthesizing nascent secretory proteins are directed to the rough endoplasmic reticulum (ER) by an ER signal sequence (pink; step 2). After translation is completed

on the ER, these proteins can move by vesicle-based transport processes to the Golgi complex and on to the plasma membrane or to lysosomes. The vesicle-based processes underlying the secretory pathway (shaded box) are discussed in Chapter 14. *Right* (signal-based targeting): Synthesis of proteins that contain no targeting sequence are released into the cytosol and remain there (step 3). Proteins with an organelle-specific targeting sequence (pink) are first released into the cytosol but are then imported into mitochondria, chloroplasts, peroxisomes, or the nucleus (steps 4–7). Mitochondrial and chloroplast proteins typically pass through the outer and inner membranes to enter the matrix or stromal space, respectively. Other proteins are sorted to other subcompartments of these organelles by additional sorting steps. Nuclear proteins enter and exit through visible pores in the nuclear envelope.

The second general sorting process, known as the **secretory pathway**, involves transport of proteins from the ER to their final destination within membrane-enclosed vesicles. For many proteins, including those that make up the extracellular matrix, the final destination is the outside of the cell (hence the name); integral membrane proteins are also transported to the Golgi complex, lysosomes, and plasma membrane by this process. The secretory pathway begins in the ER; thus all proteins slated to enter the secretory pathway are initially targeted to this organelle.

Targeting to the ER usually involves *nascent* proteins still in the process of being synthesized on a ribosome. Newly made proteins are thus extruded from the ribosome directly into the ER membrane. Once translocated across the ER membrane, proteins are assembled into their native conformation by protein-folding catalysts present in the lumen of the ER. Indeed, the ER is the location where about one-third of the proteins in a typical cell fold into their native conformations, and most of the resident ER proteins either directly or indirectly contribute to the folding process. As part of the folding process, proteins also undergo specific post-

translational modifications in the ER. These processes are monitored carefully, and only after their folding and assembly is complete are proteins permitted to be transported out of the ER to other destinations along the secretory pathway. Proteins whose final destination is the lysosome, plasma membrane, or cell exterior are transported along the secretory pathway by the action of small vesicles that bud from the membrane of one organelle and then fuse with the membrane of another (see Figure 13-1 [shaded box] and Figure 14-1). We discuss vesicle-based protein trafficking in Chapter 14 because mechanistically it differs significantly from non-vesicle-based protein targeting to intracellular organelles.

In this chapter, we examine how proteins are targeted to five intracellular organelles: the ER, mitochondria, chloroplasts, peroxisomes, and nucleus. Two features of this protein-targeting process were initially quite baffling: How could a given protein be directed to only one specific membrane, and how could relatively large hydrophilic protein molecules be translocated across a hydrophobic membrane without disrupting the function of the bilayer as a barrier to ions and small molecules? Using a combination of biochemical purification methods and genetic screens for identifying mutants unable to execute particular translocation steps, cell biologists have identified many of the cellular components required for translocation across each of the different intracellular membranes. In addition, many of the major translocation processes in the cell have been reconstituted by incorporating their purified protein components into artificial lipid bilayers, using in vitro systems that can be freely manipulated experimentally.

These studies have shown that, despite some variations, the same basic mechanisms govern protein sorting to all the various intracellular organelles. As shown in Table 13-1, the mechanism of targeting to five organelles considered in this chapter can be described by four basic elements. (1) The information to target a protein to a particular organelle destination is encoded within the amino acid sequence of the protein itself, usually within a sequence of about 20 amino acids, known generically as a **targeting sequence**; these sequences are also called **signal sequences** or **signal peptides**. Such targeting sequences are usually located at the N-terminus of a protein and are thus the first part of the protein to be synthesized. More rarely, targeting sequences are located at either the C-terminus or within the interior of a protein sequence. (2) Each organelle carries a set of receptor proteins that bind (directly or indirectly) only to specific kinds of targeting sequences, thus ensuring the specificity of targeting. (3) Once a protein containing a targeting sequence has interacted with the corresponding receptor, the polypeptide chain is transferred to some kind of *translocation channel* that allows the protein to pass into or through the membrane bilayer. (4) Finally, the unidirectional transfer of a protein into an organelle, without its sliding back out into the cytoplasm, is usually achieved by coupling translocation to an energetically favorable process such as hydrolysis of GTP or ATP. In some cases, proteins are sorted further to reach a subcompartment within the target organelle; such sorting depends on yet other signal sequences and other receptor proteins.

TABLE 13-1 • Targeting Sequences Direct Proteins from the Cytosol to Organelles

Target	Targeting	Receptor	Translocation	Energy Source

Organelle	Sequence		Channel	
Endoplasmic reticulum (lumen)	At N-terminus, 6–12 hydrophobic amino acids, often preceded by one or more basic amino acids (Arg, Lys)	SRP (ribonucleoprotein complex associated with ribosome) and SRP receptor in the ER membrane; SRP and SRP receptor are GTPases	Sec61 complex; proteins translocate as unfolded chain, as channel remains sealed to small molecules	Translation elongation powered by GTP hydrolysis
Mitochondrion[i] (matrix)	At N-terminus an amphipathic helix, 20–50 residues in length, with Arg and Lys residues on one side and hydrophobic residues on the other	Tom20/22 import receptor in the outer mitochondrial membrane	Channels composed of Tom40 in the outer mitochondrial membrane and Tim23/17 in the inner membrane	ATP hydrolysis by Hsp70 in the matrix
Chloroplast[i] (stroma)	At N-terminus, but no common motifs; generally rich in Ser, Thr, and small hydrophobic residues and	Toc159 and Toc34, GTPases in the outer membrane	Channels composed Toc75 in the outer membrane and Tic20 in the inner membrane	ATP hydrolysis by Hsp70 in the stroma

	poor in Glu and Asp			
Peroxisome (matrix)	PTS1 signal (Ser-Lys-Leu) at extreme C-terminus; PTS2 signal at N-terminus	Pex5, which cycles between the cytoplasm and peroxisomal membrane	A complex of Pex5 and the peroxisomal membrane protein Pex14; cargo proteins can be transported in a folded state	ATP hydrolysis coupled to ubiquitination and deubiquitination of Pex5
Nucleus (nucleoplasm)	NLS sequences can function at any location within a protein sequence; a common motif includes a short segment rich in Lys and Arg residues	Nuclear transport receptors, which cycle between the cytoplasm and the nuclear interior	Central channel of the nuclear pore complex which is filled with a fluidlike phase of FG-repeat proteins; proteins and RNA can traverse in folded state when bound to a nuclear transport receptor	GTP hydrolysis coupled to cycling of Ran GTPase into and out of the nucleus

[i]Targeting to subcompartments of the mitochondria (such as the inner membrane) and the chloroplast (such as the thylakoid) require additional targeting sequences, receptors, and translocation channels as discussed in this chapter.

In the first part of the chapter, we cover targeting of proteins to the ER, including the post-translational modifications that proteins undergo as they enter the secretory pathway. Targeting of proteins to the ER is the best understood example of protein targeting and will serve as a model of the process in general. We then describe targeting of proteins to mitochondria, chloroplasts, and peroxisomes. Finally, we cover the transport of proteins into and out of the nucleus through nuclear pores.

An important consequence of knowing the nature of the different kinds of targeting sequences is that from just the amino acid sequence encoded by a gene it is possible to reliably deduce the final cellular location of the gene product. Indeed, the cellular location of most proteins encoded by the human genome has been accurately predicted using the information on the nature of different kinds of targeting sequences that we will discuss in this chapter.

13.1 Targeting Proteins to and Across the ER Membrane

All eukaryotic cells have an endoplasmic reticulum (ER). The ER is a convoluted organelle, made up of tubules and flattened sacs, whose membrane is continuous with the membrane of the nucleus. The ER usually has a very large surface area, and its membrane is where cellular lipids are synthesized (see Chapter 10). The ER is also where most membrane proteins are assembled, including those of the plasma membrane and the membranes of the lysosomes, ER, and Golgi complex. In addition, all soluble proteins that will eventually be secreted from the cell — as well as those destined for the lumen of the ER, Golgi complex, or lysosomes — are initially delivered to the ER lumen (see Figure 13-1). Since the ER plays such an important role in protein secretion, we refer to the pathway of protein trafficking that flows through the ER as the *secretory pathway*. For simplicity, we will refer to all proteins initially targeted to the ER as *secretory proteins*, but keep in mind that not all proteins that are targeted to the ER are actually secreted from the cell.

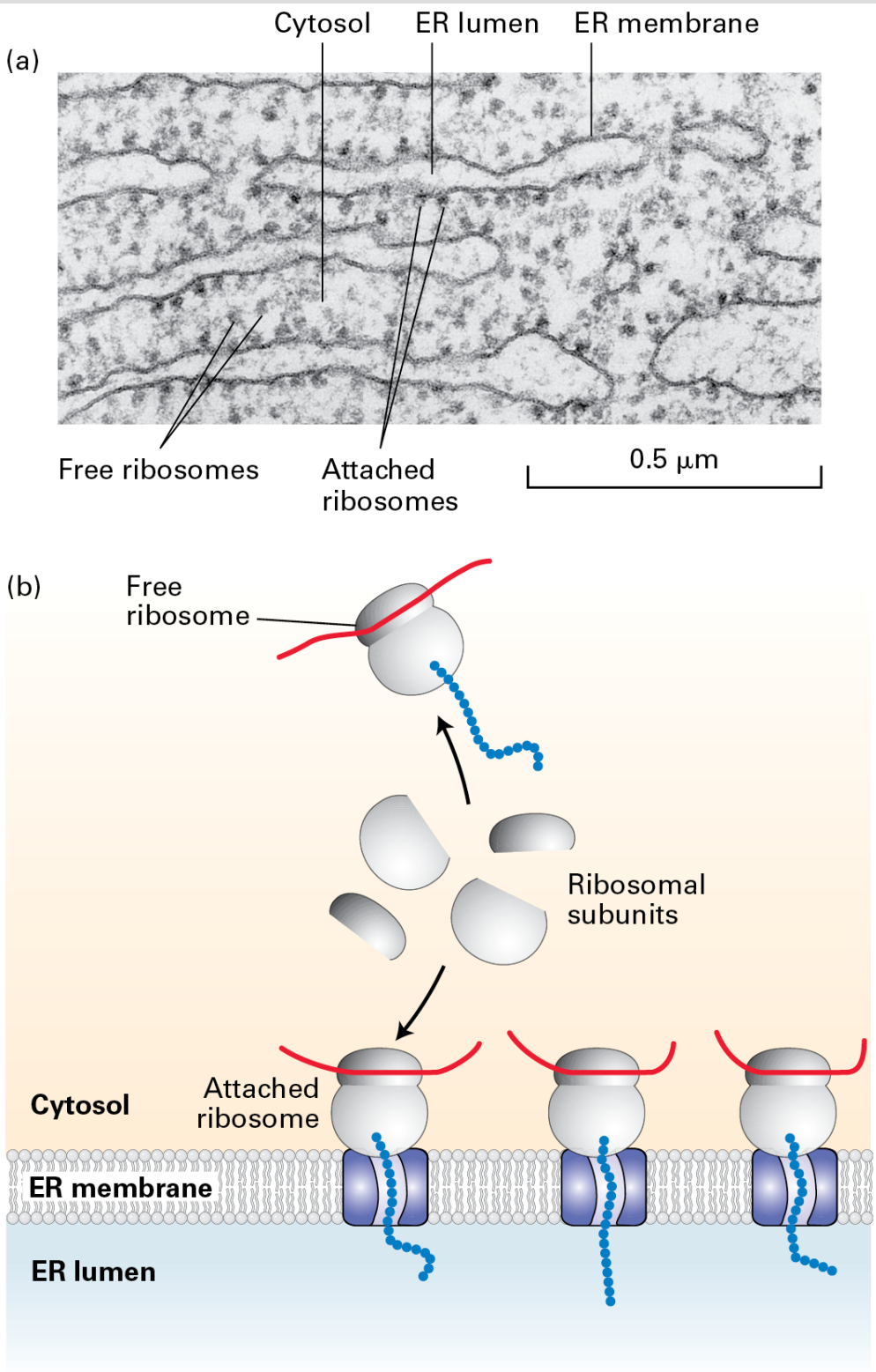
In this first section of the chapter, we discuss how proteins are initially identified as secretory proteins and how such proteins are translocated across the ER membrane. We deal first with soluble proteins — those that pass all the way through the ER membrane, into the lumen. In the next section, we discuss integral membrane proteins, which are inserted into the ER membrane.

Pulse-Chase Experiments with Purified ER Membranes Demonstrated That Secreted Proteins Cross the ER Membrane

Although all cells secrete a variety of proteins (e.g., extracellular matrix proteins), certain types of cells are specialized for secretion of large amounts of specific proteins. Pancreatic acinar cells, for instance, synthesize large quantities of several digestive enzymes, which are secreted into ductules that lead to the intestine. Because such secretory cells contain the organelles of the secretory pathway (e.g., ER and Golgi complex) in great abundance, they have been widely used in studying this pathway, including the initial steps that occur at the ER membrane.

The sequence of events that occur immediately after the synthesis of a secretory protein was first elucidated by pulse-chase experiments with pancreatic acinar cells. In these experiments, radioactively labeled amino acids were incorporated into secretory proteins as they were synthesized on ribosomes bound to the surface of the ER. The portion of the ER that receives proteins entering the secretory pathway is known as the **rough ER** because it is so densely studded with ribosomes that its surface appears morphologically distinct from other ER membranes (Figure 13-2). From these experiments, it became clear that during or immediately after their synthesis on the

ribosome, secretory proteins translocate across the ER membrane into the lumen of the ER.

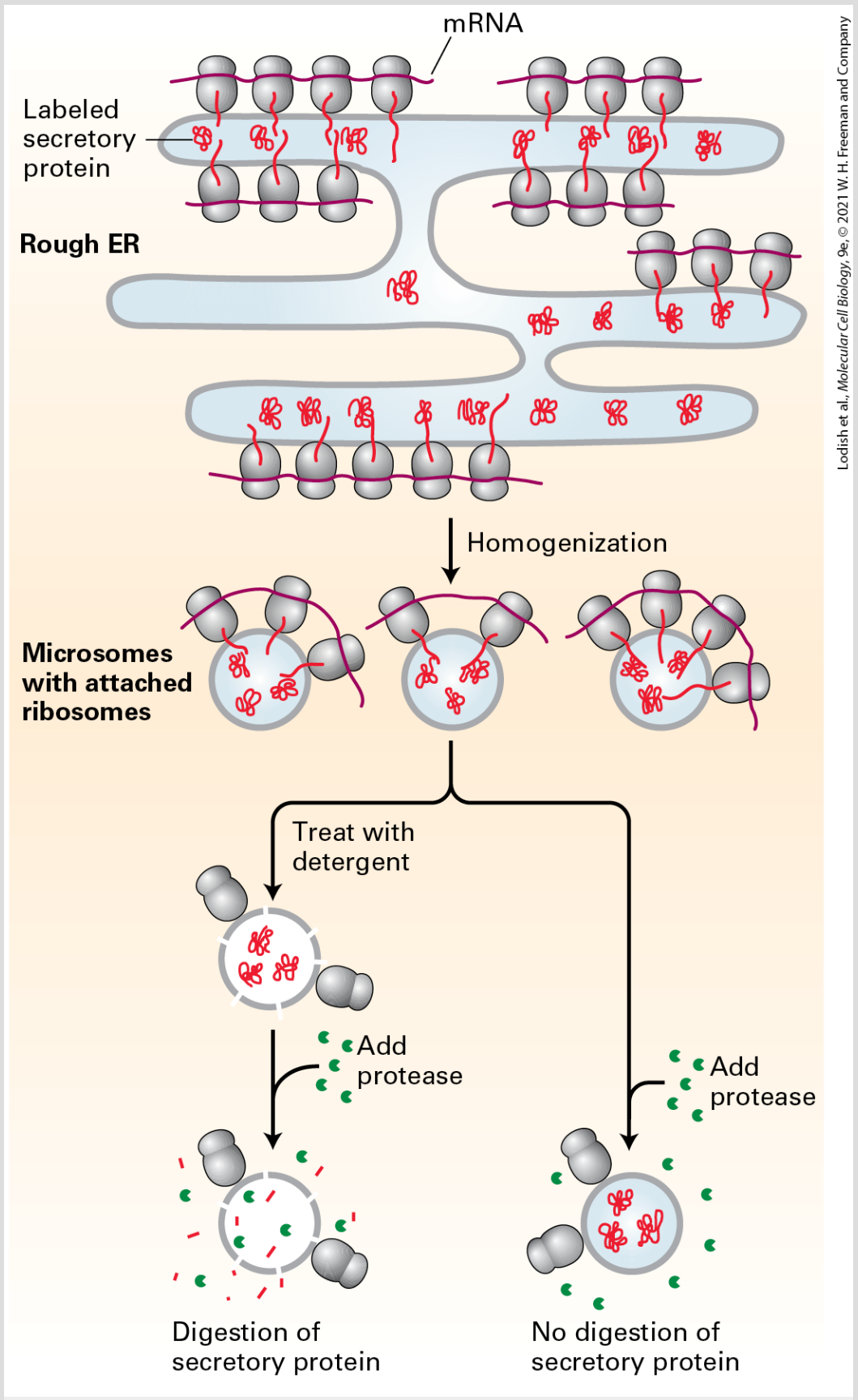


Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Part (a) Courtesy of Dr. Marilyn G. Farquhar, University of California, San Diego.

FIGURE 13-2 Structure of the rough ER. (a) Electron micrograph of ribosomes attached to the rough ER in a pancreatic acinar cell. Most of the proteins

synthesized by this type of cell are secretory proteins and are formed on membrane-attached ribosomes. A few unattached (free) ribosomes are evident; presumably, these ribosomes are synthesizing cytosolic or other nonsecretory proteins. (b) Schematic representation of protein synthesis on the ER. Note that membrane-bound and free cytosolic ribosomes are identical. Membrane-bound ribosomes are recruited to the ER during synthesis of a polypeptide containing an ER signal sequence.

To delineate the steps in the translocation process, it was necessary to isolate the ER from the rest of the cell. Isolation of intact ER, with its delicate lacelike structure and its interconnectedness with other organelles, is not feasible. However, scientists discovered that when cells are homogenized, the rough ER breaks up into small closed vesicles with ribosomes on the outside, termed **microsomes**, which retain most of the biochemical properties of the intact ER, including the capability of protein translocation. The experiments depicted in Figure 13-3, in which microsomes isolated from pulse-labeled cells were treated with a protease, demonstrate that although secretory proteins are synthesized on ribosomes bound to the cytosolic face of the ER membrane, the polypeptides produced by these ribosomes end up within the lumen of a microsome. Experiments such as these raised the question of how polypeptides are recognized as secretory proteins shortly after their synthesis begins and how a nascent secretory protein is threaded across the ER membrane.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-3 Secretory proteins enter the ER lumen. Labeling experiments demonstrated that secretory proteins are localized to the ER lumen shortly after synthesis. Cells are incubated for a brief time with radiolabeled amino acids so that only newly synthesized proteins become labeled. The cells are then homogenized, fracturing the plasma membrane and shearing the rough ER into small vesicles called *microsomes*. Because they have bound ribosomes, microsomes have a much greater buoyant density than other membranous organelles and can be separated from them by a combination of differential and sucrose density-gradient centrifugation (see Chapter 4). The purified microsomes are treated with a protease in the presence or absence of a detergent. The labeled secretory proteins associated with the microsomes are digested by the protease only if the microsomal membrane is first destroyed by treatment with detergent. This finding indicates that the newly made proteins are inside the microsomes, equivalent to the lumen of the rough ER.

A Hydrophobic N-Terminal Signal Sequence Targets Nascent Secretory Proteins to the ER

After synthesis of a secretory protein begins on free ribosomes in the cytosol, a 16–30-residue ER targeting sequence in the nascent protein directs the ribosome to the ER membrane and initiates translocation of the growing polypeptide across the ER membrane (see Figure 13-1, *left*). An ER targeting sequence, located at the N-terminus of the protein, is usually known as a *signal sequence*. The signal sequences of different secretory proteins all contain one or more positively charged amino acids adjacent to a continuous stretch of 6–12 hydrophobic residues (known as the *hydrophobic core*)

but otherwise have little in common. The signal sequence is cleaved from most secretory proteins while they are still elongating on the ribosome; thus signal sequences are usually not present in the mature proteins found in cells.

The hydrophobic core of an ER signal sequence is essential for its function. For instance, the specific deletion of several of the hydrophobic amino acids from a signal sequence or the introduction of charged amino acids into the hydrophobic core by mutation can abolish the ability of the N-terminus of a protein to function as a signal sequence. As a consequence, the modified protein remains in the cytosol, unable to cross the ER membrane into the lumen. Conversely, signal sequences can be added to normally cytosolic proteins using recombinant DNA techniques. Provided the added sequence is sufficiently long and hydrophobic, such a modified cytosolic protein can acquire the ability to be translocated into the ER lumen. The hydrophobic residues in the core of an ER signal sequence form a binding site that is critical for the interaction of the signal sequence with the machinery responsible for targeting the protein to the ER membrane.

Biochemical studies using a cell-free protein-synthesizing system, mRNA encoding a secretory protein, and microsomes stripped of their own bound ribosomes have elucidated how ER signal sequences function during protein translocation. Initial experiments with this system demonstrated that a typical secretory protein is incorporated into microsomes and has its signal sequence

removed only if the microsomes are present during protein synthesis. If instead of being present during translation, the addition of microsomes to the system is delayed until after protein synthesis is completed, no protein transport into the microsomes occurs (Figure 13-4). Subsequent experiments were designed to determine the precise stage of protein synthesis at which microsomes must be present in order for translocation to occur. In these experiments, microsomes were added to the reaction mixtures at different times after protein synthesis had begun. These experiments showed that microsomes must be added before the first 70 or so amino acids are translated in order for the completed secretory protein to be localized in the microsomal lumen. At this point, the first 40 or so amino acids protrude from the ribosome, including the signal sequence that will later be cleaved off, and the next 30 or so amino acids are still buried within a channel in the ribosome (see Figure 5-34). Thus the transport of most secretory proteins into the ER lumen begins while the incompletely synthesized (nascent) protein is still bound to the ribosome, a process referred to as **cotranslational translocation**.

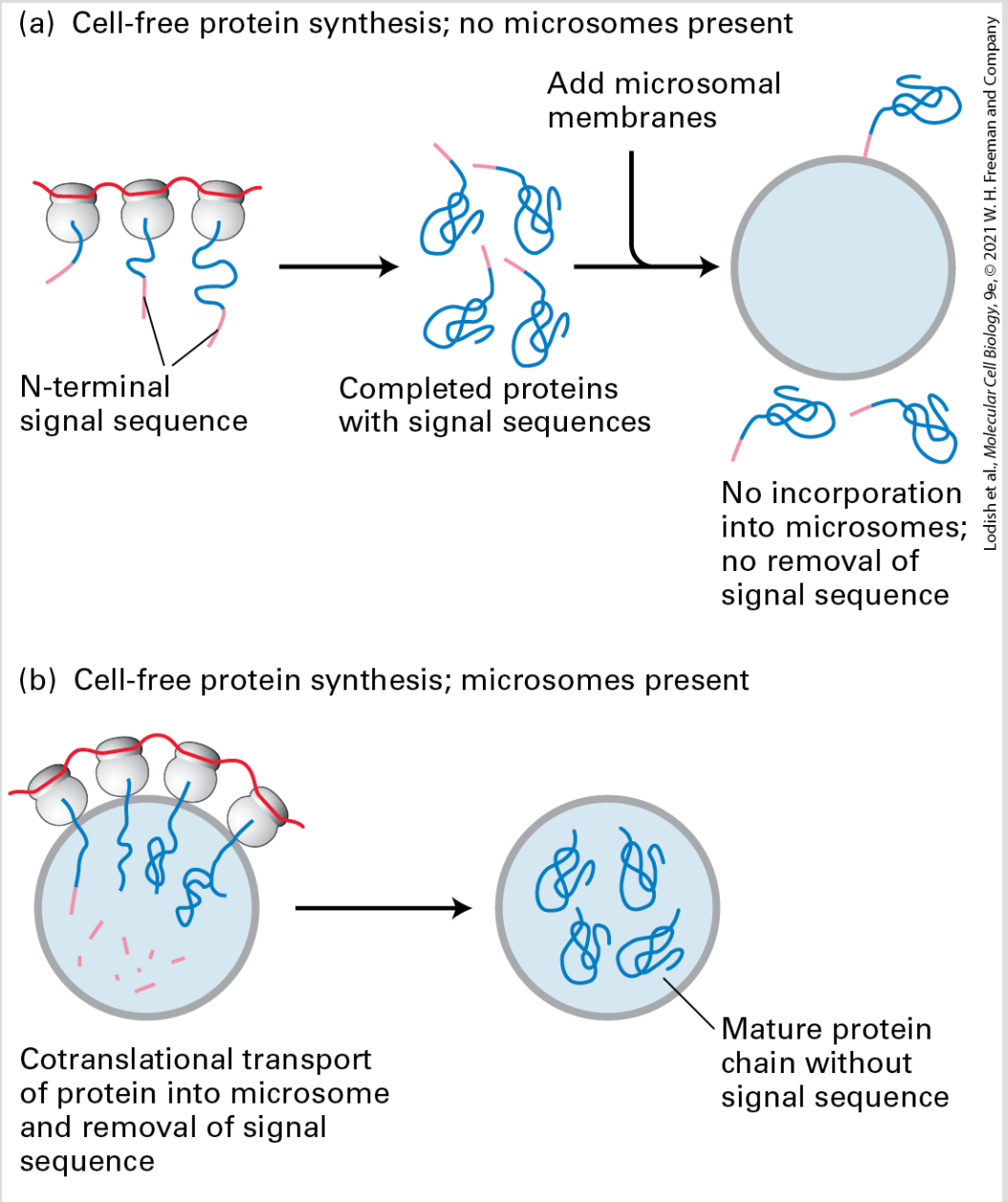


FIGURE 13-4 Translation and translocation occur simultaneously. Cell-free experiments demonstrate that translocation of secretory proteins into microsomes is coupled to translation. Treatment of microsomes with EDTA, which chelates Mg^{2+} ions, strips them of associated ribosomes, allowing isolation of ribosome-free microsomes, which are equivalent to ER membranes (see Figure 13-3). Protein synthesis is carried out in a cell-free system containing functional ribosomes, aminoacyl-tRNAs, ATP, GTP, and cytosolic enzymes, to which mRNA encoding a secretory protein is added. (a) When microsomes are added after synthesis of secretory proteins is complete, the

secretory proteins that have already been formed will not enter microsomes. (b) However, when microsomes are present during protein synthesis, secretory proteins will be translocated across the vesicle membrane and lose their signal sequences (resulting in a decrease in molecular weight) as they are synthesized.

Cotranslational Translocation Is Initiated by Two GTP-Hydrolyzing Proteins

Given that secretory proteins are synthesized in association with the ER membrane but not with any other cellular membrane, a signal-sequence recognition mechanism must target them there. The two key components in this targeting are the **signal recognition particle (SRP)** and its receptor. The SRP is a cytosolic ribonucleoprotein particle that transiently binds to both the ER signal sequence in a nascent protein and the large (60S) ribosomal subunit, forming a large complex. The SRP then targets the nascent protein–ribosome complex to the ER membrane by binding to the **SRP receptor**, which is located in the membrane.

The SRP is made up of six proteins bound to a 300-nucleotide RNA, which acts as a scaffold for the hexamer. One of the SRP proteins (P54) can be chemically cross-linked to ER signal sequences, which shows that this subunit is the one that binds to the signal sequence in a nascent secretory protein. A region of P54 known as the *M*

domain, containing many methionine and other amino acid residues with hydrophobic side chains, forms a cleft or groove whose inner surface is lined by hydrophobic side chains (Figure 13-5a). The hydrophobic core of the signal sequence binds to this cleft via hydrophobic interactions. Other polypeptides in the SRP interact with the ribosome or are required for protein translocation into the ER lumen.

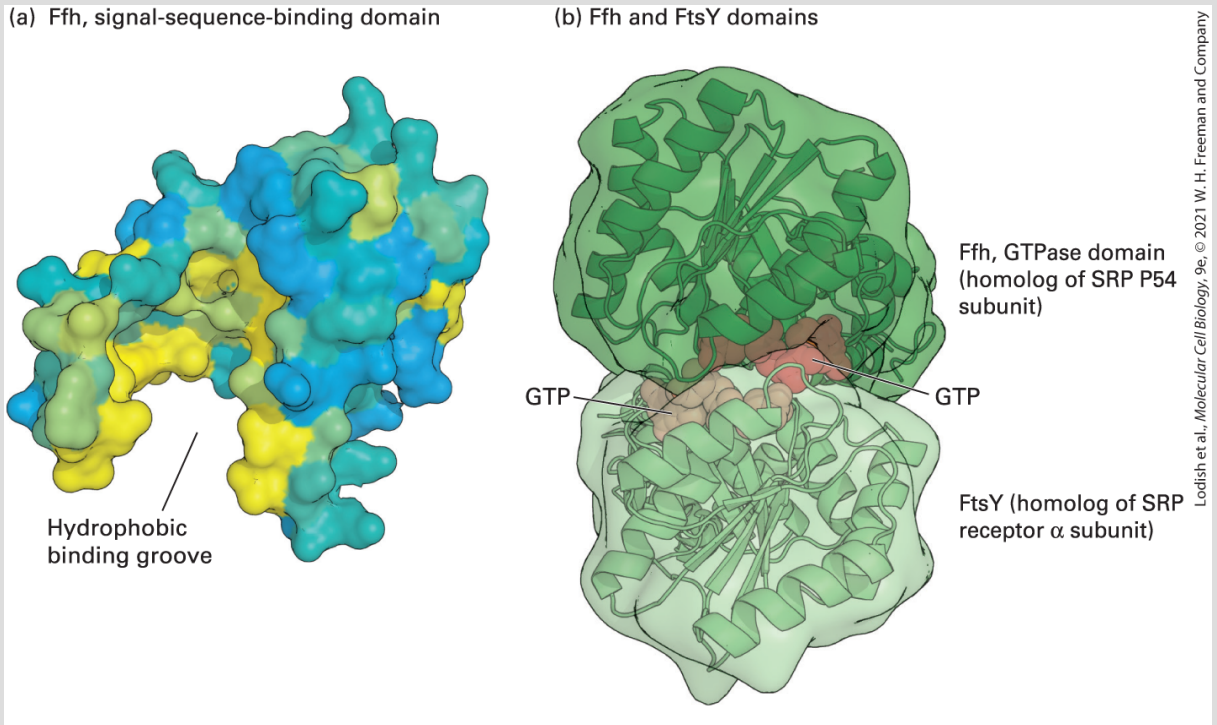


FIGURE 13-5 Structure of the signal recognition particle (SRP). (a) The signal-sequence-binding domain: the bacterial Ffh protein is homologous to the portion of P54 that binds ER signal sequences in eukaryotes. This surface model shows the binding domain in Ffh, which contains a large cleft or groove lined with hydrophobic amino acids (yellow) whose side chains interact with signal sequences. (b) GTP- and receptor-binding domain: the structure of GTP bound to FtsY (the archaeal homolog of the α subunit of the SRP receptor) and Ffh subunits from *Thermus aquaticus* illustrate how the interaction between these proteins is controlled by GTP binding and hydrolysis. Ffh and FtsY each can bind to one molecule of GTP, and when they bind to each other, the two bound molecules of GTP fit in the interface

between the protein subunits and stabilize the dimer. Assembly of the pseudosymmetric dimer allows formation of two active sites for the hydrolysis of both bound GTP molecules.

[Part (a) Data from R. J. Keenan et al., 1998, *Cell* **94**:181, PDB ID 2ffh. Part (b) data from P. J. Focia et al., 2004, *Science* **303**:373, PDB ID 1okk.]

The SRP and the nascent polypeptide chain–ribosome complex bind to the ER membrane by docking with the SRP receptor, an integral protein of the ER membrane made up of two subunits: an α subunit and a smaller β subunit. Interaction of the SRP–nascent chain–ribosome complex with the SRP receptor is strengthened when both the P54 subunit of the SRP and the α subunit of the SRP receptor are bound to GTP. The structure of the SRP P54 subunit and the SRP receptor α subunit (FtsY) from the archaean *Thermus aquaticus* provides insight into how a cycle of GTP binding and hydrolysis can drive the binding and dissociation of these proteins. Figure 13-5b shows that P54 and FtsY, each bound to a single molecule of GTP, come together to form a pseudosymmetric heterodimer. Neither subunit alone contains a complete active site for the hydrolysis of GTP, but when the two proteins come together, they form two complete active sites that are capable of hydrolyzing both bound GTP molecules. Hydrolysis to GDP destabilizes the interface, causing disassembly of the dimer.

Figure 13-6 summarizes our current understanding of secretory protein synthesis and the role of the SRP and its receptor in this process. Hydrolysis of the bound GTP accompanies disassembly of the SRP and SRP receptor and initiates transfer of the nascent chain

and ribosome to a site on the ER membrane, where translocation can take place. After dissociating from each other, the SRP and its receptor each release their bound GDP, SRP is recycled back to the cytosol, and both are ready to initiate another round of interaction between ribosomes synthesizing nascent secretory proteins and the ER membrane.

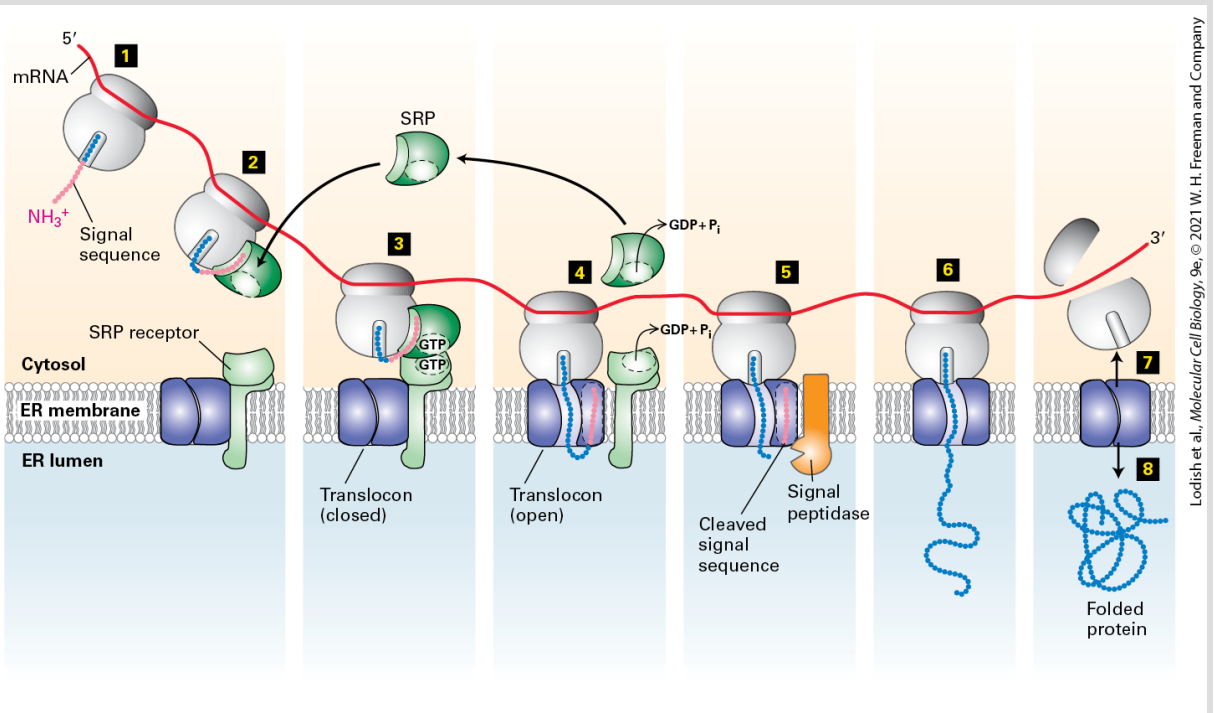


FIGURE 13-6 Cotranslational translocation. Steps 1–2: Once the ER signal sequence emerges from the ribosome, it is bound by a signal recognition particle (SRP). Step 3: The SRP and the nascent polypeptide chain–ribosome complex bind to the SRP receptor in the ER membrane. This interaction is strengthened by the binding of GTP to both the SRP and its receptor. Step 4: Transfer of the nascent polypeptide–ribosome to the translocon leads to opening of this translocation channel to admit the growing polypeptide adjacent to the signal sequence. The hydrophobic signal sequence itself is transferred to a hydrophobic binding site next to the central pore. Both the SRP and SRP receptor, once dissociated from the translocon, hydrolyze their bound GTP and then are ready to initiate the insertion of another polypeptide chain. Step 5: As the polypeptide chain elongates, it passes through the translocon channel into the ER lumen. The signal peptidase, whose active site faces the

ER lumen, cleaves the signal peptide as soon as the recognition site enters the lumen. Step 6: The peptide chain continues to elongate as the mRNA is translated toward the $3'$ end. Because the ribosome is attached to the translocon, the growing chain is extruded through the translocon into the ER lumen. Steps 7 – 8: Once translation is complete, the ribosome is released, the remainder of the protein is drawn into the ER lumen, the translocon closes, and the protein assumes its native folded conformation.

What is the purpose of coupling GTP hydrolysis by SRP and SRP receptor to delivery of a signal sequence to the translocon — the channel across the membrane? It is clear that this GTP hydrolysis does not provide the motive energy for ongoing translocation since it occurs only at the initiation of the translocation process. The GTPase cycle depicted in Figure 13-6 appears to be analogous to the GTP hydrolysis coupled to the initiation of *translation* on the ribosome as depicted in Figure 5-36. By coupling to GTP hydrolysis, a complex macromolecular assembly reaction can be made more efficient by adding a proofreading step that can increase the fidelity of the reaction.

Passage of Growing Polypeptides Through the Translocon Is Driven by Translation

Once the SRP and its receptor have targeted a ribosome synthesizing a secretory protein to the ER membrane, the ribosome and nascent polypeptide chain are rapidly transferred to the **translocon**, a

complex of proteins that forms a channel embedded within the ER membrane. As translation continues, the elongating chain passes directly from the large ribosomal subunit into the central pore of the translocon. The large ribosomal subunit is aligned with the pore of the translocon in such a way that the growing chain is never exposed to the cytoplasm and is prevented from folding until it reaches the ER lumen (see Figure 13-6).

The translocon was first identified through mutations in the yeast gene encoding a protein called *Sec61α*, which caused a block in the translocation of secretory proteins into the lumen of the ER. Subsequently, three proteins, collectively called the *Sec61 complex*, were found to form the mammalian translocon: Sec61α, an integral membrane protein with 10 membrane-spanning α helices, and two smaller proteins, termed *Sec61β* and *Sec61γ*. Chemical cross-linking experiments in a cell-free translocation system — in which amino acid side chains from a nascent secretory protein became covalently attached to the Sec61α subunit — demonstrated that the translocating polypeptide chain comes into contact with the Sec61α protein, confirming its identity as the translocon pore (Figure 13-7).

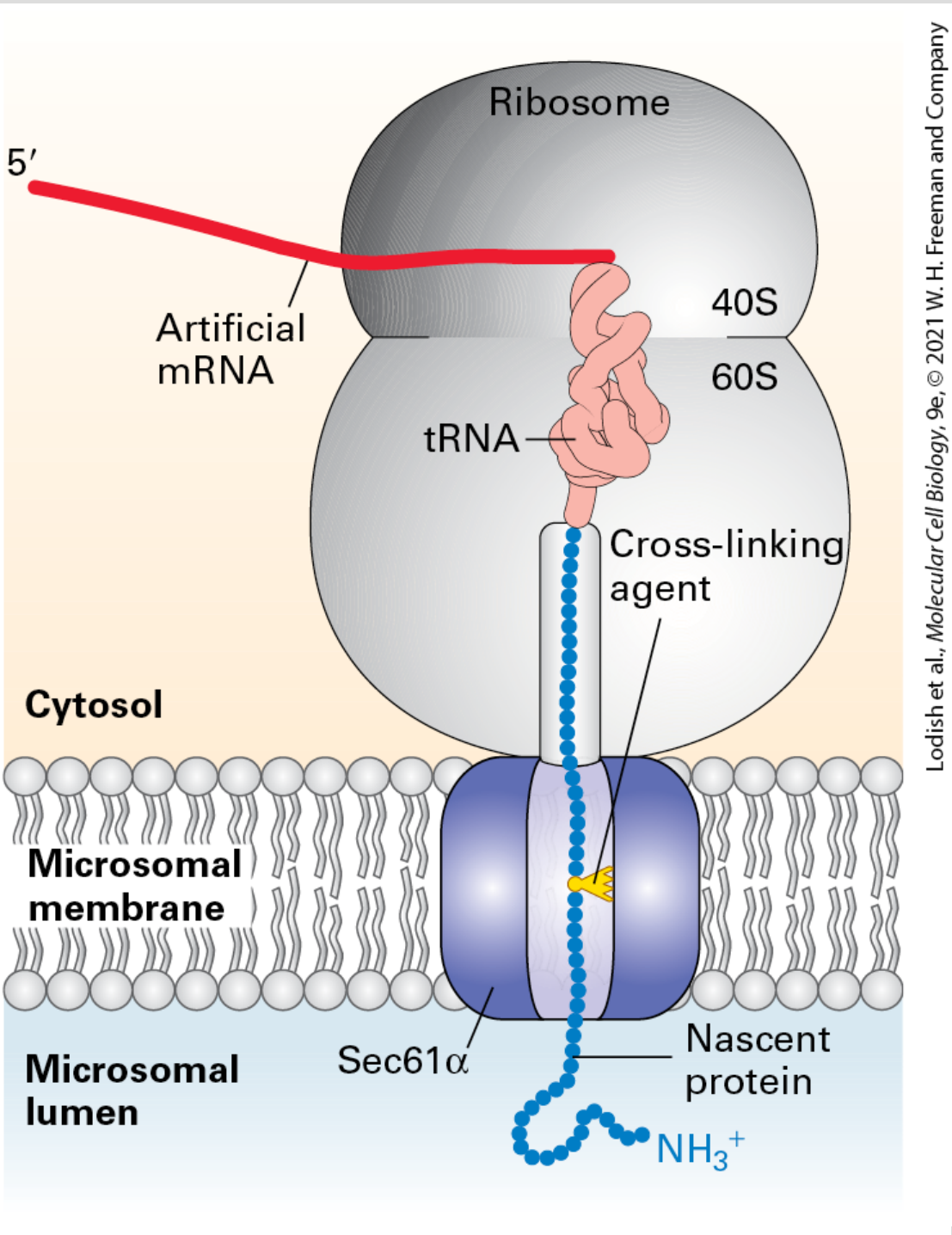


FIGURE 13-7 Sec61α is a translocon component. Experiments using added chemical cross-linkers show that nascent secretory proteins come into close proximity of the translocon component Sec61α as they pass into the ER lumen. An mRNA encoding the N-terminal 70 amino acids of the secreted protein prolactin was translated in a cell-free system containing microsomes (see Figure 13-4b). The mRNA

lacked a chain-termination codon and contained one lysine codon near the middle of the sequence. The reaction mixtures contained a chemically modified lysyl-tRNA in which a light-activated cross-linking reagent was attached to the lysine side chain. Although the entire mRNA was translated, the completed polypeptide could not be released from the ribosome without a chain-termination codon and thus became "stuck" crossing the ER membrane. The reaction mixtures were then exposed to intense light, which caused the nascent polypeptide chain to become covalently bound to whatever proteins were near it in the translocon. When the experiment was performed using microsomes from mammalian cells, the nascent chain became covalently linked to Sec61α. Different versions of the prolactin mRNA were created so that the modified lysine residue would be placed at different distances from the ribosome; cross-linking to Sec61α was observed only when the modified lysine was positioned within the translocation channel. See T. A. Rapoport, 1992, *Science* **258**:931; and D. Görlich and T. A. Rapoport, 1993, *Cell* **75**:615.

When microsomes in the cell-free translocation system were replaced with reconstituted phospholipid vesicles containing only the SRP receptor and the Sec61 complex, nascent secretory proteins were translocated from their SRP-ribosome complex into the vesicles. This finding indicates that the SRP receptor and the Sec61 complex are the only ER-membrane proteins that are absolutely required for translocation. The energy derived from chain elongation at the ribosome appears to provide the driving force to push the polypeptide chain across the membrane in one direction.

The translocon must be able to allow passage of a polypeptide chain while remaining sealed to small molecules, such as ATP, in order to maintain the permeability barrier of the ER membrane. Furthermore, there must be some way to regulate the translocon so that it is closed in its default state, opening only when a nascent

polypeptide chain–ribosome complex is bound. A high-resolution structural model of the archaeal Sec61 complex shows how the translocon preserves the integrity of the membrane (Figure 13-8). The 10 transmembrane helices of Sec61α form a central channel through which the translocating polypeptide chain passes. Two different gating steps are required for Sec61α to accept a translocating polypeptide. The 10 transmembrane helices are organized into two 5-helix bundles. In the first gating step, the bundles hinge apart like an opening clamshell to expose a hydrophobic binding pocket for the hydrophobic core of the signal sequence at the open edge. The signal sequence binds to Sec61α with its N-terminus facing the cytosol and the elongating polypeptide doubling back through the central channel. The structural model of the Sec61 complex, which was isolated without a translocating peptide and is therefore presumed to be in a closed conformation, reveals a short helical peptide plugging the central channel. Biochemical studies of the Sec61 complex have shown that, in the absence of a translocating polypeptide, the peptide that forms the plug effectively seals the translocon to prevent the passage of ions and small molecules. In the second gating step, after the signal sequence has bound to the opened channel, the translocating peptide enters the central pore of the channel, forcing away the plug peptide and allowing translocation to proceed. The middle of the central pore is lined with hydrophobic isoleucine residues that in effect form a gasket, preventing leakage of small polar molecules around the translocating peptide even as translocation proceeds.

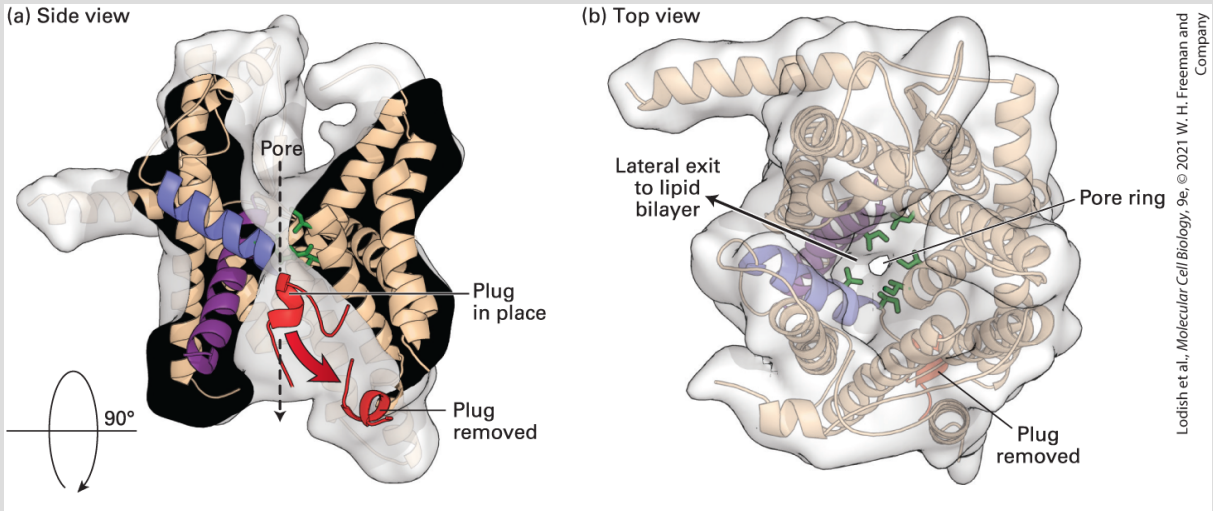


FIGURE 13-8 Structure of an archaeal Sec61 complex. The structure of the detergent-solubilized Sec61 complex from the archaeon *Methanocaldococcus jannaschii* (also known as the *SecY complex*) was determined by x-ray crystallography. (a) A side view shows the hourglass-shaped channel through the center of the pore. A ring of isoleucine residues at the constricted waist of the pore forms a gasket that keeps the channel sealed to small molecules even as a translocating polypeptide passes through the channel. When no translocating peptide is present, the channel is closed by a short helical plug (red). This plug moves out of the channel during translocation. In this view, the front half of protein has been removed to better show the pore. (b) A view looking through the center of the channel shows a region (on the left side) where helices may separate, allowing lateral passage of a hydrophobic transmembrane domain into the lipid bilayer.

[Data from B. van den Berg et al., 2004, *Nature* **427**:36–44, PDB ID 1rhz and 1rh5.]

As the growing polypeptide chain enters the lumen of the ER, the signal sequence is cleaved by *signal peptidase*, which is a transmembrane ER protein associated with the translocon (see Figure 13-6, step **5**). Signal peptidase recognizes a sequence on the C-terminal end of the hydrophobic core of the signal peptide and cleaves the chain specifically at this sequence once it has emerged into the luminal space of the ER. After the signal sequence has been

cleaved, the growing polypeptide moves through the translocon into the ER lumen. The translocon remains open until translation is complete and the entire polypeptide chain has moved into the ER lumen. After translocation is complete, the plug peptide reseals the translocon channel.

ATP Hydrolysis Powers Post-Translational Translocation of Some Secretory Proteins in Yeast

In most eukaryotes, secretory proteins enter the ER by cotranslational translocation. In yeast, however, some secretory proteins enter the ER lumen after translation has been completed. In such **post-translational translocation**, the translocating protein passes through the same Sec61 translocon that is used in cotranslational translocation. However, the SRP and SRP receptor are not involved in post-translational translocation, and in such cases a direct interaction between the translocon and the signal sequence of the completed protein appears to be sufficient for targeting to the ER membrane. In addition, the driving force for unidirectional translocation across the ER membrane is provided by an additional protein complex known as the *Sec63 complex* and a member of the Hsp70 family of molecular chaperones known as *BiP* (see Chapter 3 for further discussion of molecular chaperones). The tetrameric Sec63 complex is embedded in the ER membrane in the

vicinity of the translocon, whereas BiP is within the ER lumen. Like other members of the Hsp70 family, BiP has a peptide-binding domain and an ATPase domain. These chaperones bind and stabilize unfolded or partially folded proteins (see Figure 3-17).

The current model for post-translational translocation of a protein into the ER is outlined in Figure 13-9. Once the N-terminal segment of the protein enters the ER lumen, signal peptidase cleaves the signal sequence just as in cotranslational translocation (step **1**). Interaction of BiP·ATP with the luminal portion of the Sec63 complex causes hydrolysis of the bound ATP, producing a conformational change in BiP that promotes its binding to an exposed polypeptide chain (step **2**). Since the Sec63 complex is located near the translocon, BiP is thus activated at sites where nascent polypeptides can enter the ER. Certain experiments suggest that, in the absence of binding to BiP, an unfolded polypeptide can freely slide back and forth within the translocon channel. Such random sliding motions rarely result in the entire polypeptide's crossing the ER membrane. Binding of a molecule of BiP·ADP to the luminal portion of the polypeptide prevents backsliding of the polypeptide out of the ER. As further inward random sliding exposes more of the polypeptide on the luminal side of the ER membrane, successive binding of BiP·ADP molecules to the polypeptide chain acts as a ratchet, ultimately drawing the entire polypeptide into the ER within a few seconds (steps **3** and **4**). After a delay of some time, the BiP molecules spontaneously exchange their bound ADP for ATP, leading to release of the polypeptide, which can then fold

into its native conformation (steps 5 and 6). The recycled BiP·ATP is then ready for another interaction with Sec63. BiP and the Sec63 complex are also required for cotranslational translocation. The details of their role in this process are not well understood, but they are thought to act at an early stage of the process, such as the threading of the signal peptide into the pore of the translocon.

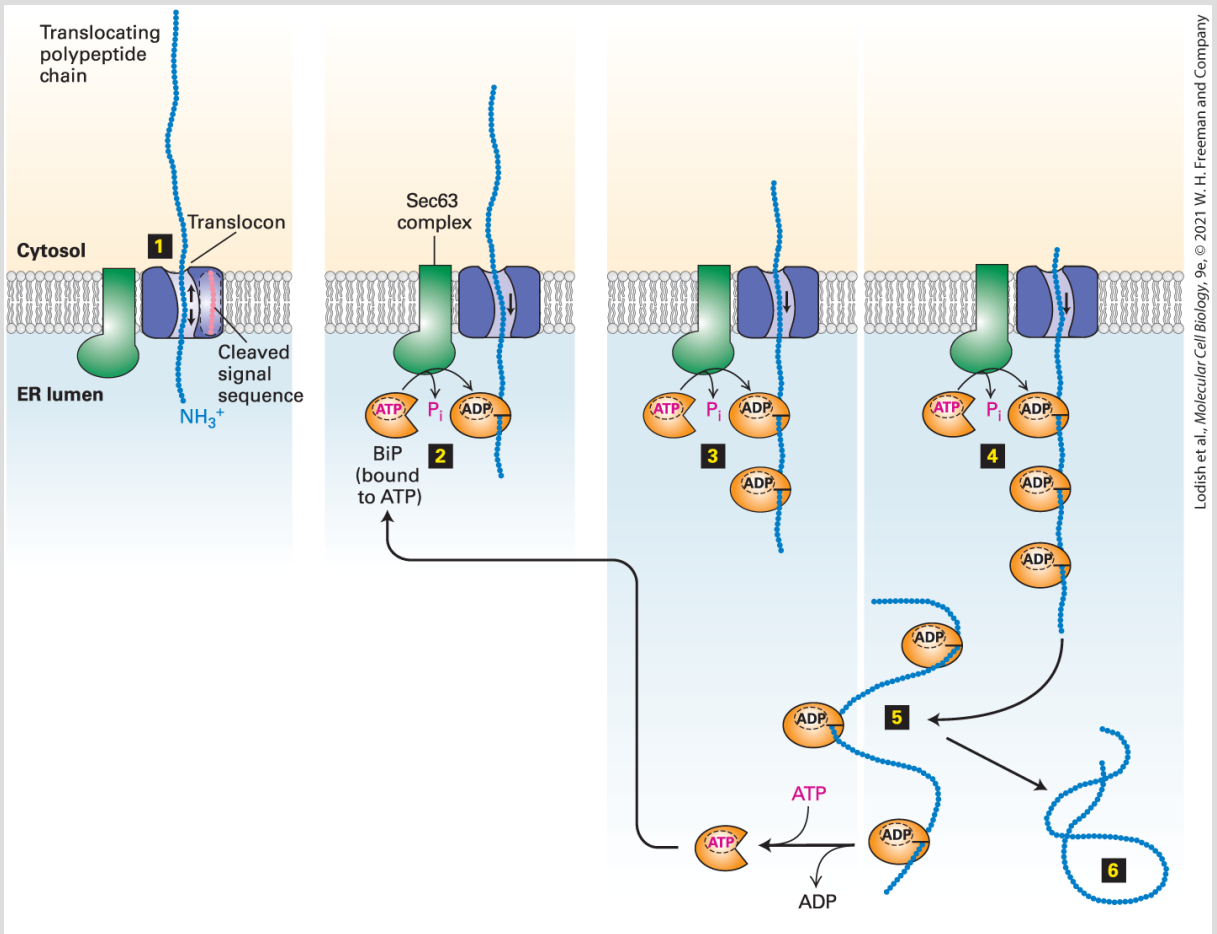


FIGURE 13-9 Post-translational translocation. This mechanism is fairly common in yeast and probably occurs occasionally in higher eukaryotes. Step 1: For yeast proteins that can be translocated post-translationally, the signal sequence of the fully translated protein can engage with the translocon, causing the signal sequence and N-terminal portion of the protein to enter the ER where the signal sequence is cleaved. Small arrows inside the

translocon represent the ability of the translocating polypeptide to slide randomly inward and outward. Step **2**: Within the lumen of the ER, BiP·ATP is converted into BiP·ADP by the Sec63 complex. BiP·ADP has a high affinity for binding to exposed hydrophobic segments of the translocating polypeptide. Once BiP·ADP is bound, the polypeptide is no longer free to slide out toward the cytosol but can still slide inward as shown by the inward pointing arrow. Steps **3**–**5**: BiP·ADP binds to successive segments of the polypeptide as they enter the ER, progressively ratcheting the polypeptide inward until it has fully entered the ER lumen. Step **6**: Relatively slow exchange of ATP for ADP regenerates BiP·ATP, which releases the translocated peptide, allowing it to complete folding in the ER lumen. See K. E. Matlack et al., 1997, *Science* **277**:938.

The overall reaction carried out by BiP is an important example of how the chemical energy released by the hydrolysis of ATP can power the mechanical movement of a protein across a membrane. As we will see, translocation of proteins into mitochondria and chloroplasts can be powered by ATP hydrolysis by ATPase chaperone proteins located inside these organelles and also typically occurs by post-translational translocation. This explains why ribosomes are typically not found bound to these other organelles, as they are to the rough ER.

KEY CONCEPTS OF SECTION 13.1

Targeting Proteins to and Across the ER Membrane

- Synthesis of secreted proteins, integral plasma-membrane proteins, and proteins destined for the ER, Golgi complex, or lysosome begins on cytosolic ribosomes, which become attached to the membrane of the ER, forming the rough ER (see Figure 13-1, *left*).
- The ER signal sequence on a nascent secretory protein is located at the N-terminus and contains a sequence of hydrophobic amino acids.

- In cotranslational translocation, the signal recognition particle (SRP) first recognizes and binds the ER signal sequence on a nascent secretory protein, then is bound in turn by an SRP receptor on the ER membrane, thereby targeting the nascent polypeptide chain–ribosome complex to the ER.
- The SRP and SRP receptor then mediate insertion of the nascent secretory protein into the translocon (Sec61 complex). Hydrolysis of two molecules of GTP by the SRP and its receptor cause the dissociation of SRP (see Figures 13-5 and 13-6). Coupling GTP hydrolysis to the loading of signal-sequence bearing proteins to the translocon is thought to increase the fidelity of this assembly reaction.
- As the ribosome attached to the translocon continues translation, the unfolded protein chain is extruded into the ER lumen.
- The translocon contains a central channel lined with hydrophobic residues that allows transit of an unfolded protein chain while remaining sealed to ions and small hydrophilic molecules. In addition, the channel is gated so that it is open only when a polypeptide is being translocated.
- In post-translational translocation, a completed secretory protein is targeted to the ER membrane by interaction of the signal sequence with the translocon. The polypeptide chain is then pulled into the ER by a ratcheting mechanism that requires ATP hydrolysis by the chaperone BiP, which stabilizes the entering polypeptide (see Figure 13-9).
- In both cotranslational and post-translational translocation, a signal peptidase in the ER membrane cleaves the ER signal sequence from a secretory protein soon after the N-terminus enters the lumen.

13.2 Insertion of Membrane Proteins into the ER

In previous chapters, we have encountered many of the vast array of integral membrane (transmembrane) proteins that are present throughout the cell. Each such protein has a unique orientation with respect to the membrane's phospholipid bilayer. Integral membrane proteins located in the ER, Golgi complex, and lysosomes, as well as in the plasma membrane, which are all synthesized on the rough ER, remain embedded in the membrane as they move to their final destinations along the same pathway that is followed by soluble secretory proteins (see Figure 13-1, *left*). During this transport, the orientation of a membrane protein is preserved; that is, the same segments of the protein always face the cytoplasmic side of the membrane, whereas other segments always face in the opposite direction, which could be considered the exoplasmic side of the membrane. Thus the final orientation of these membrane proteins is established during their biosynthesis on the ER membrane. As we will see in this section, insertion of membrane proteins into the ER membrane employs the same mechanism that has already been described for entry of soluble secretory proteins into the lumen of the ER. Membrane proteins first engage the translocon, made up of the Sec61 complex, by the interaction of a hydrophobic signal sequence with SRP and SRP receptor. Once engaged with the translocon, hydrophilic portions of the membrane protein can enter the ER lumen by translocation, whereas the hydrophobic portions

become embedded in the ER membrane by lateral opening of the translocon gate. The final topological orientation of membrane segments and hydrophilic regions of a membrane protein depends on the arrangement of hydrophobic sequences and neighboring positively charged amino acid residues. These sequences, known collectively as **topogenic sequences**, direct the membrane insertion and orientation of various classes of integral membrane proteins.

Several Topological Classes of Integral Membrane Proteins Are Synthesized on the ER

The **topology** of a membrane protein refers to the number of times its polypeptide chain spans the membrane and the orientation of those membrane-spanning segments within the membrane. The key elements of a protein that determine its topology are the membrane-spanning segments themselves, which are usually α helices containing 20–25 hydrophobic amino acids that contribute to energetically favorable interactions within the hydrophobic interior of the phospholipid bilayer.

Most integral membrane proteins fall into one of the five topological classes illustrated in Figure 13-10. Topological classes I, II, III, and the tail-anchored proteins are **single-pass membrane proteins**, which have only one membrane-spanning α-helical segment. Type I

proteins have a cleaved N-terminal ER signal sequence and are anchored in the membrane with their hydrophilic N-terminal region on the luminal face (also known as the *exoplasmic face*) and their hydrophilic C-terminal region on the cytosolic face. Type II proteins do not contain a cleavable ER signal sequence and are oriented with their hydrophilic N-terminal region on the cytosolic face and their hydrophilic C-terminal region on the exoplasmic face (i.e., opposite to type I proteins). Type III proteins have a hydrophobic membrane-spanning segment at their N-terminus and thus have the same orientation as type I proteins, but do not contain a cleavable signal sequence. Finally, tail-anchored proteins have a hydrophobic segment at their C-terminus that spans the membrane. These different topologies reflect distinct mechanisms used by the cell to establish the orientation of transmembrane segments, as we will see shortly.

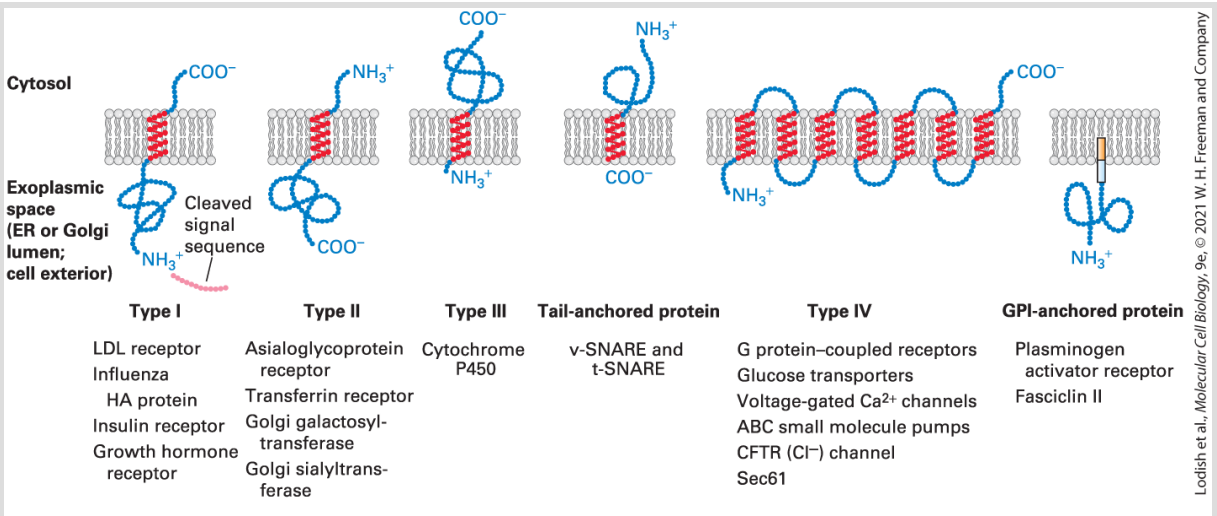


FIGURE 13-10 Classes of ER membrane proteins. Five topological classes of integral membrane proteins are synthesized on the rough ER, as is a sixth type tethered to the membrane by a phospholipid anchor. These membrane proteins are classified by their

orientation in the membrane and the types of signals they contain to direct them there. In the integral membrane proteins, hydrophobic segments of the protein chain form α helices embedded in the membrane bilayer; the regions outside the membrane are hydrophilic and fold into various conformations. All type IV proteins have multiple transmembrane α helices. The type IV topology depicted here corresponds to that of G protein–coupled receptors: seven α helices, the N-terminus on the exoplasmic side of the membrane, and the C-terminus on the cytosolic side. Other type IV proteins may have a different number of helices and various orientations of the N-terminus and C-terminus. See E. Hartmann et al., 1989, *P. Natl. Acad. Sci. USA* **86**:5786; and C. A. Brown and S. D. Black, 1989, *J. Biol. Chem.* **264**:4442.

The proteins forming topological type IV contain two or more membrane-spanning segments and are sometimes called **multipass membrane proteins**. For example, many of the membrane transport proteins discussed in Chapter 11 and the numerous G protein–coupled receptors covered in Chapter 15 belong to this class.

When considering the mechanism of assembly of integral membrane proteins it is often useful to consider the topology of an individual membrane spanning segment. An N_{cyto}-C_{exo} transmembrane segment is oriented with the N-terminal end facing the inside of the cytoplasm and the C-terminal end facing exoplasmic side of the membrane that would be the lumen of the ER or the topological equivalent that is the outside of the cell. The other possible orientation is N_{exo}-C_{cyto}. Thus the transmembrane span of a type I membrane protein is N_{exo}-C_{cyto}, whereas the transmembrane span of a type II protein is N_{cyto}-C_{exo}. Note that for a type IV protein

with multiple spans, the orientation strictly alternates between N_{cyto}-C_{exo} and N_{exo}-C_{cyto}.

Some lipid-anchored membrane proteins are also synthesized on the ER. These membrane proteins lack a hydrophobic membrane-spanning segment altogether; instead, they are linked to an amphipathic phospholipid anchor that is embedded in the membrane (Figure 13-10, *right*).

Internal Stop-Transfer Anchor and Signal-Anchor Sequences Determine Topology of Single-Pass Proteins

We begin our discussion of how membrane protein topology is determined with the insertion of integral membrane proteins that contain a single hydrophobic membrane-spanning segment. As we will see, three main types of topogenic sequences are used to direct proteins to the ER membrane and to orient them within it. We have already introduced one, the N-terminal signal sequence. The other two types of sequences, which we will introduce here, are internal sequences known as *stop-transfer anchor sequences* and *signal-anchor sequences*. Unlike signal sequences, these two types of internal topogenic sequences end up in the mature protein as membrane-spanning segments. However, the two types differ in their final orientation in the membrane.

Type I Proteins

In addition to an N-terminal signal sequence that targets them to the ER, all type I transmembrane proteins possess an internal hydrophobic sequence of approximately 22 amino acids, which becomes the membrane-spanning α helix. The N-terminal signal sequence of a nascent type I protein, like that of a soluble secretory protein, initiates cotranslational translocation of the protein through the combined action of the SRP and SRP receptor. Once the N-terminus of the growing polypeptide enters the lumen of the ER, the signal sequence is cleaved, and the growing polypeptide chain continues to be extruded across the ER membrane. However, when the sequence that will become a transmembrane domain enters the translocon, transfer of the protein through the channel stops and the transmembrane domain is allowed to move laterally from the channel into the membrane (Figure 13-11). The gating mechanism that allows lateral movement is the same as that for the opening of the translocon to accept the hydrophobic core of a signal sequence: two five-helix bundles of Sec61α hinge open to allow the hydrophobic transmembrane segment to move laterally past the hydrophobic signal-sequence-binding site through the opened edge of the translocon (see Figure 13-8). When the peptide exits the translocon in this manner, the hydrophobicity of the transmembrane segment anchors it in the hydrophobic interior of the membrane. Because such a sequence functions both to stop passage of the polypeptide chain through the translocon and to

become a hydrophobic transmembrane segment in the membrane bilayer, it is called a [stop-transfer anchor sequence (STA)](#).

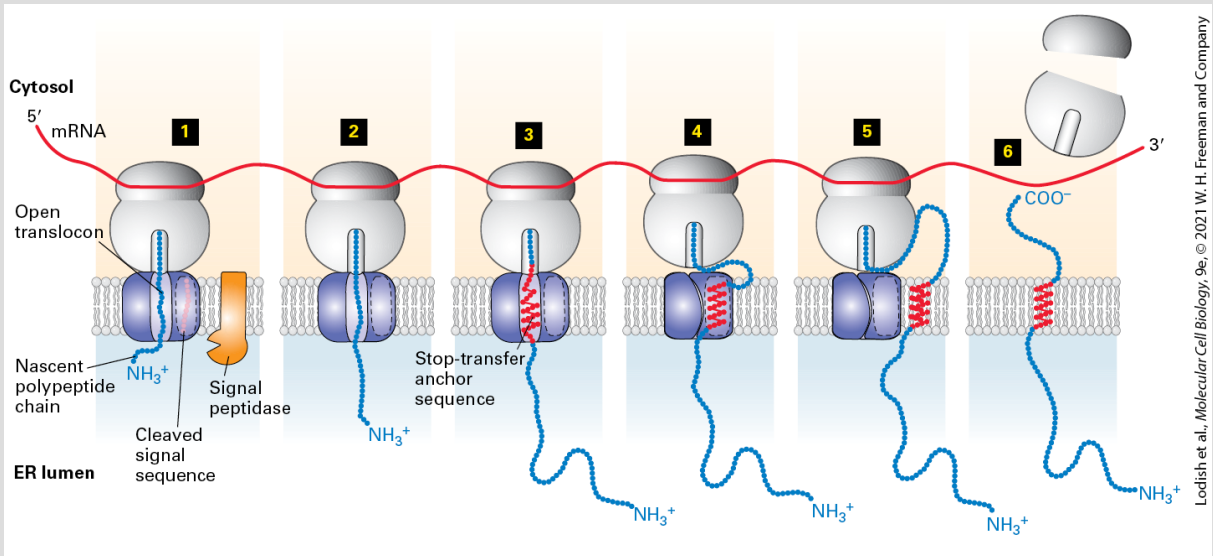


FIGURE 13-11 Membrane insertion and orientation of type I single-pass transmembrane proteins. Step 1: After the nascent polypeptide chain–ribosome complex becomes associated with a translocon in the ER membrane, the N-terminal signal sequence is cleaved. This process occurs by the same mechanism as the one for soluble secretory proteins (see Figure 13-6). Steps 2–3: The chain is elongated until the hydrophobic stop-transfer anchor sequence is synthesized and enters the translocon, where it prevents the nascent chain from extruding farther into the ER lumen. Step 4: The stop-transfer anchor sequence moves laterally through a hydrophobic cleft between translocon subunits and ultimately becomes anchored in the phospholipid bilayer. At this time, the translocon probably closes. Step 5: As synthesis continues, the elongating chain may loop out into the cytosol through the small space between the ribosome and translocon. Step 6: When synthesis is complete, the ribosomal subunits are released into the cytosol, leaving the protein free to diffuse laterally in the membrane. See H. Do et al., 1996, *Cell* **85**:369; and W. Mothes et al., 1997, *Cell* **89**:523.

Once translocation is interrupted, translation continues at the ribosome, which is still anchored to the now unoccupied and closed

translocon. As the C-terminus of the protein chain is synthesized, it loops out on the cytosolic side of the membrane. When translation is complete, the ribosome is released from the translocon, and the C-terminus of the newly synthesized type I protein remains in the cytosol. Ultimately, the stop-transfer anchor sequence resides in the membrane in an N_{exo}-C_{cyto} orientation.

Support for this mechanism has come from studies in which cDNAs encoding various mutant receptors for human growth hormone (HGH) were expressed in cultured mammalian cells. The wild-type HGH receptor, a typical type I protein, is normally transported to the plasma membrane. However, a mutant receptor that has charged residues inserted into the single membrane-spanning segment, or that is missing most of this segment altogether, is translocated entirely into the ER lumen and is eventually secreted from the cell as a soluble protein. These kinds of experiments have established that the hydrophobic membrane-spanning segment of the HGH receptor, and of other type I proteins, functions both as a stop-transfer sequence and as a membrane anchor that prevents the C-terminus of the protein from crossing the ER membrane.

Type II and Type III Proteins

Unlike type I proteins, type II and type III proteins lack a cleavable N-terminal ER signal sequence. Instead, both possess a single internal hydrophobic **signal-anchor sequence (SA)** that functions as both an ER signal sequence and a membrane anchor. Recall that

type II and type III proteins have opposite orientations in the membrane (see Figure 13-10); this difference depends on the orientation that their respective signal-anchor sequences assume within the translocon. The internal signal-anchor sequence in type II proteins directs insertion of the nascent polypeptide chain into the ER membrane so that the N-terminus of the chain faces the cytosol, using the same SRP-dependent mechanism described for signal sequences (Figure 13-12a). However, the internal signal-anchor sequence does not have a recognition sequence for signal peptidase and therefore is *not* cleaved. Because of its hydrophobicity, the signal-anchor sequence can move laterally from the signal-sequence-binding site at the edge of the translocon directly into the phospholipid bilayer, where it functions as a membrane anchor. As elongation continues, the C-terminal region of the growing chain is extruded through the translocon into the ER lumen by cotranslational translocation.

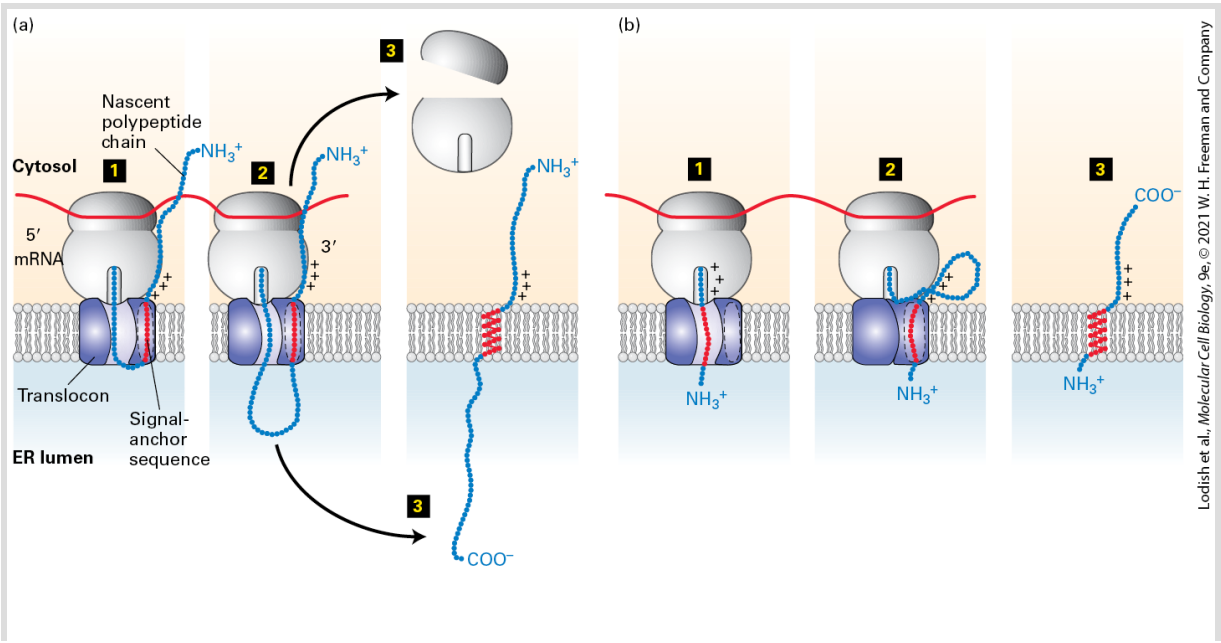


FIGURE 13-12 Membrane insertion and orientation of type II and type III single-pass transmembrane proteins. (a) Type II proteins. Step 1: After the internal signal-anchor sequence is synthesized on a cytosolic ribosome, it is bound by an SRP (not shown), which binds the SRP receptor on the ER membrane. This process is similar to the targeting of soluble secretory proteins except that the hydrophobic signal sequence is not located at the N-terminus and is not subsequently cleaved. The nascent polypeptide chain becomes oriented in the translocon with its N-terminal portion toward the cytosol. This orientation is dictated by the positively charged residues shown N-terminal to the signal-anchor sequence. Step 2: As the chain is elongated and extruded into the lumen, the internal signal-anchor sequence moves laterally through a hydrophobic cleft between translocon subunits and anchors the chain in the phospholipid bilayer. Step 3: Once protein synthesis is complete, the C-terminus of the polypeptide is released into the lumen and the ribosomal subunits are released into the cytosol. (b) Type III proteins. Step 1: Insertion is by a process similar to that of type II proteins, except that positively charged residues on the C-terminal side of the signal-anchor sequence cause the transmembrane segment to be oriented within the translocon with its C-terminal portion toward the cytosol and the N-terminal end in the ER lumen. Note that when the signal-anchor segment engages with the translocon, the preceding hydrophilic segment is short enough to spontaneously pass through the translocon channel. Steps 2–3: Elongation of the C-terminal portion of the polypeptide chain is completed in the cytosol, and the ribosomal subunits are released. See M. Spiess and H. F. Lodish, 1986, *Cell* **44**:177; and H. Do et al., 1996, *Cell* **85**:369.

In the case of type III proteins, the signal-anchor sequence, which is located near the N-terminus, directs insertion of the nascent chain into the ER membrane with its N-terminus facing the lumen, in an orientation opposite to that of the signal-anchor in type II proteins. The signal-anchor sequence of type III proteins also functions like a stop-transfer sequence and prevents further extrusion of the elongating chain into the ER lumen (Figure 13-12b). Continued elongation of the chain C-terminal to the signal-anchor sequence proceeds as it does for type I proteins, with the hydrophobic

sequence eventually moving laterally out of the translocon to anchor the polypeptide in the ER membrane (see Figure 13-11).

The key difference between type II and type III proteins is the orientation of the hydrophobic transmembrane segment as it binds to the hydrophobic signal-sequence-binding site at the edge of Sec61α. The membrane span of a type II protein is oriented N_{cyto}-C_{exo} whereas the span of a type III protein is N_{exo}-C_{cyto}. The most important feature that determines whether a protein with a single transmembrane span will assume a type II or type III orientation is the length of the hydrophilic sequence that precedes the transmembrane span. If the hydrophilic segment is more than a few amino acids in length, the energetic cost of transferring this segment across the membrane is too great for the transmembrane segment to assume an N_{exo}-C_{cyto} orientation of a type III protein. The orientation of signal-anchor sequences that are preceded by only a few hydrophilic amino acids is determined by the arrangement of positively charged amino acids adjacent to one end or the other of the hydrophobic segment. These positively charged residues tend to remain on the cytosolic side of the membrane, rather than traversing the membrane into the ER lumen. Thus the position of the charged residues dictates the orientation of the signal-anchor sequence within the translocon as well as whether the rest of the polypeptide chain continues to pass into the ER lumen: type II proteins tend to have positively charged residues on the N-terminal side of their signal-anchor sequence, orienting the N-

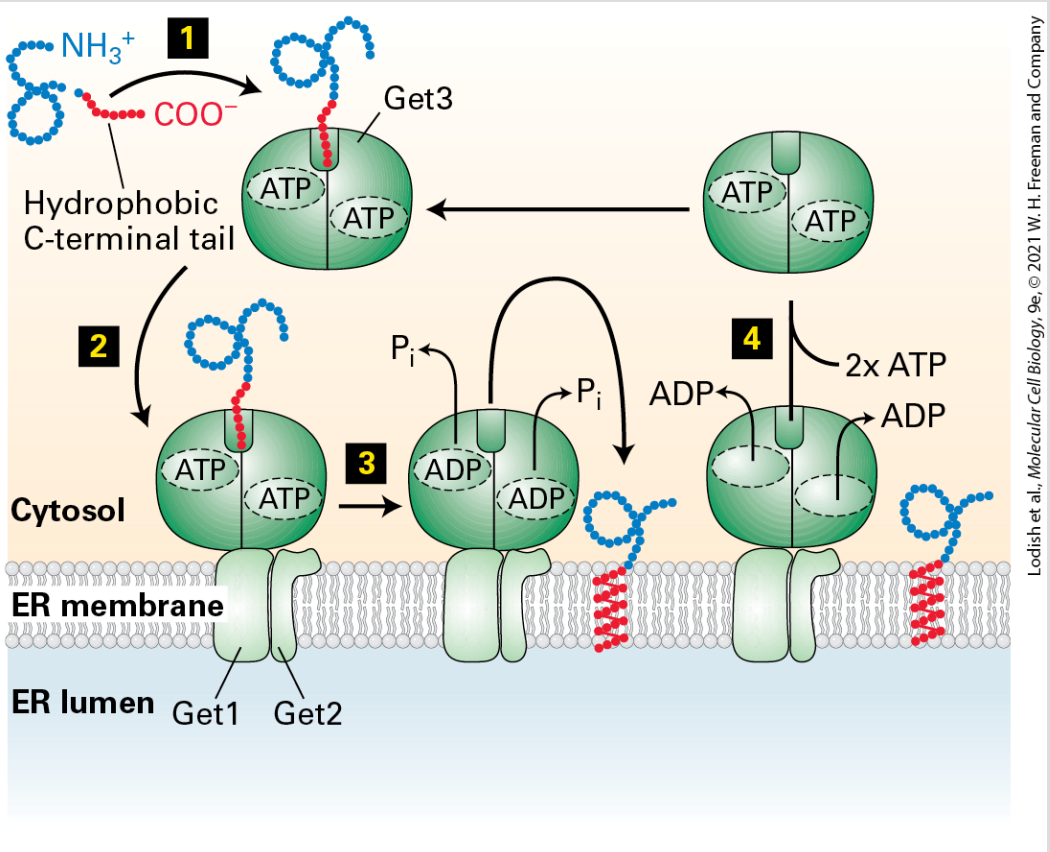
terminus in the cytosol and allowing passage of the C-terminal side into the ER (see Figure 13-12a), whereas type III proteins tend to have positively charged residues on the C-terminal side of their signal-anchor sequence, which restrict the C-terminus to the cytosol (see Figure 13-12b). Note that the hydrophobic segment of a type II signal-anchor sequence assumes the same N_{cyto}-C_{exo} orientation as the signal sequence of a secreted protein and that in most respects these signal-anchor sequences behave exactly like signal sequences, although they are not cleaved.

A striking experimental demonstration of the importance of the flanking charge in determining orientation in the membrane is provided by neuraminidase, a type II protein in the surface coat of the influenza virus. Three arginine residues are located just N-terminal to the internal signal-anchor sequence in neuraminidase. Mutation of these three positively charged residues to negatively charged glutamate residues causes neuraminidase to acquire the reverse orientation. Similar experiments have shown that other proteins, with either type II or type III orientation, can be made to "flip" their orientation in the ER membrane by mutating charged residues that flank the internal signal-anchor segment.

Tail-Anchored Proteins

For all the topological classes of proteins we have considered so far, membrane insertion begins when the SRP recognizes a hydrophobic

topogenic sequence as it emerges from the ribosome. Recognition of tail-anchored proteins, which have a single hydrophobic topogenic sequence at the C-terminus, presents a unique challenge because the hydrophobic C-terminus becomes available for recognition only after translation has been completed and the protein has been released from the ribosome. Insertion of tail-anchored proteins into the ER membrane does not employ an SRP, SRP receptor, or the translocon, but instead depends on a pathway dedicated to this purpose, as depicted in Figure 13-13. This pathway involves a dimer of an ATPase known as *Get3*, which binds to the C-terminal hydrophobic segment of a tail-anchored protein. The Get3 dimer has two ATP-binding sites at the dimer interface in a structure that is similar to the two GTP-binding sites in the heterodimeric interface between SRP and SRP receptor shown in Figure 13-5b. The complex of a Get3 dimer bound to a tail-anchored protein is recruited to the ER by a dimeric integral membrane receptor known as *Get1/Get2*. The tail-anchored protein is released from Get3, and the transmembrane portion of Get1/Get2 participates in the insertion of the tail-anchor into the ER membrane. This process is mechanistically similar to the targeting of type II and type III signal-anchor sequences to the ER by the SRP and SRP receptor; the underlying structural similarity indicates that the mechanism of targeting is fundamentally related, despite the fact that Get3 couples targeting of tail-anchored proteins to ATP hydrolysis, whereas SRP couples protein targeting to GTP hydrolysis.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-13 Insertion of tail-anchored proteins. For C-terminal tail-anchored proteins, the hydrophobic C-terminus is not available for membrane insertion until protein synthesis is complete and the protein has been released from the ribosome. Step **1**: Dimeric Get3 contains a hydrophobic binding pocket at the interface between dimer subunits and Get3 in an ATP-bound state binds to the hydrophobic C-terminal tail of the protein. This binding reaction is facilitated by a complex of three other proteins, Sgt2, Get4, and Get5, which sequester the hydrophobic C-terminal tail before transferring it to Get3·ATP (not shown). Step **2**: The ternary complex Get3·ATP bound to the protein docks onto the dimeric Get1/Get2 receptor, which is embedded in the ER membrane. Step **3**: In succession, ATP is hydrolyzed and ADP is released from Get3. At the same time, the hydrophobic C-terminal tail is released from Get3 and ultimately becomes embedded in the ER membrane in a process that is facilitated by Get1/Get2. Step **4**: Get3 binds to ATP and Get3·ATP is released from Get1/Get2 in a soluble form, ready for another round of binding to a hydrophobic C-terminal tail.

Type IV (Multipass) Proteins

As described in Chapter 11, many physiologically important proteins, such as channel proteins, membrane transporters and pumps, and some receptor proteins, can contain 12 or more membrane-spanning α helices. Although their three-dimensional structure can be quite complex, the membrane topology of multipass proteins can usually be deduced from the same principles used to predict the topology of single-pass proteins. Figure 13-14 summarizes the arrangements of topogenic sequences in single-pass and multipass transmembrane proteins. Two key principles govern the assembly of the vast majority of multipass proteins: (1) the membrane-spanning segments pass from the translocon into the membrane cotranslationally in the sequential order by which they emerge from the ribosome and (2) the first topogenic segment engages the translocon in an SRP and SRP receptor-dependent manner as for single-pass proteins, whereas all subsequent transmembrane segments engage the translocon independently of SRP. Based on these principles, the first transmembrane segment of a multipass protein acts as a topogenic sequence in the ways that we have already discussed for type I, type II, and type III proteins. Once the orientation of the first transmembrane segment is established (N_{exo}-C_{cyto} as with type I and type III proteins or N_{cyto}-C_{exo} as for type II proteins), each subsequent transmembrane segment assumes the opposite orientation as the one before. Thus an N_{exo}-C_{cyto} transmembrane segment is always followed by an N_{cyto}-C_{exo}

transmembrane segment and an N_{cyto}-C_{exo} transmembrane segment is always followed by an N_{exo}-C_{cyto} transmembrane segment. This strict alternation of orientation is a simple consequence of the fact that the hydrophilic segments cannot move across the membrane unless threaded through the channel of the translocon. If a multipass protein has an *even* number of transmembrane α helices, its N-terminus and C-terminus will be oriented toward the same side of the membrane (Figure 13-14d). Conversely, if a type IV protein has an *odd* number of α helices, its two ends will have opposite orientations (Figure 13-14e).

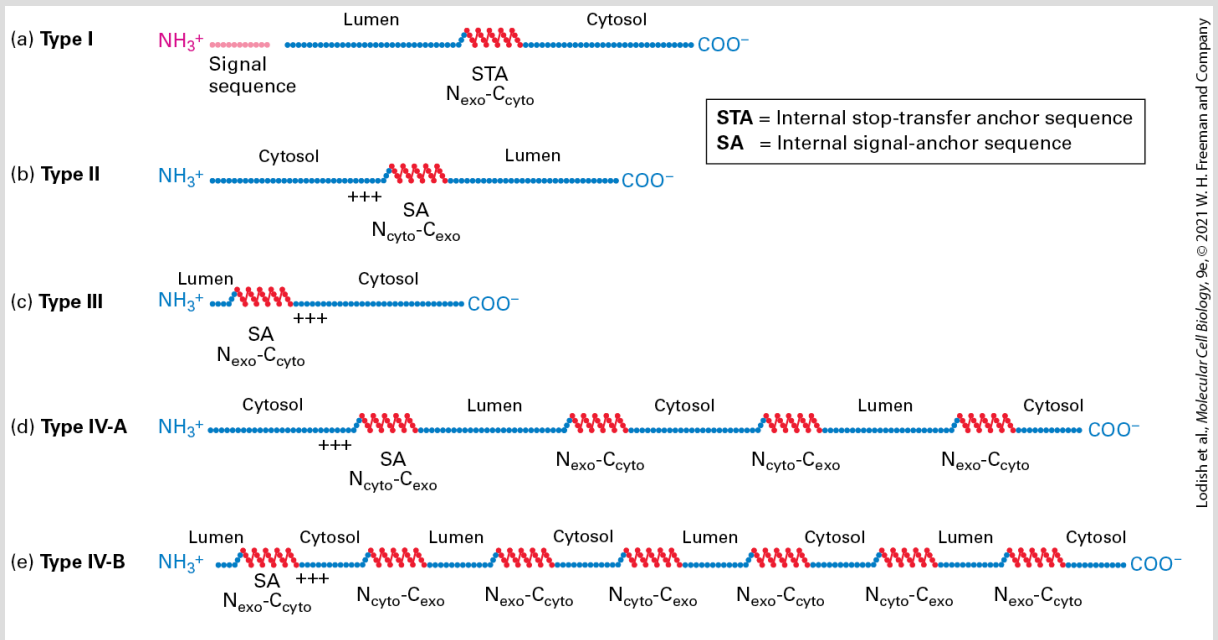


FIGURE 13-14 Topogenic sequences determine the orientation of ER membrane proteins. Topogenic sequences are shown in red; soluble, hydrophilic sequences in blue. The internal topogenic sequences form transmembrane α helices that anchor proteins or segments of proteins in the membrane. (a) Type I proteins contain a cleaved signal sequence and a single internal stop-transfer anchor (STA) sequence. (b, c) Type II and type III proteins contain a single internal signal-anchor (SA) sequence. The difference in the orientation of these protein types depends largely on whether there is a high density of positively charged amino

acids (+ + +) on the N-terminal side (type II) or on the C-terminal side of the SA sequence (type III). (d, e) Nearly all multipass proteins lack a cleavable signal sequence, as depicted in the examples shown here. Type IV-A proteins, whose N-terminus faces the cytosol, begin with an SA sequence similar to a type II protein. Type IV-B proteins, whose N-terminus faces the lumen, begin with an SA sequence similar to a type III protein. After the first topogenic sequence all subsequent transmembrane segments insert into the membrane with the opposite polarity as the segment before.

Type IV Proteins with N-Terminus in the Cytosol

Among the multipass proteins whose N-terminus extends into the cytosol are the various glucose transporters (GLUTs) and most ion-channel proteins, discussed in Chapter 11. In these proteins, the hydrophobic segment closest to the N-terminus functions like the internal signal-anchor sequence of a type II protein and is inserted into the ER membrane with an N_{cyto}-C_{exo} orientation (see Figure 13-12a). As the nascent chain following the transmembrane segment elongates, it moves through the translocon until the second hydrophobic segment is formed. This helix prevents further extrusion of the nascent chain through the translocon; thus its function is similar to that of the stop-transfer anchor sequence in a type I protein (see Figure 13-11).

After synthesis of the first two transmembrane α helices, both ends of the nascent chain face the cytosol, and the loop between them extends into the ER lumen. As the C-terminus of the nascent chain then continues to grow into the cytosol, as it does in synthesis of type III proteins, the ribosome remains attached to the translocon,

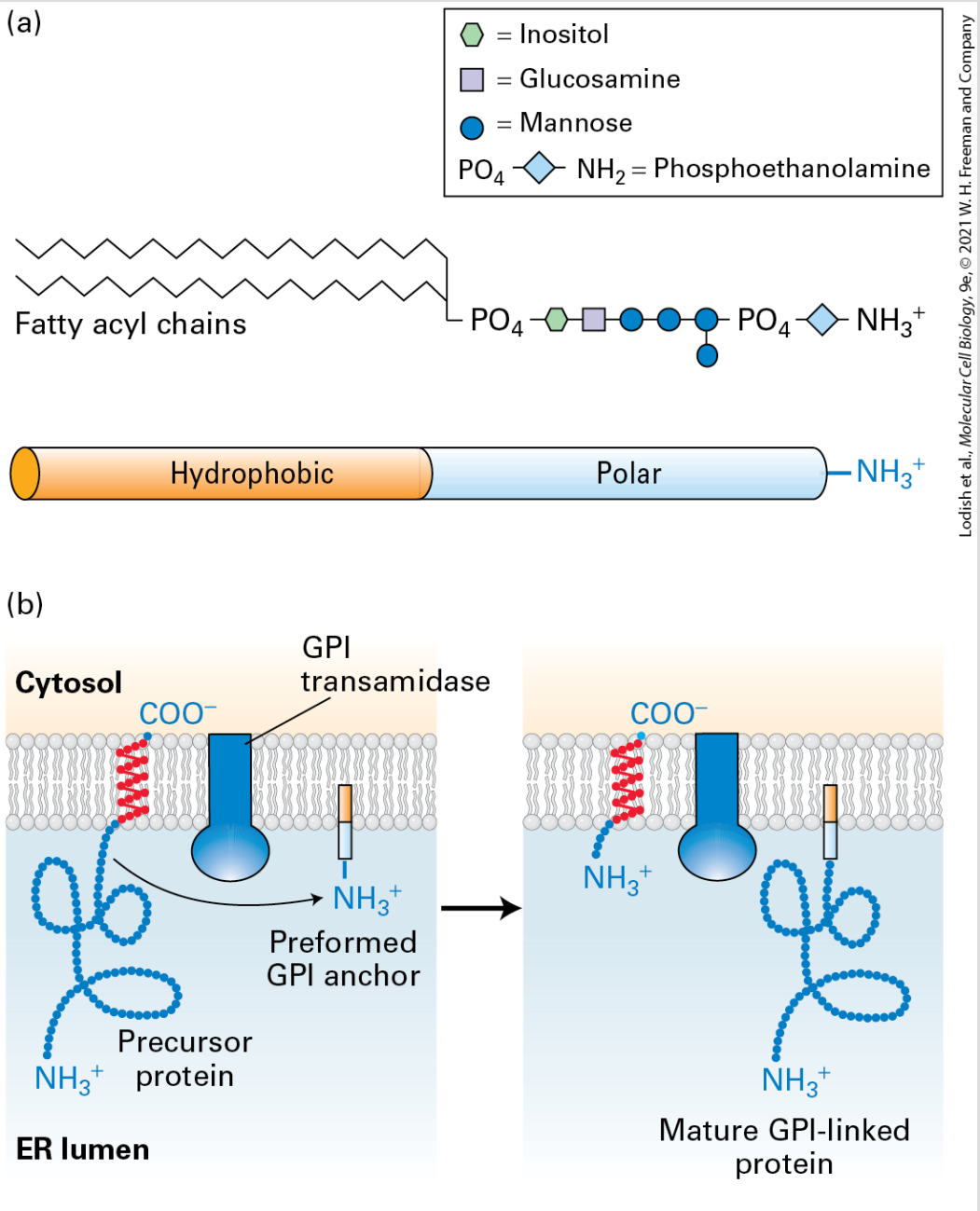
and hydrophobic sequences that subsequently emerge from the ribosome are threaded into the translocon without the need for the SRP and the SRP receptor. As new hydrophobic topogenic sequences engage the translocon, the previously engaged sequences move laterally out of the translocon using the same mechanism as for type I, type II, and type III membrane proteins.

Type IV Proteins with N-Terminus in the Exoplasmic Space

The large family of G protein–coupled receptors, all of which contain seven transmembrane α helices, constitute the most numerous type IV-B proteins, whose N-terminus extends into the exoplasmic space. In these proteins, the hydrophobic transmembrane segment closest to the N-terminus is often followed by a cluster of positively charged amino acids, like a type III signal-anchor sequence (see Figure 13-12b). As a result, the nascent polypeptide chain is inserted into the translocon with the N-terminus extending into the lumen (see Figure 13-14e). As the chain is elongated, it is inserted into the ER membrane in alternating orientation, as just described for type IV-A proteins.

A Phospholipid Anchor Tethers Some Cell-Surface Proteins to the Membrane

Some cell-surface proteins are anchored to the phospholipid bilayer not by a sequence of hydrophobic amino acids, but by a covalently attached amphipathic molecule, *glycosylphosphatidylinositol (GPI)* (Figure 13-15a; see also Chapter 7). These proteins are synthesized and initially anchored to the ER membrane exactly like type I transmembrane proteins, with a cleaved N-terminal signal sequence and an internal stop-transfer anchor sequence directing the process (see Figure 13-11). However, a short sequence of amino acids in the luminal domain, adjacent to the membrane-spanning domain, is recognized by a transamidase located within the ER membrane. This enzyme simultaneously cleaves off the original stop-transfer anchor sequence and transfers the luminal portion of the protein to a preformed GPI anchor in the membrane (Figure 13-15b).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-15 GPI-anchored proteins. (a) Structure of a glycosylphosphatidylinositol (GPI) molecule from yeast. The hydrophobic portion of the molecule is composed of fatty acyl chains, whereas the polar (hydrophilic) portion is composed of carbohydrate residues and phosphate groups. In other organisms, both the length of the acyl chains and the carbohydrate moieties may vary somewhat from the structure shown. (b) Formation of GPI-anchored proteins in the ER membrane. The protein is synthesized and initially inserted into the ER membrane like a type I

transmembrane protein, as shown in Figure 13-11. A specific transamidase simultaneously cleaves the precursor protein within the exoplasmic-facing domain, near the stop-transfer anchor sequence (red), and transfers the carboxyl group of the new C-terminus to the terminal amino group of a preformed GPI anchor. See C. Abeijon and C. B. Hirschberg, 1992, *Trends Biochem. Sci.* **17**:32; and K. Kodukula et al., 1992, *P. Natl. Acad. Sci. USA* **89**:4982.

Why change one type of membrane anchor for another? Attachment of the GPI anchor, which results in removal of the cytosol-facing hydrophilic domain from the protein, can have several consequences. Proteins with GPI anchors, for example, can diffuse relatively rapidly in the plane of the phospholipid bilayer. In contrast, many proteins anchored by membrane-spanning α helices are impeded from moving laterally in the membrane because their cytosol-facing segments interact with the cytoskeleton. In addition, the GPI anchor targets the attached protein to the apical domain of the plasma membrane (instead of the basolateral domain) in certain polarized epithelial cells, as we discuss in Chapter 14.

The Topology of a Membrane Protein Can Often Be Deduced from Its Sequence

As we have seen, various topogenic sequences in integral membrane proteins synthesized on the ER govern the interaction of the nascent polypeptide chain with the translocon. When scientists begin to

study a protein of unknown function, the identification of potential topogenic sequences within the corresponding gene sequence can provide important clues about the protein's topological class and function. Suppose, for example, that the gene for a protein known to be required for a cell-to-cell signaling pathway contains nucleotide sequences that encode an apparent N-terminal signal sequence and an internal hydrophobic sequence. These findings suggest that the protein is a type I integral membrane protein and therefore may be a cell-surface receptor for an extracellular ligand. Furthermore, the implied type I topology suggests that the N-terminal segment that lies between the signal sequence and the internal hydrophobic sequence probably constitutes the extracellular domain with a role in ligand binding, whereas the C-terminal segment that lies after the internal hydrophobic sequence is probably cytosolic and may play a role in intracellular signaling.

Identification of topogenic sequences requires a way to scan sequence databases for segments that are sufficiently hydrophobic to be either signal sequences or transmembrane anchor sequences. Topogenic sequences can often be identified with the aid of computer programs that generate a **hydropathy profile** for the protein of interest. The first step is to assign a value known as the *hydropathic index* to each amino acid in the protein. By convention, hydrophobic amino acids are assigned positive values and hydrophilic amino acids negative values. Although different scales for the hydropathic index exist, all assign the most positive values to

amino acids with side chains made up of mostly hydrocarbon residues (e.g., phenylalanine and methionine) and the most negative values to charged amino acids (e.g., arginine and aspartate). The second step is to identify long segments of sufficient overall hydrophobicity to be N-terminal signal sequences or internal stop-transfer anchor sequences and signal-anchor sequences. To accomplish this, the total hydropathic index for each segment of 20 consecutive amino acids is calculated along the entire length of the protein. Plots of these calculated values against position in the amino acid sequence yield a hydropathy profile.

Figure 13-16 shows the hydropathy profiles for three different membrane proteins. The prominent peaks in such plots identify probable topogenic sequences as well as their positions and approximate lengths. For example, the hydropathy profile of the human growth hormone receptor reveals the presence of both a hydrophobic signal sequence at the extreme N-terminus of the protein and an internal hydrophobic stop-transfer anchor sequence (Figure 13-16a). On the basis of this profile, we can deduce, correctly, that the HGH receptor is a type I integral membrane protein. The hydropathy profile of the asialoglycoprotein receptor, a cell-surface protein that mediates removal of abnormal extracellular glycoproteins, reveals a prominent internal hydrophobic signal-anchor sequence, but gives no indication of a hydrophobic N-terminal signal sequence (Figure 13-16b). Thus we can predict that the asialoglycoprotein receptor is a type II or type III membrane protein. Because the N-terminal segment of the asialoglycoprotein

receptor contains 30 hydrophilic residues that could not move across the membrane without a signal sequence to initiate engagement with the translocon, we can deduce that this segment must remain in the cytosol and we can correctly predict that this is a type II protein.

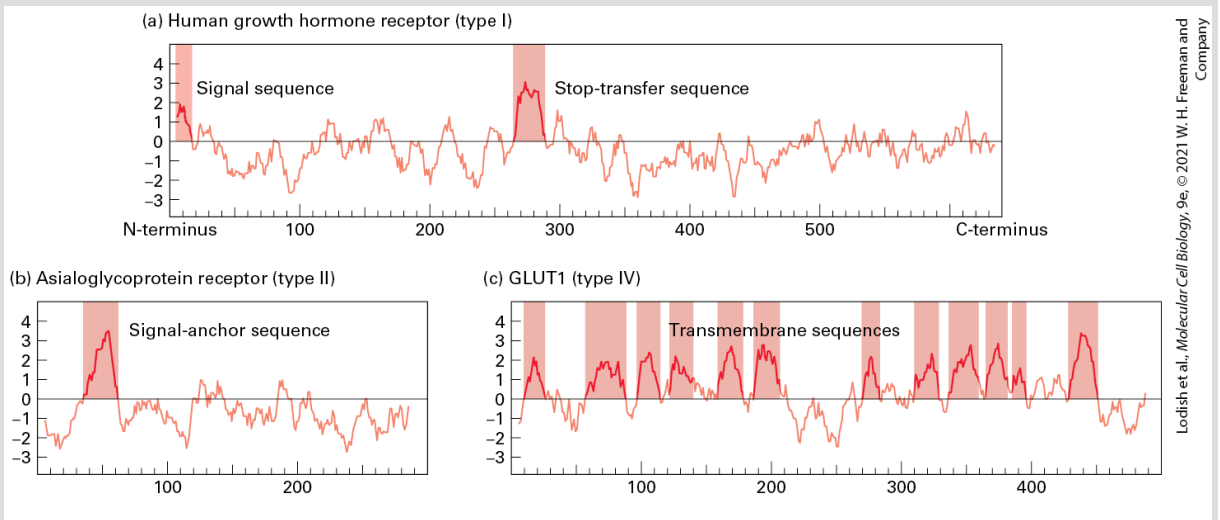


FIGURE 13-16 Hydropathy profiles for three types of proteins. Hydropathy profiles can identify likely topogenic sequences in integral membrane proteins. They are generated by plotting the total hydrophobicity of each segment of 20 contiguous amino acids along the length of a protein. Positive values indicate relatively hydrophobic portions of the protein; negative values relatively hydrophilic portions. Probable hydrophobic topogenic sequences are marked. (a) The human growth hormone receptor evidently has two topogenic sequences, a signal sequence at the N-terminus and an internal hydrophobic stop-transfer sequence. Note that the membrane-spanning stop-transfer sequence is significantly more hydrophobic than the signal sequence, which is typically the case. From this arrangement it would be possible to correctly deduce that this is a type I membrane protein with an N-terminal domain that binds the growth hormone. (b) The asialoglycoprotein receptor has a single hydrophobic topogenic sequence. This sequence can be correctly predicted to be a signal-anchor sequence for a type II protein because it is preceded by a hydrophilic segment of 30 amino acids. As a type II protein, the C-terminal domain is predicted to be the asialoglycoprotein binding domain. The complex profiles for multipass (type IV) proteins,

such as GLUT1 in part (c), must often be supplemented with other analyses to determine the topology of these proteins.

The hydropathy profile of the GLUT1 glucose transporter, a multipass transmembrane protein, shows the presence of many segments that are sufficiently hydrophobic to be membrane-spanning helices (Figure 13-16c). The complexity of this profile illustrates the difficulty both in predicting the orientation of the first topogenic sequence and in unambiguously identifying all the membrane-spanning segments. More sophisticated computer algorithms have been developed that take into account the presence of positively charged amino acids adjacent to hydrophobic segments as well as the length of and spacing between segments. Using all this information, the best algorithms can predict the complex topology of multipass proteins with an accuracy of greater than 80 percent.

Finally, sequence homology to a known protein may permit accurate prediction of the topology of a newly discovered multipass protein. For example, the genomes of multicellular organisms encode a very large number of multipass proteins with seven hydrophobic transmembrane segments. The similarities between the sequences of these proteins strongly suggest that all have the same topology as the well-studied G protein–coupled receptors, which have the N-terminus oriented to the exoplasmic side and the C-terminus oriented to the cytosolic side of the membrane.

KEY CONCEPTS OF SECTION 13.2

Insertion of Membrane Proteins into the ER

- Proteins synthesized on the rough ER include five topological classes of integral membrane proteins as well as a lipid-anchored type protein (see Figure 13-10).
- Topogenic sequences — N-terminal signal sequences, internal stop-transfer anchor sequences, and internal signal-anchor sequences — direct the insertion of nascent proteins into the ER membrane and their orientation within it. This orientation is retained during transport of the completed membrane protein to its final destination — for example, the plasma membrane.
- Single-pass membrane proteins that are initiated by interaction with SRP and are synthesized on ribosomes bound to the translocon are of three topological varieties: type I, type II, and type III. The topological type depends on the presence or absence of a signal sequence, the length of the hydrophilic N-terminal segment, and the arrangement of positively charged residues adjacent to the hydrophobic transmembrane segment (see Figure 13-14).
- Tail-anchored proteins that have a single C-terminal transmembrane domain are inserted into the membrane post-translationally. Their insertion depends on a cytosolic receptor, Get3, which is an ATPase with similarity to the GTPase domains of SRP and SRP receptor.
- The topology of multipass (type IV) proteins is established by the orientation of the first transmembrane segment according to the same principles as for single-pass proteins. Each additional transmembrane segment engages the translocon in sequential order as they emerge from the ribosome and thus are inserted in the membrane in alternating orientation.
- Some cell-surface proteins are initially synthesized as type I proteins, but then cleaved, and their luminal domains transferred to a GPI anchor (see Figure 13-15).
- The topology of membrane proteins can often be correctly predicted by computer programs that identify hydrophobic topogenic segments within the amino acid sequence and generate hydropathy profiles (see Figure 13-16).

13.3 Protein Modifications, Folding, and Quality Control in the ER

Membrane and soluble secretory proteins synthesized on the rough ER undergo four principal modifications before they reach their final destinations: (1) covalent addition and processing of carbohydrates (*glycosylation*) in the ER and Golgi complex; (2) formation of disulfide bonds in the ER; (3) proper folding of polypeptide chains and assembly of multisubunit proteins in the ER; and (4) specific proteolytic cleavages in the ER, Golgi complex, and secretory vesicles. Generally speaking, these modifications promote the folding of secretory proteins into their native structures and add structural stability to proteins exposed to the extracellular environment. Modifications such as glycosylation also allow the cell to produce a vast array of chemically distinct molecules at the cell surface that are the basis of specific molecular interactions used in cell-to-cell adhesion and communication.

The majority of proteins that are synthesized on the rough ER and enter the lumen of the ER are modified by the addition of one or more carbohydrate chains. Proteins with attached carbohydrates are known as **glycoproteins**. Carbohydrate chains in glycoproteins attached to the hydroxyl (—OH) group in serine and threonine residues are referred to as ***O*-linked oligosaccharides**, and carbohydrate chains attached to the amide nitrogen of asparagine are referred to as ***N*-linked oligosaccharides**. The various types of *O*-

linked oligosaccharides include the mucin-type *O*-linked chains (named after the abundant glycoproteins found in mucus) and the carbohydrate modifications on proteoglycans described in Chapter 20. *O*-linked chains typically consist of sugar residues in a linear chain, which are added to proteins by enzymes known as *glycoslytransferases*, located in the lumen of the Golgi complex. The more common *N*-linked oligosaccharides have a branched structure and are initially added to proteins in the ER. In this section, we focus on *N*-linked oligosaccharides, which are synthesized as a branched precursor in the ER. After the initial transfer of *N*-linked precursor to a protein in the ER, the oligosaccharide chain is modified in the ER and commonly in the Golgi complex as well.

Disulfide bond formation, protein folding, and assembly of multimeric proteins, which take place exclusively in the rough ER, are also discussed in this section. Only properly folded and assembled proteins are transported from the rough ER to the Golgi complex and ultimately to the cell surface or other final destination through the secretory pathway. Unfolded, misfolded, or partly folded and assembled proteins are selectively retained in the rough ER and marked for degradation. We consider several features of such quality control in the latter part of this section.

As discussed previously, N-terminal ER signal sequences are cleaved from soluble secretory proteins and type I membrane proteins in the ER. Some proteins also undergo other specific proteolytic

cleavages in the Golgi complex or secretory vesicles. We cover these cleavages, as well as carbohydrate modifications that occur primarily or exclusively in the Golgi complex, in the next chapter.

A Preformed *N*-Linked Oligosaccharide Is Added to Many Proteins in the Rough ER

Biosynthesis of all *N*-linked oligosaccharides begins in the rough ER with a preformed oligosaccharide precursor containing 14 residues (Figure 13-17). The structure of this precursor is the same in plants, animals, and single-celled eukaryotes: a branched oligosaccharide, containing three glucose (Glc), nine mannose (Man), and two *N*-acetylglucosamine (GlcNAc) molecules, which can be written as $Glc_3Man_9(GlcNAc)_2$. Once added to a protein, this branched carbohydrate structure is modified by addition or removal of monosaccharides in the ER and Golgi complex. The modifications to *N*-linked chains differ from one glycoprotein to another and among different organisms, but a core of 5 of the 14 residues is conserved in the structures of all *N*-linked oligosaccharides on secretory and membrane proteins.

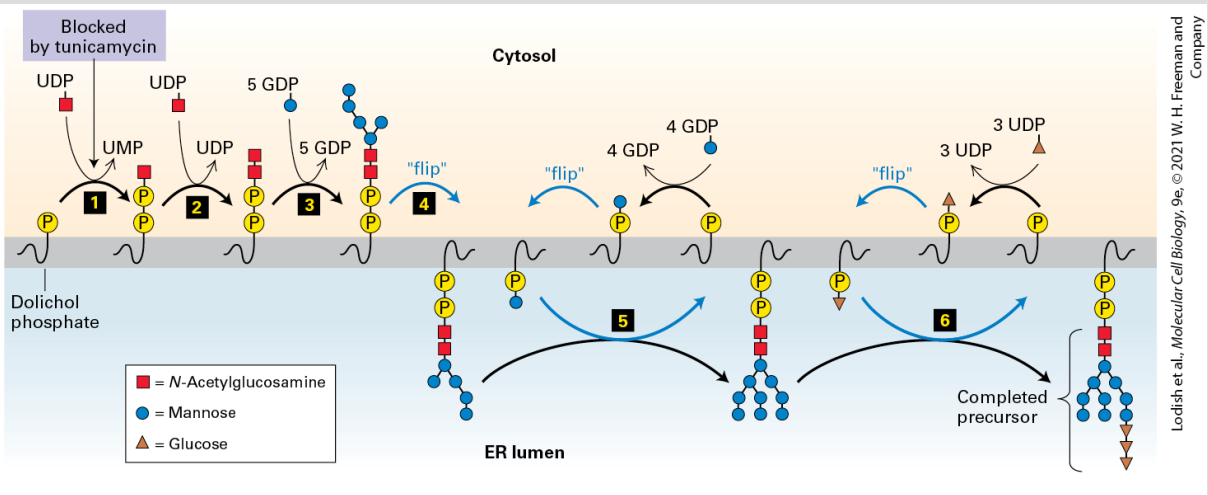


FIGURE 13-17 Biosynthesis of the oligosaccharide precursor. Dolichol phosphate is a strongly hydrophobic lipid, containing 75–95 carbon atoms, that is embedded in the ER membrane. Two *N*-acetylglucosamine (GlcNAc) and five mannose residues are added one at a time to a dolichol phosphate on the cytosolic face of the ER membrane (steps **1**–**3**). The nucleotide-sugar donors in these and later reactions are synthesized in the cytosol. Note that the first sugar residue is attached to dolichol by a high-energy pyrophosphate linkage. Tunicamycin, which blocks the first enzyme in this pathway, inhibits the synthesis of all *N*-linked oligosaccharides in cells. After the seven-residue dolichol pyrophosphoryl intermediate is flipped to the luminal face (step **4**), the remaining four mannose residues and all three glucose residues are added one at a time (steps **5**–**6**). In the later reactions, the sugar to be added is first transferred from a nucleotide sugar to a carrier dolichol phosphate on the cytosolic face of the ER; the carrier is then flipped to the luminal face, where the sugar is transferred to the growing oligosaccharide, after which the "empty" carrier is flipped back to the cytosolic face. See C. Abeijon and C. B. Hirschberg, 1992, *Trends Biochem. Sci.* **17**:32.

Prior to transfer to a nascent chain in the lumen of the ER, the oligosaccharide precursor is assembled on a membrane-attached anchor called **dolichol phosphate**, a long-chain polyisoprenoid lipid (see Chapter 10). After the first sugar, GlcNAc, is attached to the dolichol phosphate by a pyrophosphate bond, the other sugars are added by glycosidic bonds in a complex set of reactions catalyzed by

enzymes attached to the cytosolic or luminal faces of the rough ER membrane (see Figure 13-17). The final dolichol pyrophosphoryl oligosaccharide is oriented so that the oligosaccharide portion faces the ER lumen. The three glucose residues, which are the last residues added during synthesis of the precursor on the dolichol carrier, appear to act as a signal that the oligosaccharide is complete and ready to be transferred to a protein.

The entire 14-residue precursor is transferred from the dolichol carrier to an asparagine residue on a nascent polypeptide as it emerges into the ER lumen (Figure 13-18, step **1**). Only asparagine residues in the tripeptide sequences Asn-X-Ser and Asn-X-Thr (where X is any amino acid except proline) are substrates for *oligosaccharyl transferase*, the enzyme that catalyzes this reaction. Two of the three subunits of this enzyme are ER membrane proteins whose cytosol-facing domains bind to the ribosome, localizing a third subunit of the transferase, the catalytic subunit, near the growing polypeptide chain in the ER lumen. Not all Asn-X-Ser/Thr sequences become glycosylated, and it is not possible to predict from the amino acid sequence alone which potential *N*-linked glycosylation sites will be modified; for instance, rapid folding of a segment of a protein containing an Asn-X-Ser/Thr sequence may prevent transfer of the oligosaccharide precursor to it. Immediately after the entire precursor, $Glc_3Man_9(GlcNAc)_2$, is transferred to a nascent polypeptide, three different enzymes, called *glycosidases*,

remove all three glucose residues and one particular mannose residue (see Figure 13-18, steps 2 – 4).

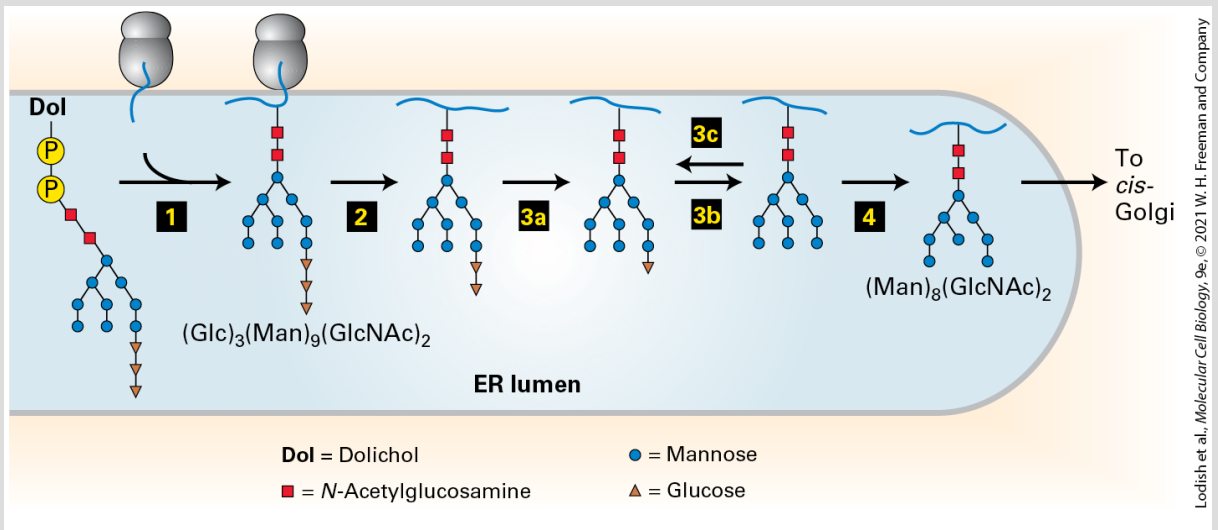


FIGURE 13-18 Addition and initial processing of *N*-linked oligosaccharides. In the rough ER of vertebrate cells, the $Glc_3Man_9(GlcNAc)_2$ precursor is transferred from the dolichol carrier to a susceptible asparagine residue on a nascent protein as soon as the asparagine crosses to the luminal side of the ER (step 1). In three separate reactions, first one glucose residue (step 2), then two glucose residues (steps 3a and 3b), and finally one mannose residue (step 4) are removed. Re-addition of one glucose residue (step 3c) plays a role in the correct folding of many proteins in the ER, as discussed later. The process of *N*-linked glycosylation of a soluble secretory protein is shown here, but the luminal portions of an integral membrane protein can be modified on asparagine residues by the same mechanism. See R. Kornfeld and S. Kornfeld, 1985, *Annu. Rev. Biochem.* **45**:631; and M. Sousa and A. J. Parodi, 1995, *EMBO J.* **14**:4196.

Oligosaccharide Side Chains May Promote Folding and Stability of Glycoproteins

The oligosaccharides attached to glycoproteins serve various functions. For example, some proteins require *N*-linked oligosaccharides in order to fold properly in the ER. This function has been demonstrated in studies with the antibiotic tunicamycin, which blocks the first step in the formation of the dolichol-linked oligosaccharide precursor and therefore inhibits synthesis of all *N*-linked oligosaccharides in cells (see Figure 13-17, *top left*). For example, in the presence of tunicamycin, the influenza virus hemagglutinin precursor polypeptide (HA_0) is synthesized, but it cannot fold properly and form a normal trimer; in this case, the misfolded protein remains in the rough ER. Moreover, mutation of a particular asparagine in the hemagglutinin sequence to a glutamine residue prevents addition of an *N*-linked oligosaccharide to that site and causes the protein to accumulate in the ER in an unfolded state.

In addition to promoting proper folding, *N*-linked oligosaccharides confer stability on many secreted glycoproteins. Many secretory proteins fold properly and are transported to their final destination even if the addition of all *N*-linked oligosaccharides is blocked, for example, by treatment with tunicamycin. However, such nonglycosylated proteins have been shown to be less stable than their glycosylated forms. For instance, glycosylated fibronectin, a normal component of the extracellular matrix, is degraded much more slowly by tissue proteases than is nonglycosylated fibronectin.

Oligosaccharides on certain cell-surface glycoproteins also play a role in cell-cell adhesion. For example, the plasma membrane of white blood cells (leukocytes) contains cell-adhesion molecules (CAMs) that are extensively glycosylated. The oligosaccharides in these molecules interact with a sugar-binding domain in certain other CAMs found on endothelial cells lining blood vessels. This interaction tethers the leukocytes to the endothelium and assists in their movement into tissues during an inflammatory response to infection (see Figure 20-40). Other cell-surface glycoproteins possess oligosaccharide side chains that can induce an immune response. A common example is the ABO blood group antigens, which are *O*-linked oligosaccharides attached to glycoproteins and glycolipids on the surfaces of erythrocytes and other cell types (see Figure 10-20). In both cases, oligosaccharides are added to the luminal face of these membrane proteins, in a manner similar to what is shown in Figure 13-18 for soluble proteins. The luminal face of these membrane proteins is topologically equivalent to the exterior face of the plasma membrane, where these proteins eventually end up.

Disulfide Bonds Are Formed and Rearranged by Proteins in the ER Lumen

In Chapter 3, we learned that both intramolecular and intermolecular **disulfide bonds** (**–S–S–**) help stabilize the tertiary

and quaternary structure of many proteins. These covalent bonds form by the oxidative linkage of *sulfhydryl groups (–SH)*, also known as *thiol* groups, on two cysteine residues in the same or different polypeptide chains. This reaction can proceed only when a suitable oxidant is present. In eukaryotic cells, disulfide bonds are formed in the lumen of the rough ER and are typically found in soluble secretory proteins and in the exoplasmic domains of membrane proteins. The cytosol does not contain enzymes for disulfide bond formation and therefore cytosolic proteins and most organelle proteins (i.e., those destined for mitochondria, chloroplasts, peroxisomes, etc.) usually lack disulfide bonds.

The efficient formation of disulfide bonds in the lumen of the ER depends on the enzyme **protein disulfide isomerase (PDI)**, which is present in all eukaryotic cells. This enzyme is especially abundant in the ER of secretory cells in organs such as the liver and pancreas, where large quantities of proteins that contain disulfide bonds are produced. As shown in Figure 13-19a, the disulfide bond in the active site of PDI can be readily transferred to a protein by two sequential thiol-disulfide transfer reactions. The reduced PDI generated by this reaction is returned to an oxidized form by the action of an ER-resident protein, called *Ero1*, which carries a disulfide bond that can be transferred to PDI. Ero1 itself becomes oxidized by reaction with molecular oxygen that has diffused into the ER.

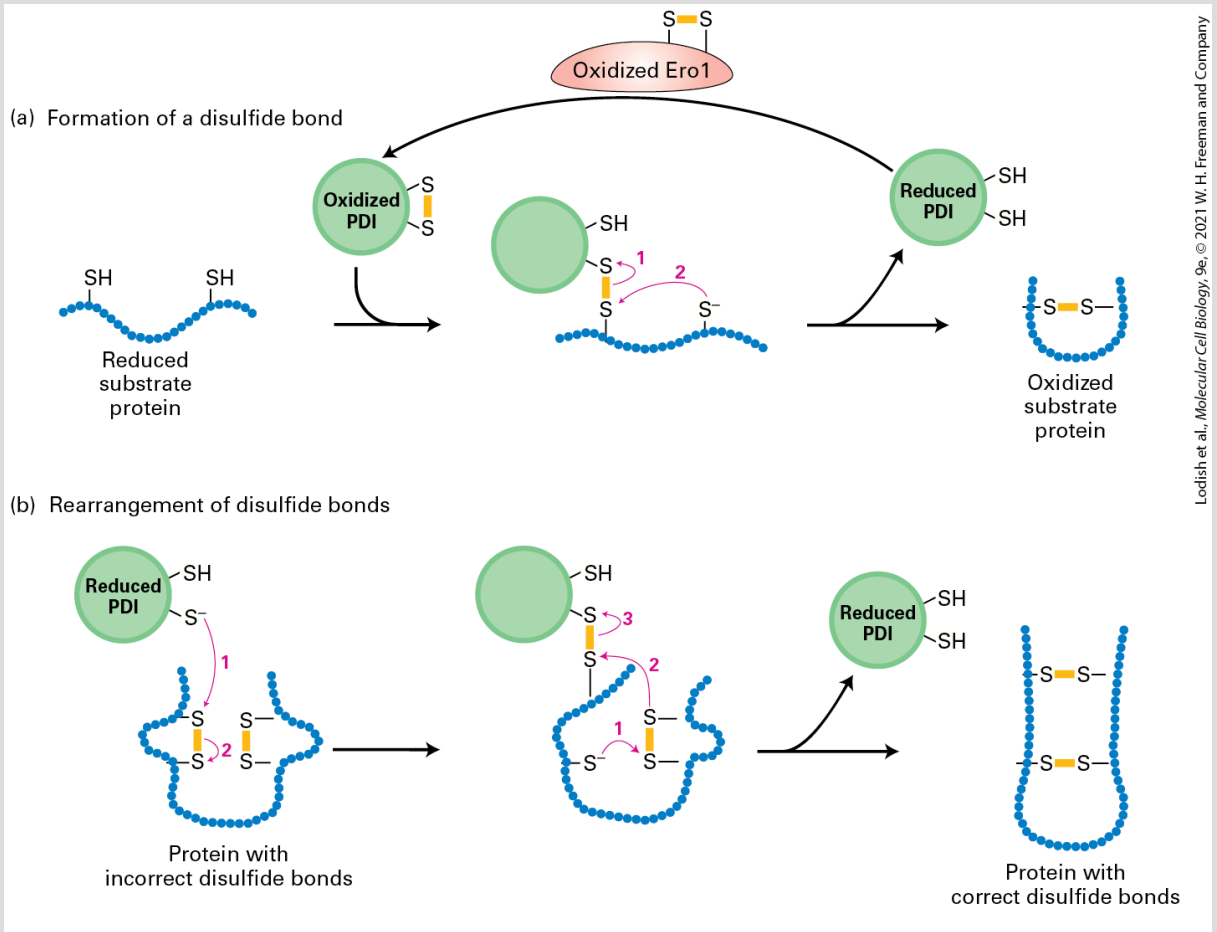


FIGURE 13-19 Action of protein disulfide isomerase (PDI). PDI forms and rearranges disulfide bonds via an active site with two closely spaced cysteine residues that are easily interconverted between the reduced dithiol form and the oxidized disulfide form. Numbered red arrows indicate the sequence of electron transfers. Yellow bars represent disulfide bonds. (a) In the formation of disulfide bonds, the ionized ($-S^-$) form of a cysteine thiol in the substrate protein reacts with the disulfide (S–S) bond in oxidized PDI to form a disulfide-bonded PDI–substrate protein intermediate. A second ionized thiol in the substrate protein then reacts with this intermediate, forming a disulfide bond within the substrate protein and releasing reduced PDI. PDI, in turn, transfers electrons to a disulfide bond in the luminal protein Ero1, thereby regenerating the oxidized form of PDI. (b) Reduced PDI can catalyze rearrangement of improperly formed disulfide bonds by similar thiol-disulfide transfer reactions. In this case, reduced PDI both initiates and is regenerated in the reaction pathway. These reactions are repeated until the most stable conformation of the protein is achieved. See M. M. Lyles and H. F. Gilbert, 1991, *Biochemistry* **30**:619.

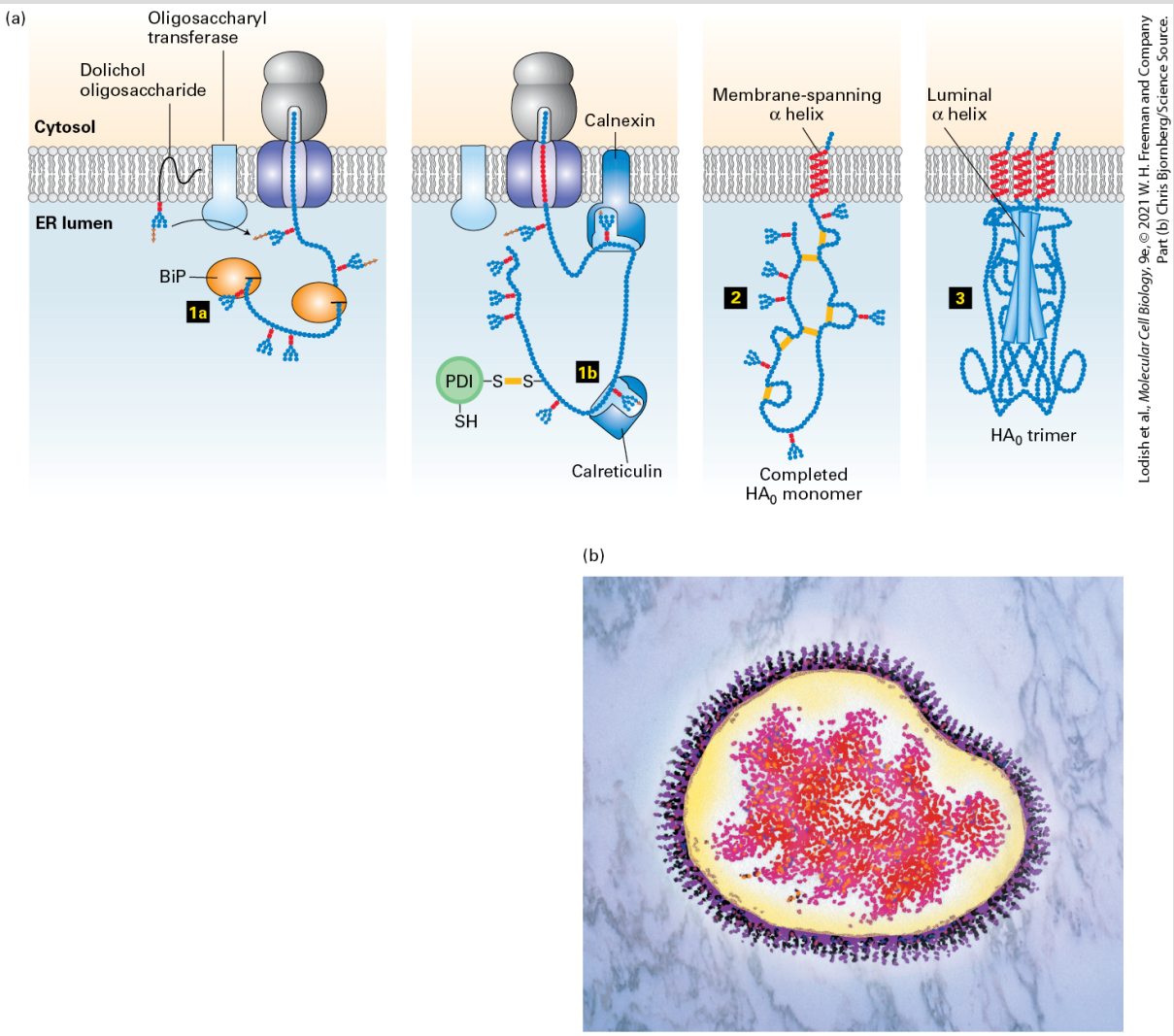
In proteins that contain more than one disulfide bond, the proper pairing of cysteine residues is essential for normal structure and activity. Disulfide bonds are commonly formed between cysteines that occur sequentially in the amino acid sequence while a polypeptide is still growing on the ribosome. Such sequential formation, however, sometimes yields disulfide bonds between the wrong cysteines. For example, proinsulin, a precursor to the peptide hormone insulin, has three disulfide bonds that link cysteines 1 and 4, 2 and 6, and 3 and 5. In this case, a disulfide bond that initially formed sequentially (e.g., between cysteines 1 and 2) would have to be rearranged for the protein to achieve its proper folded conformation. In cells, the rearrangement of disulfide bonds is also accelerated by PDI, which acts on a broad range of protein substrates, allowing them to reach their most thermodynamically stable conformations (Figure 13-19b). Disulfide bonds generally form in a specific order, first stabilizing small domains of a polypeptide, then stabilizing the interactions of more distant segments; this phenomenon is illustrated by the folding of the influenza hemagglutinin (HA) protein, discussed in the next section.

Chaperones and Other ER Proteins Facilitate Folding and Assembly of Proteins

Although many denatured proteins can spontaneously refold into their native state in vitro, such refolding usually requires hours to reach completion. Yet proteins produced in the ER generally fold into their proper conformation within minutes after their synthesis. The rapid folding of these newly synthesized proteins in cells depends on the sequential action of several proteins present within the ER lumen. We have already seen how the molecular chaperone BiP can drive post-translational translocation in yeast by binding fully synthesized polypeptides as they enter the ER (see Figure 13-9). BiP can also bind transiently to nascent polypeptide chains as they enter the ER during cotranslational translocation. Bound BiP is thought to prevent segments of a nascent chain from misfolding or forming aggregates, thereby promoting folding of the entire polypeptide into the proper conformation. PDI also contributes to proper folding because the correct three-dimensional conformation is stabilized by disulfide bonds in many proteins.

As illustrated in Figure 13-20, two other ER proteins, the homologous lectins (carbohydrate-binding proteins) *calnexin* and *calreticulin*, bind selectively to certain *N*-linked oligosaccharides on growing polypeptide chains. The ligand for these two proteins, which resembles the *N*-linked oligosaccharide precursor but has only a single glucose residue $[Glc_1Man_9(GlcNAc)_2]$, is generated by a specific glucosyltransferase in the ER lumen (see Figure 13-18, step **3a**). This enzyme acts only on polypeptide chains that are unfolded or misfolded, and in this respect, the glucosyltransferase acts as one

of the primary surveillance mechanisms to ensure quality control of protein folding in the ER. Unfolded proteins often expose hydrophobic segments that in a properly folded state are buried in the hydrophobic core of the protein. The glucosyltransferase is thought to specifically recognize unfolded proteins by binding to these exposed hydrophobic segments. Binding of calnexin and calreticulin to unfolded nascent chains marked with glucosylated *N*-linked oligosaccharides prevents aggregation of adjacent segments of a protein as it is being made on the ER. Thus calnexin and calreticulin, like BiP, help prevent premature, incorrect folding of segments of a newly made protein.

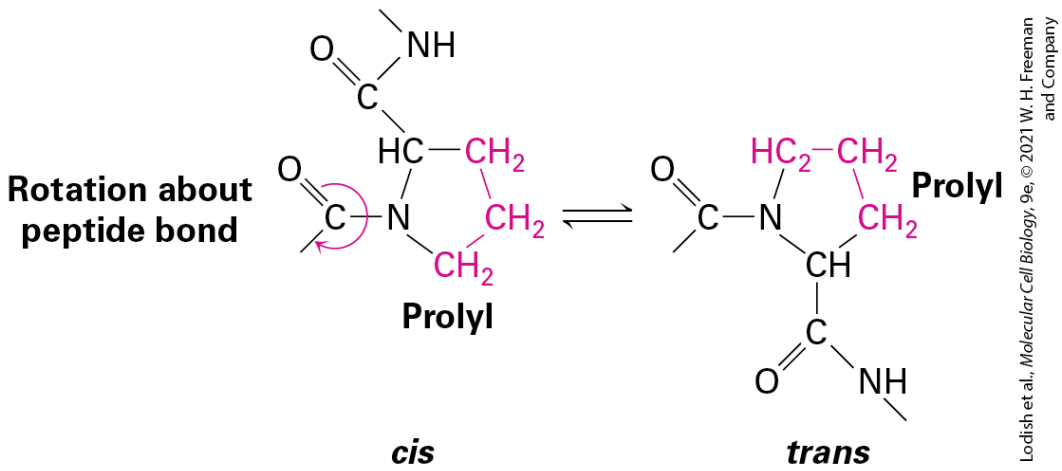


Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Part (b) Chris Bjornberg/Science Source.

FIGURE 13-20 Hemagglutinin folding and assembly. (a) Mechanism of (HA_0) trimer assembly. Transient binding of the chaperone BiP (step **1a**) to the nascent polypeptide chain and of two lectins, calnexin and calreticulin, to certain oligosaccharide chains (step **1b**) promotes proper folding of adjacent segments of HA_0. A total of seven *N*-linked oligosaccharide chains are added to the luminal portion of the nascent chain during cotranslational translocation, and PDI catalyzes the formation of six disulfide bonds per monomer. Completed HA_0 monomers are anchored in the membrane by a single membrane-spanning α helix with the N-terminus in the lumen (step **2**). Interaction of three HA_0 chains with one another, initially via their transmembrane α helices, apparently triggers formation of a long stem containing one α helix from the luminal part of each HA_0 polypeptide. Finally, interactions occur among the three globular heads, generating a stable HA_0 trimer (step **3**). (b) Electron micrograph (false color) of a complete influenza virion

showing trimers of HA protein protruding as spikes from the surface of the viral membrane. See U. Tatu et al., 1995, *EMBO J.* **14**:1340; and D. Hebert et al., 1997, *J. Cell Biol.* **139**:613.

Other important protein-folding catalysts in the ER lumen are *peptidyl-prolyl isomerases*, a family of enzymes that accelerate the rotation about peptidyl-prolyl bonds at proline residues in unfolded segments of a polypeptide:



Such isomerizations are sometimes the rate-limiting step in the folding of protein domains. Many peptidyl-prolyl isomerases can catalyze the rotation of exposed peptidyl-prolyl bonds indiscriminately in numerous proteins, but some have very specific protein substrates.

Many important soluble secretory and membrane proteins synthesized on the ER are built of two or more polypeptide subunits. In all cases, the assembly of the subunits constituting these multisubunit (multimeric) proteins occurs in the ER. The

immunoglobulins, which contain two heavy (H) and two light (L) chains, all linked by intrachain disulfide bonds, are assembled in this way. Hemagglutinin (HA) is another multimeric protein that provides a good illustration of folding and subunit assembly (see Figure 13-20). This trimeric protein forms the spikes that protrude from the surface of an influenza virus particle. The HA trimer is formed within the ER of an infected host cell from three copies of a precursor protein termed HA_0, which has a single membrane-spanning α helix. In the Golgi complex, each of the three HA_0 proteins is cleaved to form two polypeptides, HA_1 and HA_2; thus each HA molecule that eventually resides on the viral surface contains three copies of HA_1 and three of HA_2 (see Figure 3-10). The trimer is stabilized by interactions between the large exoplasmic domains of the constituent polypeptides, which extend into the ER lumen; after HA is transported to the cell surface, these domains extend into the extracellular space. Interactions between the smaller cytosolic and membrane-spanning portions of the HA subunits also help stabilize the trimeric protein. Studies have shown that it takes just 10 minutes for the HA_0 polypeptides to fold and assemble into their proper trimeric conformation.

Improperly Folded Proteins in the ER Induce Expression of Protein-Folding Catalysts

Wild-type proteins that are synthesized on the rough ER cannot exit this compartment until they achieve their completely folded conformation. Likewise, almost any mutation that prevents proper folding of a protein in the ER also blocks the movement of that protein from the ER lumen or membrane to the Golgi complex. The mechanisms that retain unfolded or incompletely folded proteins within the ER probably increase the overall efficiency of folding by keeping intermediate forms in proximity to folding catalysts, which are most abundant in the ER. Improperly folded proteins retained within the ER are generally seen bound to the ER chaperones BiP and calnexin. Thus these luminal folding catalysts perform two related functions: assisting in the folding of normal proteins by preventing their aggregation and binding to misfolded proteins to retain them in the ER.

Both mammalian cells and yeasts respond to the presence of unfolded proteins in the rough ER by increasing transcription of several genes that encode ER chaperones and other folding catalysts. A key participant in this **unfolded-protein response** is Ire1, an ER membrane protein that exists both as a monomer and as a dimer. The dimeric form, but not the monomeric form, promotes formation of Hac1, a transcription factor in yeast that activates expression of the genes induced in the unfolded-protein response. As depicted in Figure 13-21, binding of BiP to the luminal domain of monomeric Ire1 prevents formation of the Ire1 dimer. Thus the quantity of free BiP in the ER lumen determines the relative proportions of monomeric and dimeric Ire1. Accumulation of

unfolded proteins within the ER lumen sequesters BiP molecules, making them unavailable for binding to Ire1. As a result, the level of dimeric Ire1 increases, leading to an increase in the level of Hac1 and production of proteins that assist in protein folding.

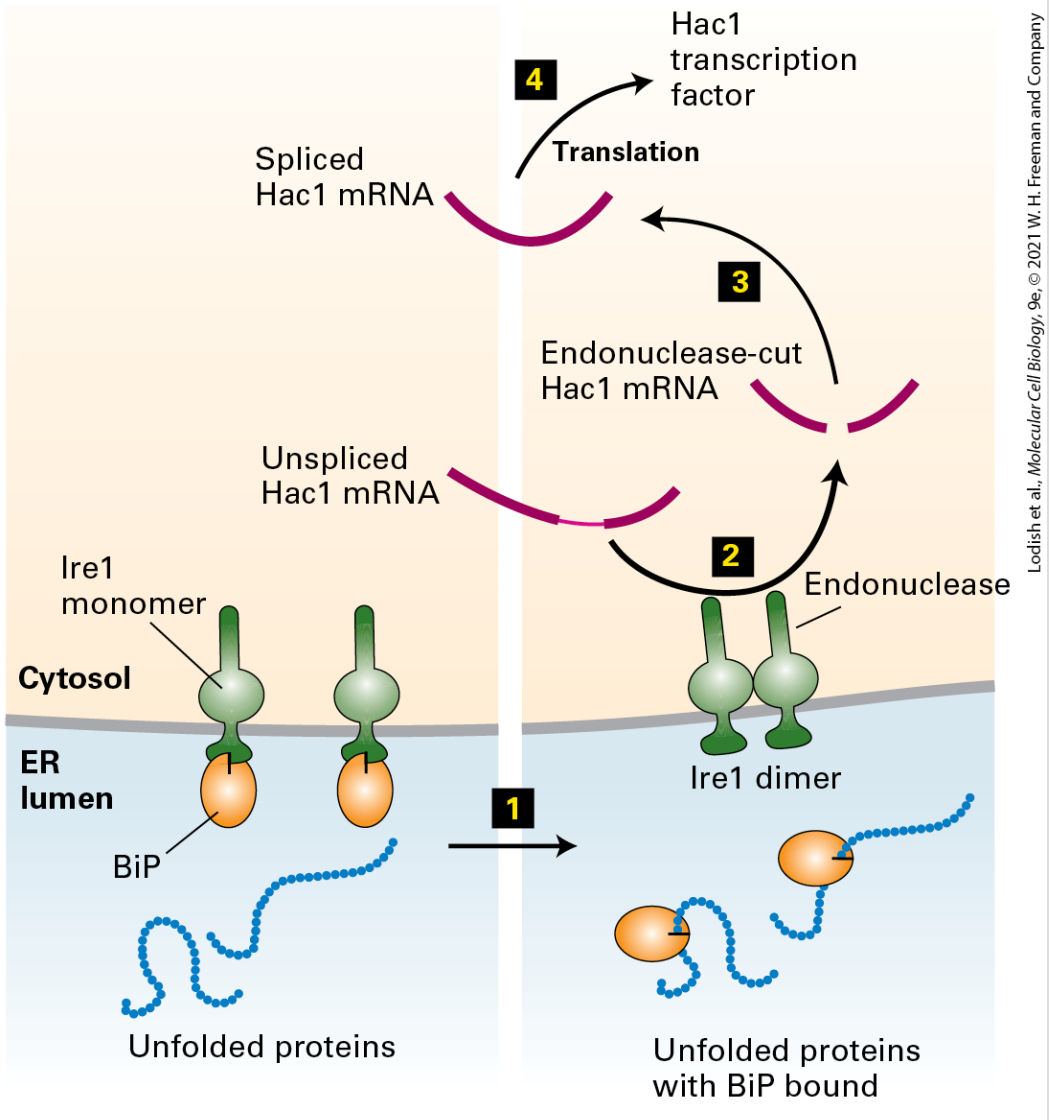


FIGURE 13-21 The unfolded-protein response. Ire1, a transmembrane protein in the ER membrane, has a binding site for BiP on its luminal domain; the cytosolic domain contains a specific RNA endonuclease. Step 1: Accumulating unfolded proteins in the ER lumen bind BiP molecules, releasing them from monomeric Ire1. Dimerization of Ire1 then activates its

endonuclease activity. Steps 2 – 3 : The unspliced mRNA precursor encoding the transcription factor Hac1 is cleaved by dimeric Ire1, and the two exons are joined to form functional Hac1 mRNA. Current evidence indicates that this processing occurs in the cytosol, although pre-mRNA processing generally occurs in the nucleus. Step 4 : Hac1 is translated into Hac1 protein, which then moves back into the nucleus and activates transcription of genes encoding several protein-folding catalysts. See U. Ruegsegger et al., 2001, *Cell* **107**:103; A. Bertolotti et al., 2000, *Nat. Cell Biol.* **2**:326; and C. Sidrauski and P. Walter, 1997, *Cell* **90**:1031.

Mammalian cells contain an additional regulatory pathway that operates in response to unfolded proteins in the ER. In this pathway, accumulation of unfolded proteins in the ER triggers proteolysis of ATF6, a transmembrane protein in the ER membrane, at a site within the membrane-spanning segment. The cytosolic domain of ATF6 released by proteolysis then moves to the nucleus, where it stimulates transcription of the genes encoding ER chaperones. Activation of a transcription factor by such *regulated intramembrane proteolysis* also occurs in the Notch signaling pathway and during activation of the cholesterol-responsive transcription factor SREBP (see Figures 16-25 and 21-6).

A hereditary form of emphysema illustrates the detrimental effects that can result from misfolding of proteins in the ER. This disease is caused by a point mutation in α_1-antitrypsin, which is normally secreted by hepatocytes and macrophages. The wild-type protein binds to and inhibits trypsin as well as the blood protease elastase. In the absence of normal α_1-antitrypsin, elastase degrades

the fine tissue in the lung that participates in the absorption of oxygen, eventually producing the symptoms of emphysema. Although the mutant α_1-antitrypsin is synthesized on the rough ER, it does not fold properly, forming an almost crystalline aggregate that is not exported from the ER. In hepatocytes, the secretion of other proteins also becomes impaired as the rough ER is filled with aggregated α_1-antitrypsin. ■

Unassembled or Misfolded Proteins in the ER Are Often Transported to the Cytosol for Degradation

Misfolded soluble secretory and membrane proteins, as well as the unassembled subunits of multimeric proteins, are often degraded within an hour or two after their synthesis in the rough ER. Initially, it was thought that proteolytic enzymes within the ER lumen catalyzed degradation of misfolded or unassembled polypeptides, but such proteases were never found. More recent studies have shown that misfolded secretory proteins are recognized by specific ER membrane proteins and are targeted for transport from the ER lumen into the cytosol by a process known as **dislocation**.

The dislocation and degradation of misfolded proteins depends on a set of proteins located in the ER membrane and in the cytosol that perform three basic functions. The first function is recognition of

misfolded proteins. One mechanism for recognition involves the trimming of *N*-linked carbohydrate chains by mannosidases located in the ER (Figure 13-22). Trimmed glycans with the structure $\mathrm{Man}_{5\text{–}6}(\mathrm{GlcNAc})_2$ are recognized by a protein known as *OS-9*, which targets the trimmed glycoprotein for dislocation. It is not known precisely how the ER α-mannosidases distinguish the partially folded states of defective proteins that can never completely fold properly and are thus legitimate substrates for the dislocation process, from normal proteins that have not yet acquired their fully folded conformation. One possibility is that the ER α-mannosidases act slowly, so that only those glycoproteins that remain misfolded in the ER lumen for a sufficiently long time are trimmed and therefore targeted for degradation. Luminal proteins that lack oligosaccharides altogether can also be targeted for degradation, indicating that other processes for the recognition of unfolded proteins must also exist that do not involve trimming of *N*-linked oligosaccharides.

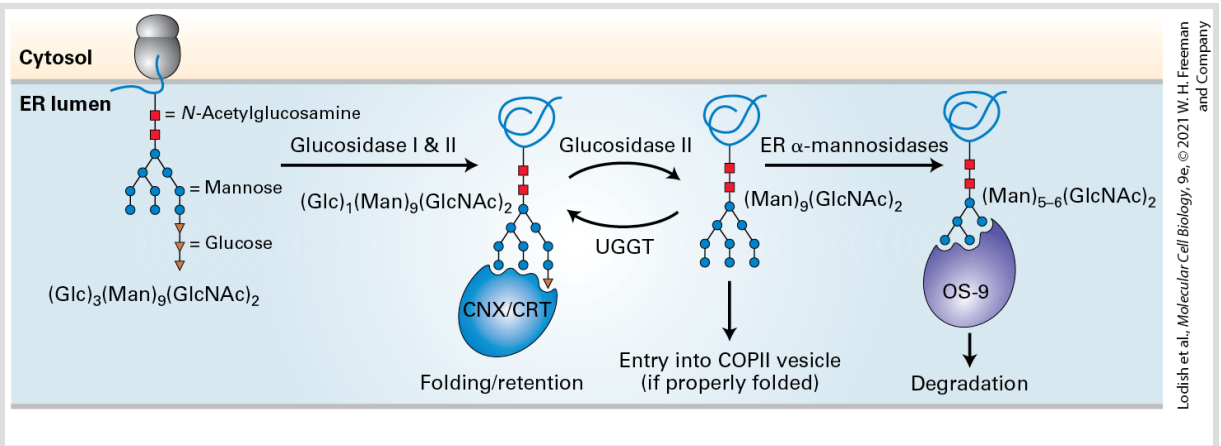


FIGURE 13-22 Modifications of *N*-linked oligosaccharides are used to monitor folding and for quality control. After a $\mathrm{Glc}_3\mathrm{Man}_9(\mathrm{GlcNAc})_2$ *N*-linked oligosaccharide is added to a

newly synthesized and incompletely folded protein in the ER, three glucose residues are removed by Glucosidases I and II to form $\mathrm{Glc_1Man_9(GlcNAc)_2}$ (see Figure 13-18, steps 2 and 3). This modified *N*-linked oligosaccharide binds the lectins calnexin (CNX) and calreticulin (CRT) for retention in the ER and engagement of folding chaperones. When the *N*-linked oligosaccharide is released from $\mathrm{CNX/and}$ CRT, removal of glucose by Glucosidase II yields $\mathrm{Man_9(GlcNAc)_2}$, which is further processed to $\mathrm{Man_8(GlcNAc)_2}$ before entering a COPII vesicle bound for the *cis*-Golgi (see Figure 13-18, steps 3 and 4). If the protein has not folded completely, a glucosyltransferase known as *UGGT* can add a glucose residue to reform $\mathrm{Glc_1Man_9(GlcNAc)_2}$, which will bind again to CNX and CRT for another attempt at folding. Proteins that cannot fold properly, and are therefore retained in the ER for longer times, eventually undergo mannose trimming by ER α-mannosidases to form $\mathrm{Man_{5\text{–}6}(GlcNAc)_2}$, which is recognized by OS-9. Recognition by OS-9 leads to dislocation of the misfolded protein out of the ER, ubiquitinylation, and degradation by proteasomes.

Misfolded ER proteins that are marked for degradation are transported across the ER membrane by a process known as *dislocation*, since it occurs in the reverse direction from translocation, which occurs during synthesis. A complex of at least four integral membrane proteins, known as the **ERAD (*ER-associated degradation*)** complex, enables dislocation of misfolded proteins across the ER membrane, but there is no evidence that the ERAD complex forms a protein channel for dislocation, suggesting other possible mechanisms. One possibility is that the Sec61 translocation channel can be repurposed for the dislocation of misfolded proteins from the ER lumen to the cytosol. Dislocation involves powerful pulling and unfolding protein machines in the cytosol and it is possible that misfolded proteins are dragged directly through the lipid bilayer without a channel dedicated for this purpose.

As segments of a dislocated polypeptide emerge on the cytosolic face of the membrane, they encounter p97, a member of a protein family known as the **AAA ATPase family** that couples the energy of ATP hydrolysis to disassembly of protein complexes. In dislocation of polypeptides from the ER, hydrolysis of ATP by p97 may provide the driving force to pull misfolded proteins from the ER membrane into the cytosol. As the misfolded proteins enter the cytosol, specific ubiquitin ligase enzymes that are components of the ERAD complex add ubiquitin residues to these dislocated peptides. Like the action of p97, ubiquitinylation is coupled to ATP hydrolysis and this reaction may also contribute to the energetically favorable trapping of the proteins in the cytosol. The resulting polyubiquitinylated polypeptides, now fully in the cytosol, are targeted for degradation in proteasomes. The role of polyubiquitinylation in targeting proteins to proteasomes is discussed more fully in Chapter 3 (see Figure 3-32 and Figure 3-39).

KEY CONCEPTS OF SECTION 13.3

Protein Modifications, Folding, and Quality Control in the ER

- All *N*-linked oligosaccharides, which are bound to asparagine residues, contain a core of two *N*-acetylglucosamine and at least three mannose residues and usually have several branches. *O*-linked oligosaccharides, which are bound to serine or threonine residues, are generally short, often containing only one to four sugar residues.
- Formation of *N*-linked oligosaccharides begins with assembly of a conserved 14-residue high-mannose oligosaccharide precursor on dolichol, a lipid in the membrane of the rough ER (see Figure 13-17). After this preformed oligosaccharide is transferred

to specific asparagine residues of nascent polypeptide chains in the ER lumen, three glucose residues and one mannose residue are removed (see Figure 13-18).

- Oligosaccharide side chains may assist in the proper folding of glycoproteins, help protect the mature proteins from proteolysis, participate in cell-cell adhesion, and function as antigens.
- Disulfide bonds are added to many soluble secretory proteins and to the exoplasmic domain of membrane proteins in the ER. Protein disulfide isomerase (PDI), present in the ER lumen, catalyzes both the formation and the rearrangement of disulfide bonds (see Figure 13-19).
- The chaperone BiP, the lectins calnexin and calreticulin, and peptidyl-prolyl isomerases work together to ensure proper folding of newly made secretory and membrane proteins in the ER. The subunits of multimeric proteins also assemble in the ER (see Figure 13-20).
- Only properly folded proteins and assembled subunits are transported from the rough ER to the Golgi complex in vesicles.
- The accumulation of abnormally folded proteins and unassembled subunits in the ER can induce increased expression of ER protein-folding catalysts via the unfolded-protein response (see Figure 13-21).
- Unassembled or misfolded proteins in the ER are often transported back to the cytosol, where they are degraded in the ubiquitin-proteasome pathway (see Figure 13-22).

13.4 Targeting of Proteins to Mitochondria and Chloroplasts

In the remainder of this chapter, we examine how proteins synthesized on cytosolic ribosomes are sorted to discrete organelles: mitochondria, chloroplasts, peroxisomes, and the nucleus (see Figure 13-1). Mitochondria and chloroplasts are closely related organelles that contain an internal lumen — called the *matrix* in mitochondria and the *stroma* in chloroplasts — that is surrounded by a double membrane. Within the stroma, chloroplasts contain a third membrane known as the *thylakoid*, which is where the light-harvesting reactions of photosynthesis take place. In contrast to mitochondria and chloroplasts, peroxisomes are bounded by a single membrane and have a single luminal matrix compartment. The mechanism of protein transport into and out of the nucleus, which differs in many respects from sorting to other organelles, is discussed in the last section of the chapter.

In addition to being bounded by two membranes, mitochondria and chloroplasts share similar types of electron-transporting proteins and use an F-class ATPase to synthesize ATP (see Figure 12-26). Remarkably, these characteristics are shared by gram-negative bacteria. Also like bacterial cells, mitochondria and chloroplasts contain their own DNA, which encodes organelle rRNAs, tRNAs, and some proteins (see Chapter 8). Moreover, growth and division of mitochondria and chloroplasts are not coupled to nuclear division.

Rather, these organelles grow by the incorporation of cellular proteins and lipids, and new organelles form by division of preexisting organelles. These numerous similarities to free-living bacterial cells have led to the understanding that mitochondria and chloroplasts arose when bacteria were incorporated into ancestral eukaryotic cells, forming endosymbiotic organelles (see Figure 12-7). The sequence similarity of the many membrane translocation proteins shared by mitochondria, chloroplasts, and bacteria provides the most striking evidence for this ancient evolutionary relationship. In this section, we examine these membrane translocation proteins in detail.

Proteins encoded by mitochondrial DNA or chloroplast DNA are synthesized on ribosomes within the organelles and directed to the correct subcompartment within the parent organelle immediately after synthesis. However, the majority of proteins located in mitochondria and chloroplasts are not encoded by organelle DNA — they are encoded by genes in the nucleus and are imported into the organelles after their synthesis in the cytosol. Apparently, as eukaryotic cells evolved over a billion years, much of the genetic information from the ancestral bacterial DNA in these endosymbiotic organelles moved, by an unknown mechanism, to the nucleus. Precursor proteins synthesized in the cytosol that are destined for the matrix of mitochondria or the equivalent stroma in chloroplasts usually contain specific N-terminal targeting sequences that specify binding to receptor proteins on the organelle surface. Generally these sequences are cleaved once the protein reaches the

matrix or stroma. Clearly these targeting sequences are similar in their location and general function to the signal sequences that direct nascent proteins to the ER lumen. Although the three types of signals share some sequence features, their specific sequences differ considerably, as summarized in Table 13-1.

In both mitochondria and chloroplasts, protein import requires energy and occurs at points where the outer and inner organelle membranes are in direct contact. Because mitochondria and chloroplasts contain multiple membranes and membrane-limited spaces, the sorting of many proteins to their correct locations requires the sequential action of two targeting sequences and two membrane-bound translocation systems: one to direct the protein into the organelle and the other to direct it into the correct organelle subcompartment or membrane. As we will see, the mechanisms for sorting various proteins to mitochondria and chloroplasts are related to SRP-dependent recognition of a signal sequence as discussed previously.

Amphipathic N-Terminal Targeting Sequences Direct Proteins to the Mitochondrial Matrix

All proteins that travel from the cytosol to the same mitochondrial destination have targeting signals that share common motifs,

although the signal sequences are not identical. Thus the receptors that recognize such signals are able to bind to a number of different but related sequences. The most extensively studied sequences for localizing proteins to mitochondria are the *matrix-targeting sequences*. These sequences, located at the N-terminus, are usually 20–50 amino acids in length. They are rich in hydrophobic amino acids, positively charged basic amino acids (arginine and lysine), and hydroxylated ones (serine and threonine) but tend to lack negatively charged acidic residues (aspartate and glutamate).

Mitochondrial matrix–targeting sequences are thought to assume an α-helical conformation in which positively charged amino acids predominate on one side of the helix and hydrophobic amino acids on the other. Sequences such as these that contain both hydrophobic and hydrophilic regions are said to be amphipathic. Mutations that disrupt the amphipathic character of these sequences usually disrupt targeting to the matrix, although many other amino acid substitutions do not. These findings indicate that the amphipathicity of matrix-targeting sequences is critical to their function.

The cell-free assay outlined in Figure 13-23 has been widely used to define the biochemical steps in the import of mitochondrial precursor proteins. Respiring (energized) mitochondria extracted from cells can incorporate mitochondrial precursor proteins that have been synthesized in the absence of mitochondria if they are carrying appropriate targeting sequences. Successful incorporation

of the precursor into the organelle can be assayed by resistance to digestion by an added protease such as trypsin. In other assays, successful import of a precursor protein can be shown by the proper cleavage of the N-terminal targeting sequences by specific mitochondrial proteases. The uptake of completely presynthesized mitochondrial precursor proteins by the organelle in this system contrasts with the cell-free cotranslational translocation of secretory proteins into the ER, which generally occurs only when microsomal (ER-derived) membranes are present during synthesis (see Figure 13-4).

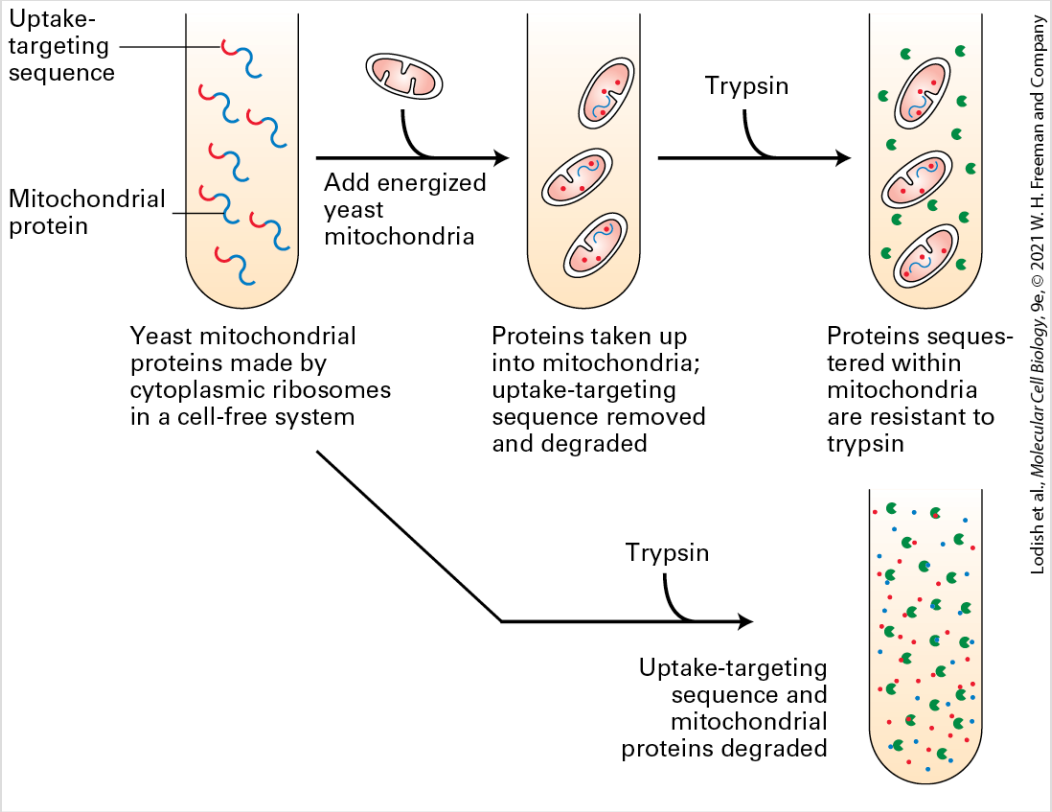


FIGURE 13-23 Import of mitochondrial precursor proteins is assayed in a cell-free system. Mitochondrial precursor proteins with attached uptake-targeting signals can be synthesized on ribosomes in a cell-free reaction. When respiring

mitochondria are added to these synthesized mitochondrial precursor proteins (*top*), the proteins are taken up by mitochondria. Inside mitochondria, the proteins are protected from the action of proteases such as trypsin. When no mitochondria are present (*bottom*), the mitochondrial proteins are degraded by added protease. Only energized (respiring) mitochondria, which have a proton electrochemical gradient (proton-motive force) across the inner membrane, take up these proteins. Furthermore, the imported proteins must contain an appropriate uptake-targeting sequence. Uptake also requires ATP and a cytosolic extract containing chaperone proteins that maintain the precursor proteins in an unfolded conformation. This assay has been used to study targeting sequences and other features of the translocation process.

Mitochondrial Protein Import Requires Outer-Membrane Receptors and Translocons in Both Membranes

Figure 13-24 presents an overview of protein import from the cytosol into the mitochondrial matrix, the route into the mitochondrion followed by most imported proteins. Here we discuss in detail each step of protein transport into the matrix, then consider how some proteins are subsequently targeted to other compartments of the mitochondrion.

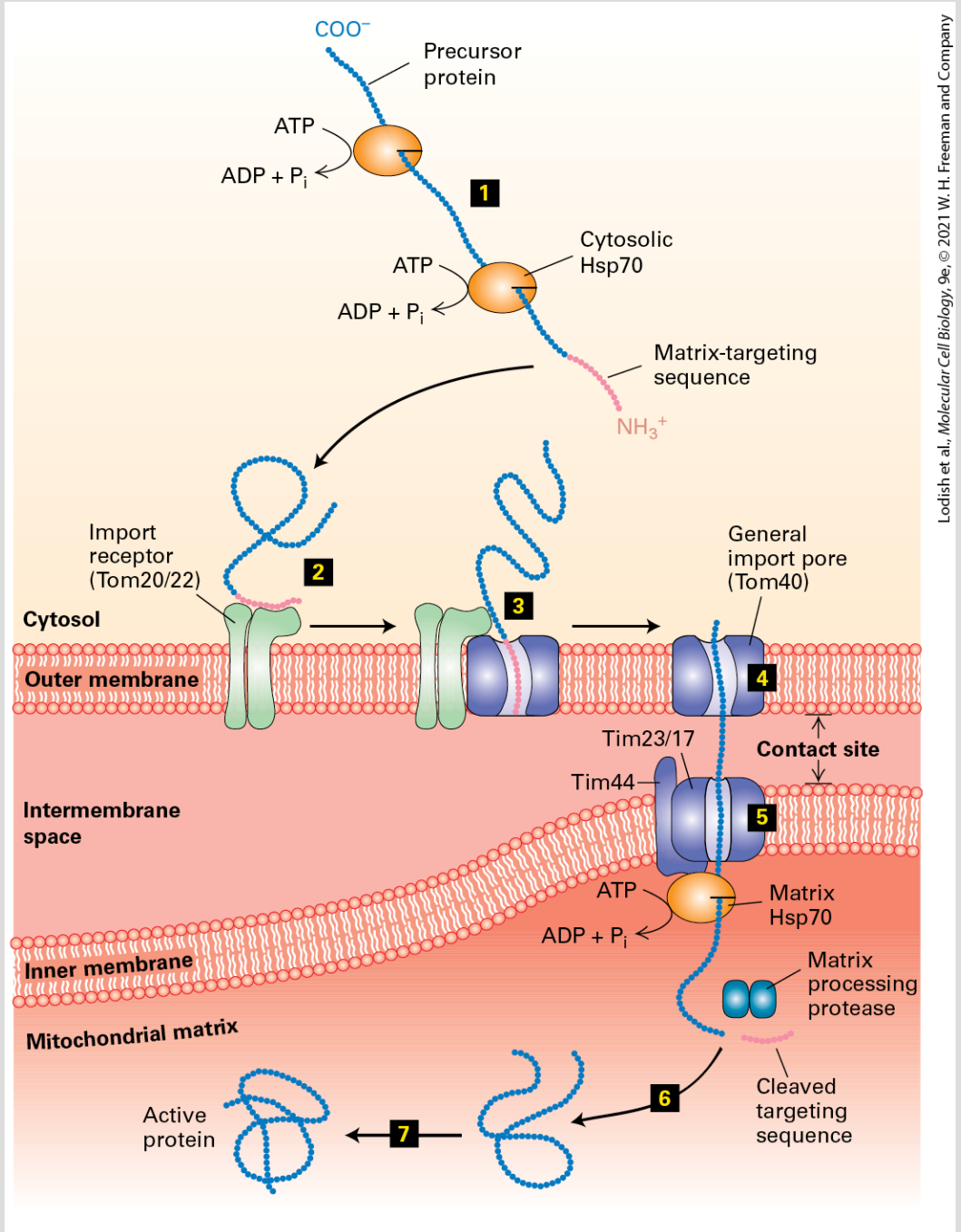


FIGURE 13-24 Protein import into the mitochondrial matrix. Precursor proteins synthesized on cytosolic ribosomes are maintained in an unfolded or partially folded state by bound chaperones, such as cytosolic Hsp70 (step 1). After a precursor protein binds to an import receptor near a site of contact with the inner membrane (step 2), it is transferred into the general import pore (step 3). The translocating protein then moves through this channel and an adjacent channel in the inner membrane (steps 4–5). Note that

translocation occurs at rare “contact sites” at which the inner and outer membranes appear to touch. Binding of the translocating protein by matrix Hsp70 and subsequent ATP hydrolysis by Hsp70 helps drive import into the matrix. Once the targeting sequence is removed by a matrix protease and Hsp70 is released from the newly imported protein (step **6**), the protein folds into its mature, active conformation within the matrix (step **7**). Folding of some proteins depends on matrix chaperonins. See G. Schatz, 1996, *J. Biol. Chem.* **271**:31763; and N. Pfanner et al., 1997, *Annu. Rev. Cell Dev. Biol.* **13**:25.

After synthesis in the cytosol, the soluble precursors of mitochondrial proteins (including hydrophobic integral membrane proteins) can interact directly with the mitochondrial membrane. Import of an unfolded mitochondrial precursor protein is initiated by the binding of a mitochondrial targeting sequence to an *import receptor* in the outer mitochondrial membrane. These import receptors were first identified by experiments in which antibodies to specific proteins of the outer mitochondrial membrane were shown to inhibit protein import into isolated mitochondria. Subsequent genetic experiments in which the genes for specific mitochondrial outer-membrane proteins were mutated showed that specific receptor proteins were responsible for the import of different classes of mitochondrial proteins. For example, N-terminal matrix-targeting sequences are recognized by a complex of Tom20 and Tom22 (proteins in the outer mitochondrial membrane involved in targeting and import, designated *Tom* proteins for *t*ranslocon of the *o*uter *m*embrane).

Many proteins can be imported into the mitochondrion only in an unfolded state. Chaperone proteins such as cytosolic Hsp70 and Hsp90 use energy derived from ATP hydrolysis to keep nascent and newly made proteins in a disaggregated state so that they are available to be taken up by mitochondria. For some mitochondrial precursor proteins, the mitochondrial outer-membrane protein Tom70 serves as an import receptor by binding to both Hsp90 and portions of the unfolded precursor protein.

The import receptors subsequently transfer the precursor protein to an import channel in the outer membrane. This channel, composed mainly of the Tom40 protein, is known as the **general import pore** because all known mitochondrial precursor proteins gain access to the interior compartments of the mitochondrion through it. When Tom40 is purified and incorporated into liposomes, it forms a transmembrane channel with a pore wide enough to accommodate an unfolded polypeptide chain. The general import pore forms a largely passive channel through the outer mitochondrial membrane; the driving force for unidirectional transport comes from within the mitochondrion, as we will see shortly. In the case of precursors destined for the mitochondrial matrix, transfer through the outer membrane occurs simultaneously with transfer through an inner-membrane channel composed of the Tim23 and Tim17 proteins. (*Tim* stands for *t*ranslocon of the *i*nner *m*embrane.) Translocation into the matrix thus occurs at contact sites where the outer and inner membranes are in close proximity.

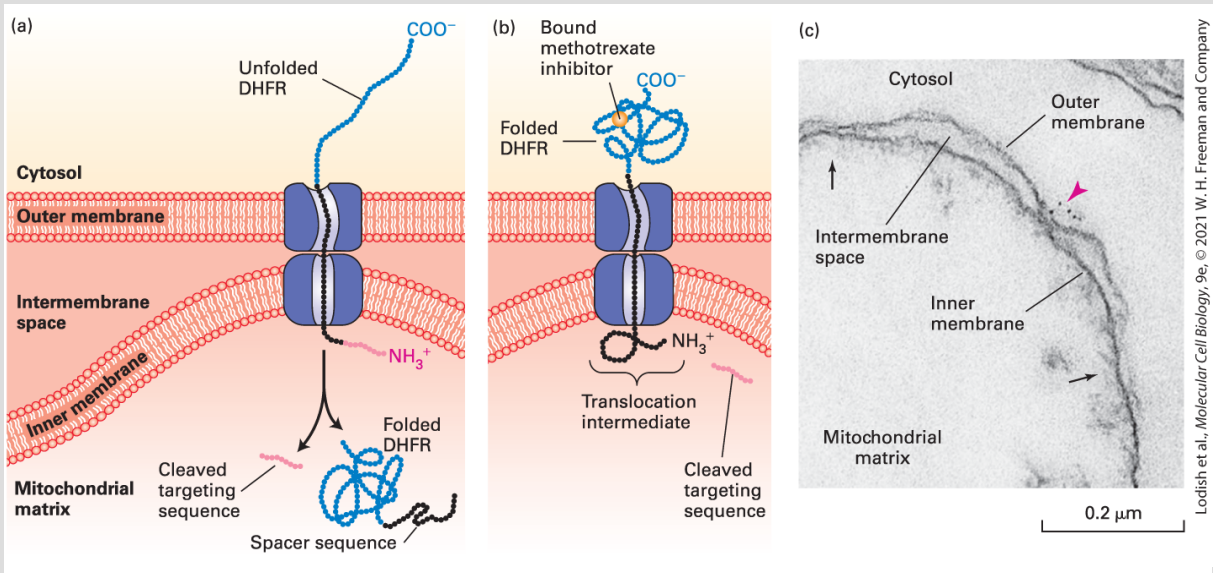
Soon after the N-terminal matrix-targeting sequence of a protein enters the mitochondrial matrix, it is removed by a protease that resides within the matrix. The emerging protein is also bound by matrix Hsp70, a chaperone that is localized to the translocation channels in the inner mitochondrial membrane by interaction with transmembrane protein Tim44. Together, Tim44 and Hsp70 are thought to power translocation of proteins into the matrix.

Some imported proteins can fold into their final, active conformation without further assistance. Final folding of many matrix proteins, however, requires chaperonins. As discussed in Chapter 3, chaperonin proteins actively facilitate protein folding by a process that depends on ATP. Yeast mutants defective in Hsc60, a chaperonin in the mitochondrial matrix, can import matrix proteins and cleave their targeting sequences normally, but the imported polypeptides fail to fold and assemble into their native tertiary and quaternary structures.

Studies with Chimeric Proteins Demonstrate Important Features of Mitochondrial Protein Import

Dramatic evidence for the ability of mitochondrial matrix–targeting sequences to direct the import of any protein to the matrix was shown by adding the matrix-targeting sequence of alcohol

dehydrogenase to the enzyme dihydrofolate reductase (DHFR), which normally resides in the cytosol. Cell-free translocation assays show that in the presence of chaperones, which prevent the DHFR segment from folding in the cytosol, the chimeric protein is transported into the matrix (Figure 13-25a). However, the DHFR inhibitor methotrexate, which binds tightly to the active site of DHFR and greatly stabilizes its folded conformation, renders the chimeric protein resistant to unfolding by cytosolic chaperones and prevents the chimeric protein from completely entering the translocon of the matrix. This finding demonstrates that a precursor must be unfolded in order to traverse the import pores in both mitochondrial membranes.



EXPERIMENTAL FIGURE 13-25 Experiments with chimeric proteins elucidate mitochondrial protein import processes. These experiments show that a matrix-targeting sequence alone directs proteins to the mitochondrial matrix and that only unfolded proteins are translocated across both mitochondrial membranes. The chimeric protein in these experiments contained a matrix-targeting signal at its N-terminus (red), followed by a spacer sequence of no particular function (black), and then by dihydrofolate reductase (DHFR), an

enzyme normally present only in the cytosol. (a) When the DHFR segment is unfolded, the chimeric protein moves across both membranes to the matrix of an energized mitochondrion, and the matrix-targeting signal is then removed. (b) When the C-terminus of the chimeric protein is locked in the folded state by binding of methotrexate, translocation is blocked. If the spacer sequence is long enough to extend across both transport channels, a stable translocation intermediate with the targeting sequence cleaved off is generated in the presence of methotrexate, as shown here. (c) The C-terminus of the translocation intermediate in (b) can be detected by incubating the mitochondria with antibodies that bind to the DHFR segment, followed by gold particles coated with bacterial protein A, which binds nonspecifically to antibody molecules (see Figure 4-33). An electron micrograph of a sectioned sample reveals gold particles (red arrowhead) bound to the translocation intermediate at a contact site between the inner and outer membranes. Other contact sites (black arrows) are also evident. See J. Rassow et al., 1990, *FEBS Lett.* **275**:190.

[Part (c) M. Schwaiger, V. Herzog, and W. Neupert, 1987, *J. Cell Biol*. **105**:235–246; https://doi.org/10.1083/jcb.105.1.235.]

Additional studies revealed that if a sufficiently long spacer sequence separates the N-terminal matrix-targeting sequence and the DHFR portion of the chimeric protein, then, in the presence of methotrexate, a translocation intermediate that spans both membranes can be trapped (Figure 13-25b). Microscopy studies of these stable trapped intermediates show that they accumulate at sites where the inner and outer mitochondrial membranes are close together (Figure 13-25c). Since roughly a thousand stuck chimeric proteins can be observed at these *contact sites* in a typical yeast mitochondrion, it is thought that mitochondria have approximately a thousand general import pores for the uptake of mitochondrial proteins.

Three Energy Inputs Are Needed to Import Proteins into Mitochondria

As noted previously and as indicated in Figure 13-24, ATP hydrolysis by Hsp70 chaperone proteins in both the cytosol and the mitochondrial matrix is required for the import of mitochondrial proteins. Cytosolic Hsp70 expends energy to maintain bound precursor proteins in an unfolded state so that they can be translocated into the matrix. The importance of ATP to this function was demonstrated in studies in which a mitochondrial precursor protein was purified and then denatured (unfolded) by urea. When tested in the cell-free mitochondrial translocation system, the denatured protein was incorporated into the matrix in the absence of ATP. In contrast, the same precursor protein in its native, undenatured state was not imported in the absence of ATP, even in the presence of cytosolic chaperones.

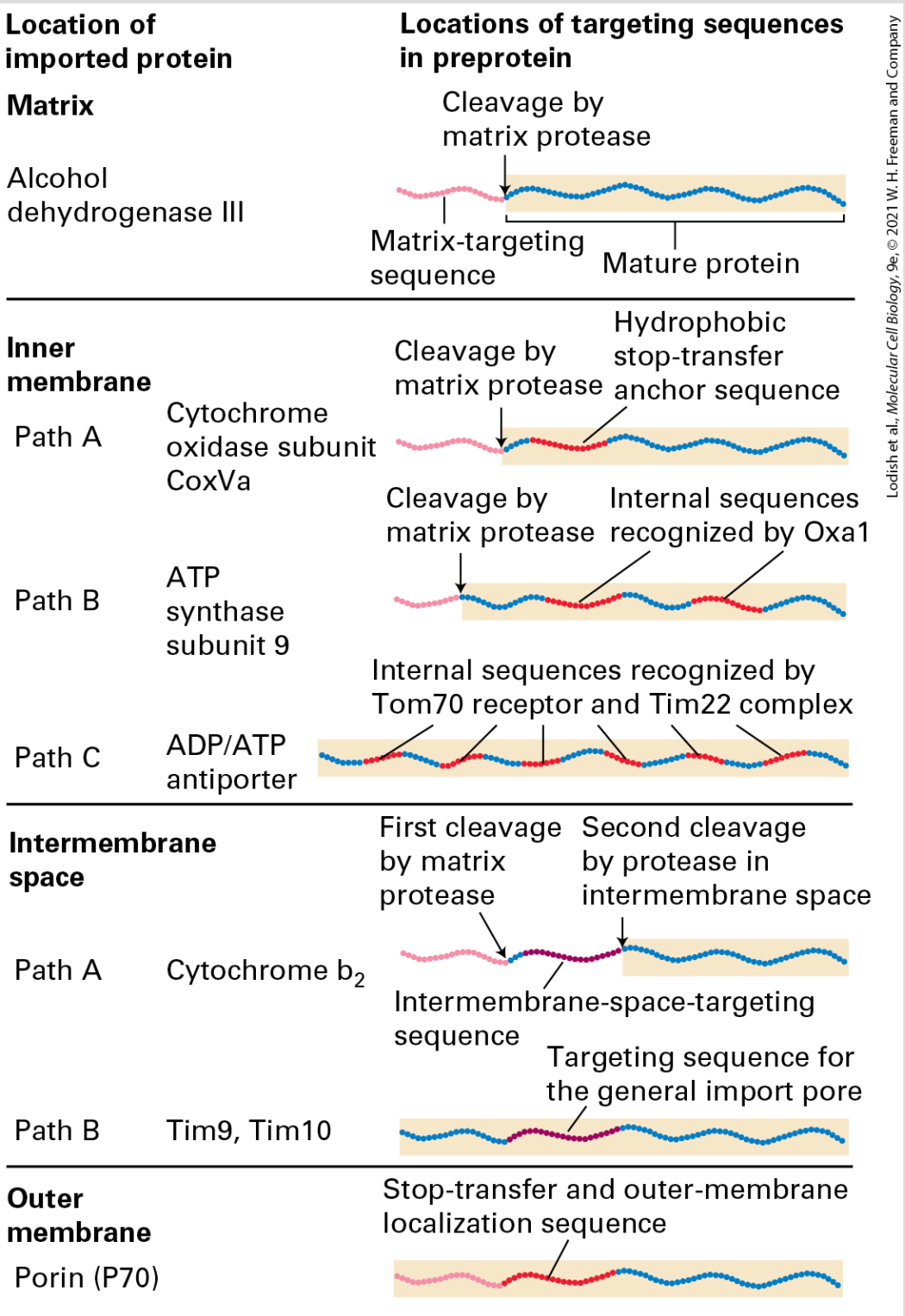
The sequential binding to and ATP-driven release of multiple matrix Hsp70 molecules from a translocating protein may simply trap the unfolded protein in the matrix. In this case, the functions of matrix Hsp70 and Tim44 would be analogous to those of the chaperone BiP and Sec63 complex, respectively, in post-translational translocation into the ER lumen (see Figure 13-9).

The third energy input required for mitochondrial protein import is an H^+ electrochemical gradient, or *proton-motive force*, across the

inner membrane. Recall from Chapter 12 that protons are pumped from the matrix into the intermembrane space during electron transport, creating an electric potential across the inner membrane. In general, only mitochondria that are actively undergoing respiration, and therefore have generated a proton-motive force across the inner membrane, are able to translocate precursor proteins from the cytosol into the mitochondrial matrix. Treatment of mitochondria with inhibitors or uncouplers of oxidative phosphorylation, such as cyanide or dinitrophenol, dissipates this proton-motive force. Although precursor proteins can still bind tightly to receptors on such poisoned mitochondria, the proteins cannot be imported, either in intact cells or in cell-free systems, even in the presence of ATP and chaperone proteins. Once a protein is partially inserted into the inner membrane, it is subjected to a membrane potential of 200 mV (matrix negative). This seemingly small potential is established across the very narrow hydrophobic core of the lipid bilayer, which gives an enormous electrochemical gradient, equivalent to about 400,000 V/cm. The positive charges in the amphipathic matrix-targeting sequence may simply be pulled into the matrix by the inside-negative membrane potential.

Multiple Signals and Pathways Target Proteins to Submitochondrial Compartments

Unlike targeting to the matrix, targeting of proteins to the intermembrane space, inner membrane, or outer membrane of mitochondria generally requires more than one targeting sequence and occurs via one of several pathways. Figure 13-26 summarizes the organization of targeting sequences in proteins sorted to different mitochondrial locations.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-26 Targeting sequences in imported mitochondrial proteins. Most mitochondrial proteins have an N-terminal matrix-targeting sequence (pink) that is similar, though not identical, in different proteins. Proteins destined for

the inner membrane, the intermembrane space, or the outer membrane have one or more additional targeting sequences that function to direct the proteins to these locations by several different pathways. The lettered pathways correspond to those illustrated in Figures 13-27 and 13-28. See W. Neupert, 1997, *Annu. Rev. Biochem.* **66**:863.

Inner-Membrane Proteins

Three separate pathways are known to target proteins to the inner mitochondrial membrane. One pathway makes use of the same machinery that is used for the targeting of matrix proteins (Figure 13-27, path A). A cytochrome oxidase subunit called *CoxVa* is one protein transported by this pathway. The precursor form of CoxVa, which contains an N-terminal matrix-targeting sequence recognized by the Tom20/22 import receptor, is transferred through the Tom40 general import pore of the outer membrane and the inner-membrane Tim23/17 translocation complex. In addition to the matrix-targeting sequence, which is cleaved during import, CoxVa contains a hydrophobic stop-transfer anchor sequence. As the protein passes through the Tim23/17 channel, the stop-transfer anchor sequence blocks translocation of the C-terminus across the inner membrane. The membrane-anchored intermediate is then transferred laterally into the bilayer of the inner membrane, much like type I integral membrane proteins are incorporated into the ER membrane (see Figure 13-11).

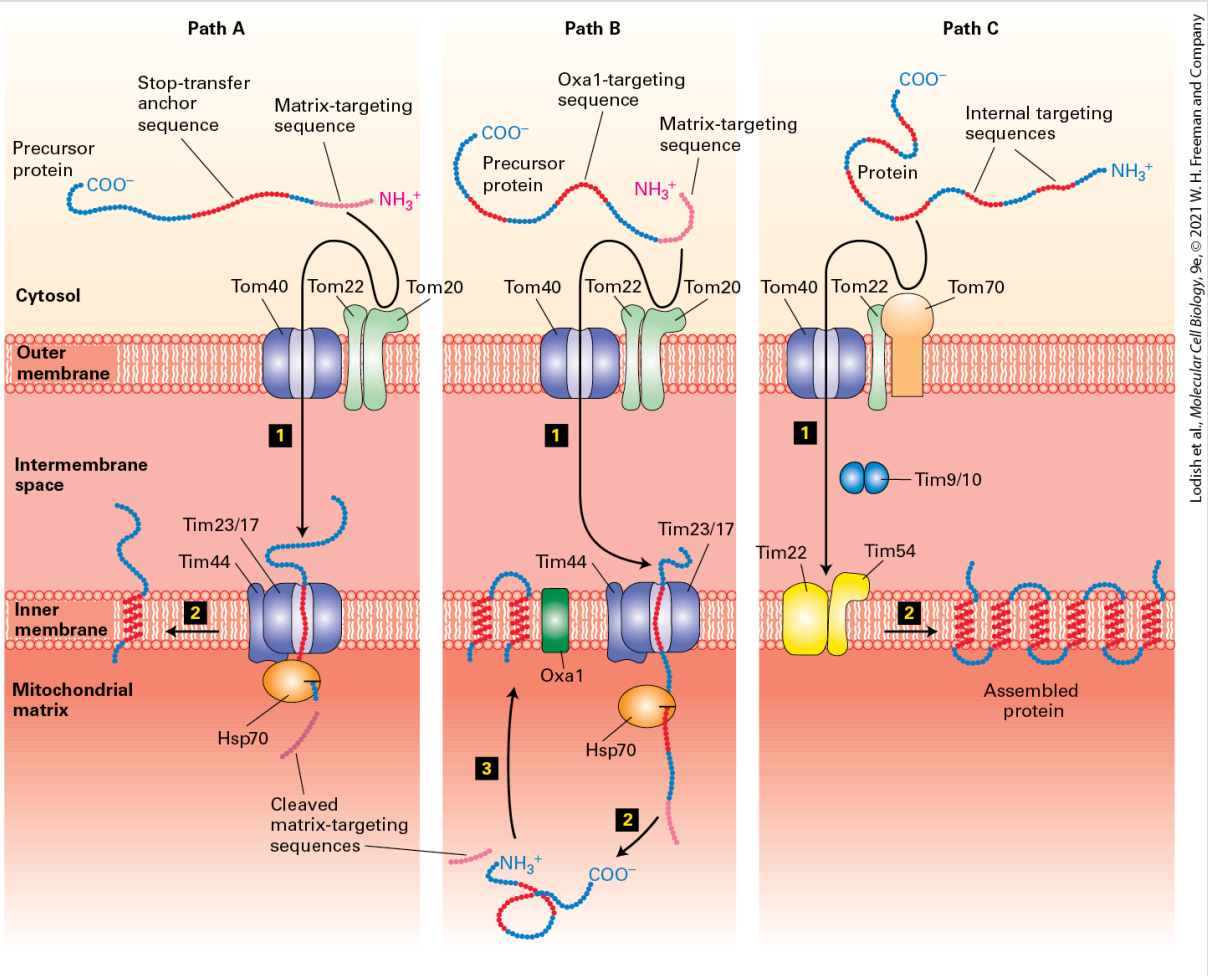


FIGURE 13-27 Three pathways to the inner mitochondrial membrane from the cytosol. Proteins with different targeting sequences are directed to the inner membrane via different pathways. In all three pathways, proteins cross the outer membrane via the Tom40 general import pore. Proteins delivered by pathways A and B contain an N-terminal matrix-targeting sequence that is recognized by the $\mathrm{Tom20/22}$ import receptor in the outer membrane. Although both these pathways use the $\mathrm{Tim23/17}$ inner-membrane channel, they differ in that the entire precursor protein enters the matrix and is then redirected to the inner membrane in pathway B. Matrix Hsp70 plays a role similar to its role in the import of soluble matrix proteins (see Figure 13-23). Proteins delivered by pathway C contain internal sequences that are recognized by the $\mathrm{Tom70/Tom22}$ import receptor; a different inner-membrane translocation channel ($\mathrm{Tim22/54}$) is used in this pathway. Two intermembrane proteins (Tim9 and Tim10) facilitate transfer between the outer and inner channels. See the text for discussion. See R. E. Dalbey and A. Kuhn, 2000, *Annu. Rev. Cell Dev. Biol.* **16**:51; and N. Pfanner and A. Geissler, 2001, *Nat. Rev. Mol. Cell Biol.* **2**:339.

A second pathway to the inner membrane is used by proteins such as ATP synthase subunit 9, whose precursors contain both a matrix-targeting sequence and internal hydrophobic domains recognized by an inner-membrane protein termed *Oxa1*. This pathway is thought to involve translocation of at least a portion of the precursor into the matrix via the Tom40 and Tim23/17 channels. After cleavage of the matrix-targeting sequence, the protein is inserted into the inner membrane by a process that requires interaction with Oxa1 and perhaps other inner-membrane proteins (see Figure 13-27, path B). Oxa1 is related to a protein involved in inserting some plasma-membrane proteins in bacteria. This relatedness suggests that Oxa1 may have descended from the translocation machinery in the endosymbiotic bacterium that eventually became the mitochondrion. However, the proteins that form the inner-membrane channels in mitochondria are not related to the proteins in bacterial translocons. Oxa1 also participates in the inner-membrane insertion of certain proteins, such as subunit II of cytochrome oxidase, that are encoded by mitochondrial DNA and synthesized in the matrix by mitochondrial ribosomes.

The final pathway for insertion in the inner mitochondrial membrane is followed by multipass proteins that contain six membrane-spanning domains, such as the ATP/ADP antiporter. These proteins, which lack the usual N-terminal matrix-targeting sequence, contain multiple internal mitochondrial targeting sequences. After the internal sequences are recognized by a second

import receptor composed of outer-membrane proteins Tom70 and Tom22, the imported protein passes through the outer membrane via the general import pore (see Figure 13-27, path C). The protein then is transferred to a second translocation complex in the inner membrane composed of the Tim22, Tim18, and Tim54 proteins. Transfer to the Tim22/18/54 complex depends on a multimeric complex of two small proteins, Tim9 and Tim10, which reside in the intermembrane space. These small Tim proteins are thought to act as chaperones, guiding imported protein precursors from the general import pore to the Tim22/18/54 complex in the inner membrane by binding to their hydrophobic regions, preventing them from forming insoluble aggregates in the aqueous environment of the intermembrane space. Ultimately the Tim22/18/54 complex is responsible for incorporating the multiple hydrophobic segments of the imported protein into the inner membrane.

Intermembrane-Space Proteins

Two pathways deliver cytosolic proteins to the space between the inner and outer mitochondrial membranes. The major pathway is followed by proteins such as cytochrome b_2, whose precursors carry two different N-terminal targeting sequences, both of which are ultimately cleaved. The most N-terminal of the two sequences is a matrix-targeting sequence, which is removed by the matrix protease. The second targeting sequence is a hydrophobic segment

that blocks complete translocation of the protein across the inner membrane (Figure 13-28, path A). After the resulting membrane-embedded intermediate diffuses laterally away from the Tim23/17 translocation channel, a protease in the membrane cleaves the protein near the hydrophobic transmembrane segment, releasing the mature protein in a soluble form into the intermembrane space. Except for the second proteolytic cleavage, this pathway is similar to that of inner-membrane proteins such as CoxVa (see Figure 13-27, path A).

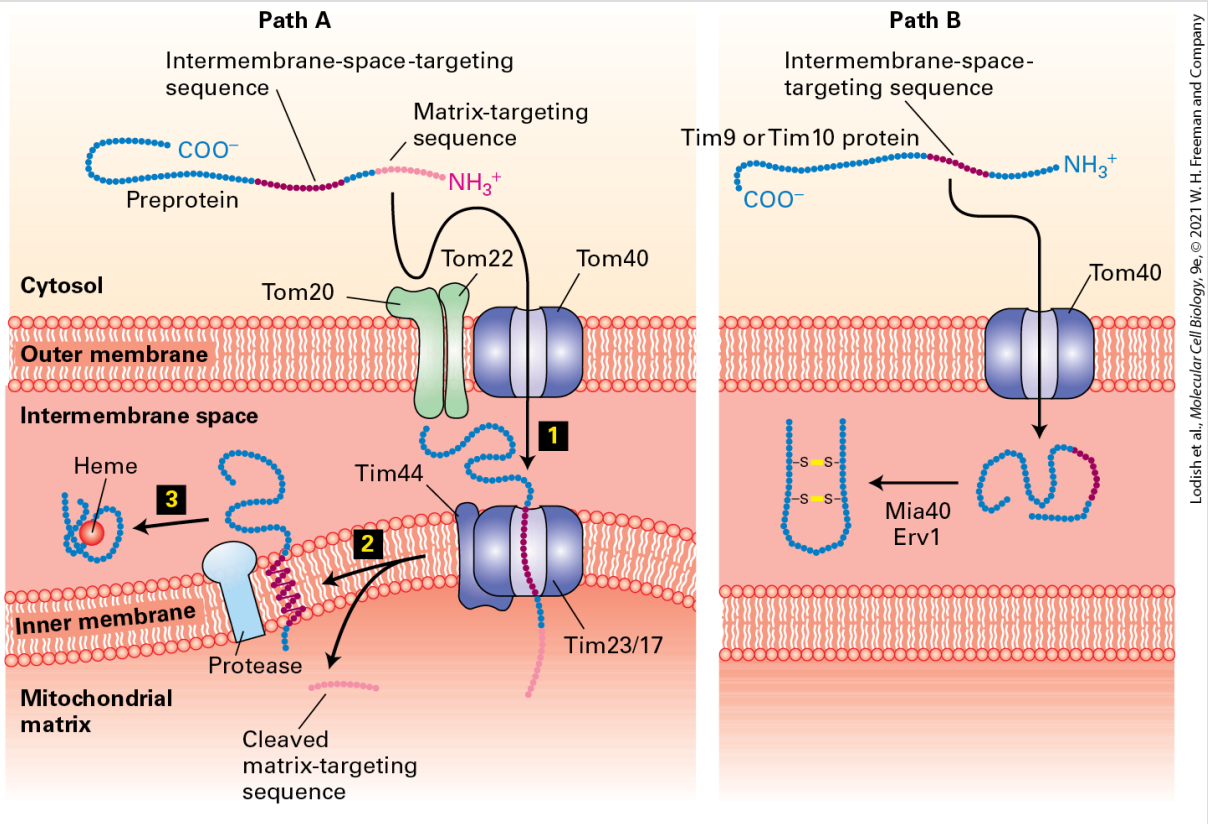


FIGURE 13-28 Two pathways to the mitochondrial intermembrane space. Pathway A, the major one for delivery of proteins from the cytosol to the intermembrane space, is similar to pathway A for delivery to the inner membrane (see Figure 13-26). The major difference is that the internal targeting sequence in these proteins, such as cytochrome b_2, is recognized

by an inner-membrane protease, which cleaves the protein on the intermembrane-space side of the membrane. The released protein then folds and binds to its heme cofactor within the intermembrane space. Pathway B is a specialized pathway for delivery of the proteins Tim9 and Tim10 to the intermembrane space. These proteins readily pass through the Tom40 general import pore, and once they are in the intermembrane space, they fold and form disulfide bonds that prevent reverse translocation through Tom40. The disulfide bonds are generated by Erv1 and are transferred to Tim9 and Tim10 by Mia40. See R. E. Dalbey and A. Kuhn, 2000, *Annu. Rev. Cell Dev. Biol.* **16**:51; N. Pfanner and A. Geissler, 2001, *Nat. Rev. Mol. Cell Biol.* **2**:339; and K. Tokatlidis, 2005, *Cell* **121**:965–967.

The small Tim9 and Tim10 proteins, which reside in the intermembrane space, illustrate a second pathway for targeting to the intermembrane space. Tim9 and Tim10 do not contain an N-terminal matrix-targeting sequence and are delivered directly to the intermembrane space via the general import pore without involvement of any inner-membrane translocation factors (see Figure 13-28, path B). Translocation through the Tom40 general import pore does not seem to be coupled to any energetically favorable process; however, once located in the intermembrane space, Tim9 and Tim10 proteins acquire two disulfide bonds each and fold into compact, stable structures. Apparently, the mechanism that drives their unidirectional translocation through the outer membrane involves passive diffusion through the outer membrane followed by folding and disulfide bond formation, which irreversibly traps the proteins in the intermembrane space. The process of disulfide bond formation in the intermembrane space resembles that in the ER lumen and involves a disulfide bond–generating protein, Erv1, and a disulfide transfer protein, Mia40.

Outer-Membrane Proteins

Many of the proteins that reside in the mitochondrial outer membrane, including the Tom40 pore itself and mitochondrial porin, have a β-barrel structure in which antiparallel β strands form hydrophobic transmembrane segments surrounding a central channel. Such proteins are incorporated into the outer membrane by first interacting with the general import pore, Tom40; they are then transferred to the SAM (*s*orting and *a*ssembly *m*achinery) complex, which is composed of at least three outer-membrane proteins. Presumably it is the very stable hydrophobic nature of β-barrel proteins that ultimately causes them to be stably incorporated into the outer membrane, but precisely how the SAM complex facilitates this process is not known.

Import of Chloroplast Stromal Proteins Is Similar to Import of Mitochondrial Matrix Proteins

The most abundant proteins found in the chloroplast stroma are the enzymes of the Calvin cycle, which function in fixing carbon dioxide during photosynthesis (see Chapter 12). The large (L) subunit of ribulose 1,5-bisphosphate carboxylase (rubisco) is encoded by chloroplast DNA and synthesized on chloroplast ribosomes in the stromal space. The small (S) subunit of rubisco and all the other

Calvin cycle enzymes are encoded by nuclear genes and transported to chloroplasts after their synthesis in the cytosol. The precursor forms of these stromal proteins contain an N-terminal *stromal-import* sequence (see Table 13-1).

Experiments with isolated chloroplasts, similar to those with mitochondria, have shown that the general process of stromal import depicted in the top half of Figure 13-29 is similar to that of protein import into the mitochondrial matrix. At least three chloroplast outer-membrane proteins are known to be essential for directing proteins to the stroma. These include a receptor that binds the stromal-import sequence made up of proteins, known as *Toc159* and *Toc34*, each of which has a GTPase domain on the cytosolic side of the outer membrane reminiscent of SRP and SRP receptor. The receptor transfers the stromal-import sequence to a translocation channel made up of Toc75. A complex of at least five chloroplast inner-membrane proteins known as the *Tic complex* forms the translocation channel for passage of proteins into the stromal space. Although these proteins are functionally analogous to the receptor and channel proteins of the mitochondria, they are not structurally homologous. The lack of sequence similarity between these chloroplast and mitochondrial proteins suggests that they may have arisen independently during evolution.

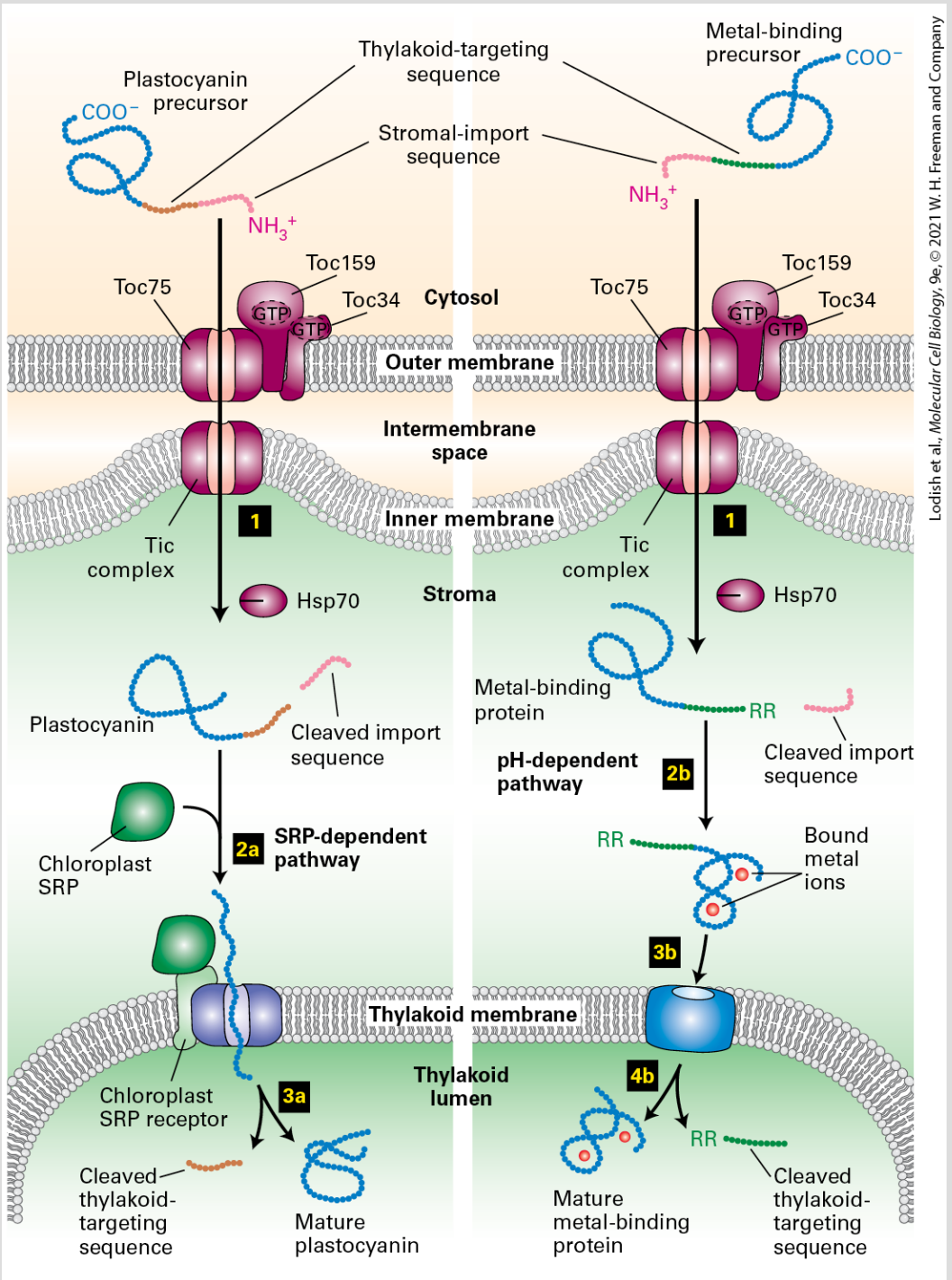


FIGURE 13-29 Transporting proteins to chloroplast thylakoids. Two of the four pathways for transporting proteins from the cytosol to the thylakoid lumen are shown here. In these pathways, unfolded precursors are delivered to the stroma via the same outer-membrane proteins that import stromal-localized proteins. Cleavage of the N-terminal stromal-import sequence by a stromal

protease then reveals the thylakoid-targeting sequence (step **1**). At this point the two pathways diverge. In the SRP-dependent pathway (*left*), plastocyanin and similar proteins are kept unfolded in the stromal space by a set of chaperones (not shown), and the thylakoid-targeting sequence binds to proteins that are closely related to the bacterial SRP, SRP receptor, and SecY translocon, which mediate movement into the thylakoid lumen (step **2a**). After the thylakoid-targeting sequence is removed in the thylakoid lumen by a separate endoprotease, the protein folds into its mature conformation (step **3a**). In the pH-dependent pathway (*right*), metal-binding proteins fold in the stroma, and complex redox cofactors are added (step **2b**). Two arginine residues (RR) at the N-terminus of the thylakoid-targeting sequence and a pH gradient across the inner membrane are required for transport of the folded protein into the thylakoid lumen (step **3b**). The translocon in the thylakoid membrane is composed of at least four proteins related to proteins in the bacterial plasma membrane. The thylakoid-targeting sequence containing the two arginine residues is cleaved in the thylakoid lumen (step **4b**). See R. Dalbey and C. Robinson, 1999, *Trends Biochem. Sci.* **24**:17; R. E. Dalbey and A. Kuhn, 2000, *Annu. Rev. Cell Dev. Biol.* **16**:51; and C. Robinson and A. Bolhuis, 2001, *Nat. Rev. Mol. Cell Biol.* **2**:350.

The available evidence suggests that chloroplast stromal proteins, like mitochondrial matrix proteins, are imported in the unfolded state. Import into the stroma depends on ATP hydrolysis catalyzed by a stromal Hsp70 chaperone whose function is similar to that of Hsp70 in the mitochondrial matrix and BiP in the ER lumen. Unlike mitochondria, chloroplasts do not generate an electrochemical gradient (proton-motive force) across their inner membrane. Thus protein import into the chloroplast stroma appears to be powered solely by ATP hydrolysis.

In addition to Hsp70, the stroma contains two other ATPases that may participate in protein import. These are a chloroplast version of the chaperone Hsp90, and a ring forming AAA+ ATPase known as *Hsp93*. Homologs of both proteins are known to catalyze energy-dependent protein-folding reactions in other contexts and it may be that chloroplast Hsp90 and Hsp93 participate in protein folding in the stroma, protein import into the stroma, or both.

Proteins Are Targeted to Thylakoids by Mechanisms Related to Bacterial Protein Translocation

In addition to the double membrane that surrounds them, chloroplasts contain a series of internal interconnected membranous sacs, the **thylakoids** (see Figure 12-33). All of the chemical reactions of photosynthesis take place in the thylakoid membrane or lumen and are catalyzed by the proteins that are localized to this specialized subcompartment. Many of these proteins are synthesized in the cytosol as precursors containing multiple targeting sequences. For example, plastocyanin and other proteins destined for the thylakoid lumen require the successive action of two targeting sequences. The first is an N-terminal stromal-import sequence that directs the protein to the stroma by the same pathway that imports the rubisco S subunit. The second sequence targets the protein from the stroma to the thylakoid

lumen. The role of these targeting sequences has been shown in experiments measuring the uptake of mutant proteins generated by recombinant DNA techniques into isolated chloroplasts. For instance, mutant plastocyanin that lacks the thylakoid-targeting sequence but contains an intact stromal-import sequence accumulates in the stroma and is not transported into the thylakoid lumen.

Four separate pathways for transporting proteins from the stroma into the thylakoid have been identified. All four pathways have been found to be closely related to analogous transport mechanisms in bacteria, illustrating the close evolutionary relationship between the thylakoid membrane and the bacterial plasma membrane. Transport of plastocyanin and related proteins into the thylakoid lumen from the stroma occurs by an SRP-dependent pathway that uses a translocon similar to SecY, the bacterial version of the Sec61 complex (Figure 13-29, *left*). A second pathway for transporting proteins into the thylakoid lumen involves a protein related to bacterial protein SecA, which uses the energy from ATP hydrolysis to drive protein translocation through the SecY translocon. A third pathway, which targets proteins to the thylakoid membrane, depends on a protein related to the mitochondrial Oxa1 protein and the homologous bacterial protein (see Figure 13-27, path B). Some proteins encoded by chloroplast DNA and synthesized in the stroma or transported into the stroma from the cytosol are inserted into the thylakoid membrane via this pathway.

Finally, thylakoid proteins that bind metal-containing cofactors follow another pathway into the thylakoid lumen (see Figure 13-29, *right*). The unfolded precursors of these proteins are first targeted to the stroma, where the N-terminal stromal-import sequence is cleaved off, and the protein then folds and binds its cofactor. A set of thylakoid-membrane proteins assists in translocating the folded protein and bound cofactor into the thylakoid lumen. This process is powered by the H^+ electrochemical gradient normally maintained across the thylakoid membrane. The thylakoid-targeting sequence that directs a protein to this pathway includes two closely spaced arginine residues that are crucial for recognition. Bacterial cells also have a mechanism for translocating folded proteins with a similar arginine-containing sequence across the plasma membrane, known as the *Tat* (*t*win-*a*rginine *t*ranslocation) pathway. The molecular mechanism whereby these large folded globular proteins are translocated across the thylakoid membrane is not fully understood, but the presence of a folded protein with an appropriate twin arginine signal appears to induce the oligomerization of Tat proteins in the membrane to form pore-like structures. In this respect, the Tat pathway resembles the pathway for the import of folded proteins into peroxisomes, described in the next section.

KEY CONCEPTS OF SECTION 13.4

Targeting of Proteins to Mitochondria and Chloroplasts

- Most mitochondrial and chloroplast proteins are encoded by nuclear genes, synthesized on cytosolic ribosomes, and imported post-translationally into the

organelles.

- All the information required to target a precursor protein from the cytosol to the mitochondrial matrix or chloroplast stroma is contained within its N-terminal targeting sequence. After protein import, the targeting sequence is removed by proteases within the matrix or stroma.
- Cytosolic chaperones maintain the precursors of mitochondrial and chloroplast proteins in an unfolded state. Only unfolded proteins can be imported into the organelles.
- Proteins destined for the mitochondrial matrix bind to receptors on the outer mitochondrial membrane and are then transferred to the general import pore (Tom40) in the outer membrane. Translocation takes place at sites where the outer and inner membranes of the organelles are close together and occurs through the outer and inner membranes concurrently. Translocation is driven by ATP hydrolysis by Hsp70 in the matrix (see Figure 13-24) and by the proton-motive force across the inner membrane.
- Proteins sorted to mitochondrial destinations other than the matrix usually contain two or more targeting sequences, one of which may be an N-terminal matrix-targeting sequence (see Figure 13-26).
- Some mitochondrial proteins destined for the intermembrane space or inner membrane are first imported into the matrix and then redirected; others never enter the matrix, but go directly to their final location.
- Protein import into the chloroplast stroma occurs through outer-membrane and inner-membrane translocation channels that are analogous in function to mitochondrial channels, but composed of proteins unrelated in sequence to the corresponding mitochondrial proteins.
- Proteins destined for the thylakoid have secondary targeting sequences. After entry of these proteins into the stroma, cleavage of the stromal-targeting sequences reveals the thylakoid-targeting sequences.
- The four known pathways for moving proteins from the chloroplast stroma to the thylakoid closely resemble translocation across the bacterial plasma membrane (see Figure 13-29). One of these systems can translocate folded proteins.

13.5 Targeting of Peroxisomal Proteins

Peroxisomes are small organelles bounded by a single membrane which surrounds a luminal space known as the *peroxisomal matrix*. Although peroxisomes have a simple structure, the biogenesis of peroxisomes is similar in fundamental outline to the biogenesis of the more complicated mitochondria and chloroplasts. Peroxisomal proteins, encoded by nuclear genes and synthesized on free ribosomes in the cytosol, are recognized by specific peroxisomal-targeting sequences. These sequences bind to a peroxisome-specific receptor protein and are then delivered either to the peroxisomal matrix or to the peroxisomal membrane by an energy-dependent targeting process.

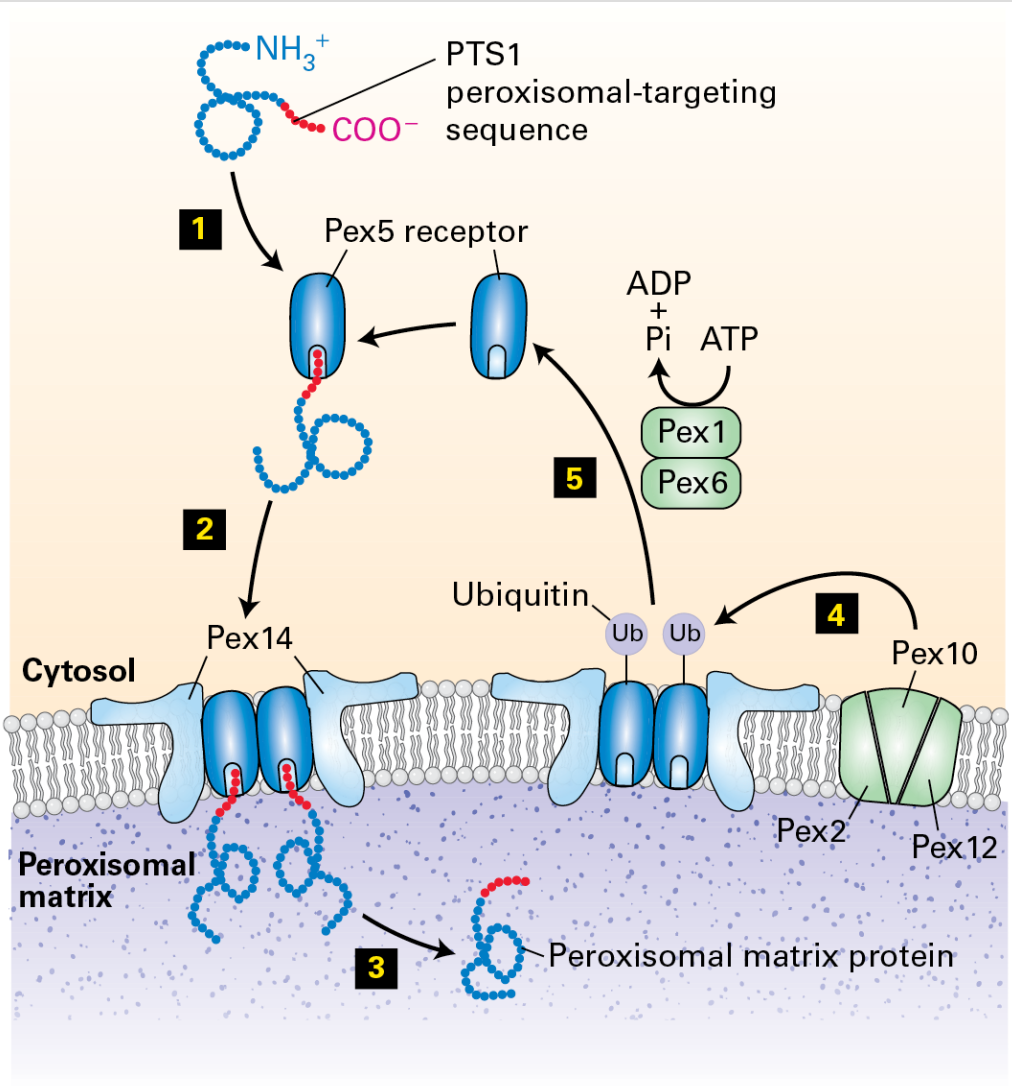
The size and enzyme content of peroxisomes vary considerably among different kinds of cells. However, all peroxisomes contain enzymes that use molecular oxygen to oxidize various substrates such as fatty acids, breaking them down into smaller components for use in biosynthetic pathways. The hydrogen peroxide (H_2O_2) generated by these oxidation reactions is extremely reactive and potentially harmful to cellular components; however, the peroxisome contains other enzymes, such as catalase, that efficiently convert H_2O_2 into H_2O. In mammals, peroxisomes are most abundant in liver cells, where they constitute about 1–2 percent of the cell volume.

A Cytosolic Receptor Targets Proteins with an SKL Sequence at the C-Terminus to the Peroxisomal Matrix

Peroxisomal-targeting sequences were first identified by testing of peroxisomal proteins with deletions for a specific defect in peroxisomal targeting. In one early study, the gene for firefly luciferase was expressed in cultured insect cells, and the resulting protein was shown to be properly targeted to the peroxisome. However, expression of a gene missing a small portion of the sequence encoding the C-terminus of the protein led to luciferase that failed to be targeted to the peroxisome and remained in the cytoplasm. By testing various mutant luciferase proteins in this system, researchers discovered that the sequence Ser-Lys-Leu (SKL in one-letter code) or a related sequence at the C-terminus is necessary for peroxisomal targeting. Furthermore, addition of the SKL sequence to the C-terminus of a normally cytosolic protein leads to uptake of the altered protein by peroxisomes in cultured cells. All but a few of the many different peroxisomal proteins bear a sequence of this type, known as *peroxisomal-targeting sequence 1*, or simply *PTS1*.

The pathway for the import of catalase and other PTS1-bearing proteins into the peroxisomal matrix is depicted in Figure 13-30. In the cytosol, PTS1 binds to a receptor called *Pex5*. Pex5 has the remarkable property of being able to switch from a monomeric

soluble form to an oligomeric form embedded in the peroxisomal membrane in a complex with the membrane protein Pex14. In a manner that is not well understood, the PTS1-bearing protein is released from the oligomeric form of Pex5 into the interior of the peroxisome. This peroxisome import machinery, unlike most systems that mediate protein import into the ER, mitochondria, and chloroplasts, can translocate folded proteins across the membrane. For example, catalase assumes a folded conformation and binds to heme in the cytoplasm before traversing the peroxisomal membrane. Cell-free studies have shown that the peroxisome import machinery can transport large macromolecular objects, including gold particles about 9 nm in diameter, as long as they have a PTS1 tag attached to them. There is evidence that the size of oligomers of Pex5 bound to PTS1-bearing cargo molecules and Pex14 adjusts according to the size of the PTS1-bearing cargo molecules. The dynamic formation of oligomers is apparently the key mechanism by which PTS1-bearing cargo molecules can be accommodated without the formation of large stable pores that would disrupt the integrity of the peroxisomal membrane.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-30 PTS1-directed import of peroxisomal matrix proteins. Step 1: Most peroxisomal matrix proteins contain a C-terminal PTS1 targeting sequence (red), which binds to the cytosolic receptor Pex5. Step 2: Pex5 with the bound matrix protein forms a multimeric complex with the Pex14 receptor located on the peroxisomal membrane. Step 3: After assembly of the matrix protein-Pex5-Pex14 complex, the matrix protein dissociates from Pex5 and is released into the peroxisomal matrix. Steps 4 and 5: Pex5 is then returned to the cytosol by a process that involves ubiquitinylation by the membrane proteins Pex2, Pex10, and Pex12, followed by ATP-dependent removal from the membrane by the AAA-ATPase proteins Pex1 and Pex6. Note that folded proteins can be imported into peroxisomes and that the targeting sequence is not removed in the matrix. See P. E. Purdue and P. B. Lazarow, 2001, *Annu. Rev.*

Cell Dev. Biol. **17**:701; S. Subramani et al., 2000, *Annu. Rev. Biochem.* **69**:399; and V. Dammai and S. Subramani, 2001, *Cell* **105**:187.

Once the PTS1-bearing cargo molecule is released into the interior of the peroxisome, the oligomeric complex of Pex5 and Pex14 is actively disassembled, thus releasing Pex5 back into the cytoplasm in a soluble state. Pex5 recycling involves modification of membrane-bound Pex5 by ubiquitinylation. A complex of the peroxisomal membrane proteins Pex10, Pex12, and Pex2 transfers a ubiquitin moiety to Pex5. The AAA ATPases Pex1 and Pex6, anchored to the peroxisomal membrane by Pex15, recognize ubiquitinylated Pex5 and use the energy from ATP hydrolysis to remove it from the oligomeric complex with Pex14 and release it into the cytosol. After removal of the ubiquitin modification, cytosolic Pex5 is ready to carry out another cycle of binding to a PTS1-bearing protein.

Peroxisomal import studies with purified components have shown that the binding of Pex5 to a PTS1-bearing protein, the assembly of an oligomeric complex of Pex5 and Pex14, and release of the PTS1-bearing protein into the interior of the peroxisome can all occur spontaneously without a source of chemical energy such as ATP. In contrast, both ubiquitin modification of Pex5 and the recycling of Pex5 by the AAA-ATPase require ATP hydrolysis. Evidently, in the process by which Pex5 delivers PTS1-bearing proteins to the peroxisome matrix by cycling between the cytosol and the peroxisomal membrane, it is the recycling step that uses energy to

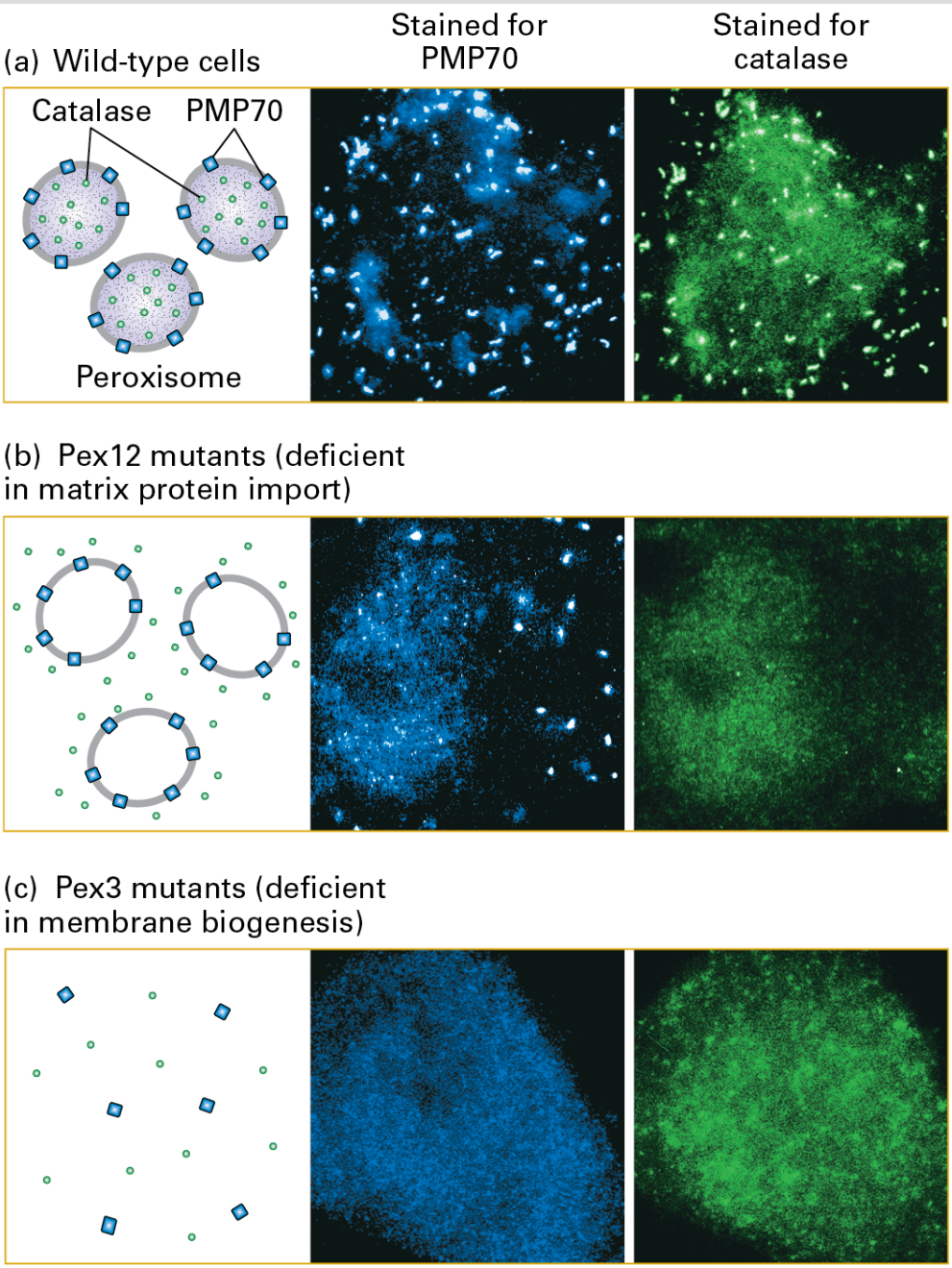
power unidirectional translocation of cargo molecules into the peroxisomal matrix.

A few peroxisomal matrix proteins, such as thiolase, are synthesized as precursors with an N-terminal targeting sequence known as *PTS2*. These proteins bind to a different cytosolic receptor protein, but otherwise their import is thought to occur by the same mechanism as for PTS1-containing proteins.

Peroxisomal Membrane and Matrix Proteins Are Incorporated by Different Pathways

Autosomal recessive mutations that cause defective peroxisome assembly occur naturally in the human population. Such mutations can lead to severe developmental defects often associated with craniofacial abnormalities. In *Zellweger syndrome* and related disorders, for example, the transport of many or all proteins into the peroxisomal matrix is impaired: newly synthesized peroxisomal enzymes remain in the cytosol and are eventually degraded. Genetic analyses of cultured cells from patients with Zellweger syndrome and of yeast cells carrying similar mutations have identified more than 20 genes that are required for peroxisome biogenesis.

Studies with peroxisome-assembly mutants have shown that different pathways are used for importing peroxisomal matrix proteins and for inserting proteins into the peroxisomal membrane (Figure 13-31). For example, analysis of cells from some patients with Zellweger syndrome led to the identification of the gene encoding Pex5 as well as many of the Pex genes needed for recycling of Pex5. Mutant cells that are defective in any one of these proteins cannot incorporate matrix proteins into peroxisomes; nonetheless, the cells contain empty peroxisomes that have a normal complement of peroxisomal membrane proteins (Figure 13-31b). Mutations in any one of three other genes were found to block insertion of peroxisomal membrane proteins as well as import of matrix proteins (Figure 13-31c). These findings demonstrate that one set of proteins translocates soluble proteins into the peroxisomal matrix, but a different set is required for insertion of proteins into the peroxisomal membrane. This situation differs markedly from that of the ER, mitochondrion, and chloroplast, whose membrane proteins and soluble proteins share many of the same components for their import into these organelles.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Stephen Gould, Johns Hopkins University School of Medicine.

EXPERIMENTAL FIGURE 13-31 Studies reveal different pathways for incorporation of peroxisomal membrane and matrix proteins. Cells were stained with fluorescent antibodies to PMP70, a peroxisomal membrane protein, or with fluorescent antibodies to catalase, a peroxisomal matrix protein, then viewed in a fluorescence microscope. (a) In wild-type cells, both peroxisomal membrane and matrix proteins are visible as bright foci in numerous peroxisomal bodies. (b) In cells from a Pex12-deficient patient,

PMP70 is localized normally to peroxisomal bodies, but catalase is not delivered to peroxisomes and is distributed uniformly throughout the cytosol with no bright foci, whereas (c) in cells from a Pex3-deficient patient, peroxisomal membranes cannot assemble, and as a consequence, peroxisomal bodies do not form. Thus both catalase and PMP70 are mislocalized to the cytosol.

KEY CONCEPTS OF SECTION 13.5

Targeting of Peroxisomal Proteins

- All luminal peroxisomal proteins are synthesized on free cytosolic ribosomes and incorporated into the organelle post-translationally.
- Most peroxisomal matrix proteins contain a C-terminal targeting sequence known as *PTS1*; a few have an N-terminal PTS2 targeting sequence. Neither targeting sequence is cleaved after import.
- All proteins destined for the peroxisomal matrix bind to a cytosolic receptor protein, which is Pex5 for PTS1-bearing proteins, and then are directed to common translocation machinery on the peroxisomal membrane (see Figure 13-30).
- Translocation of matrix proteins across the peroxisomal membrane depends on ATP hydrolysis, which is coupled to recycling of Pex5 from the peroxisomal membrane back to the cytosol.
- Unlike proteins imported to the ER, mitochondrion, or chloroplast, many peroxisomal matrix proteins fold in the cytosol and traverse the membrane in a folded conformation.
- Proteins destined for the peroxisomal membrane contain different targeting sequences than peroxisomal matrix proteins and are imported by a different pathway.

13.6 Transport into and out of the Nucleus

The nucleus is separated from the cytoplasm by two membranes, which form the nuclear envelope (see Figure 1-12a). The nuclear envelope is continuous with the ER and forms a part of it. Transport of proteins from the cytoplasm into the nucleus and movement of macromolecules, including mRNAs, tRNAs, and ribosomal subunits, out of the nucleus occurs through *nuclear pores*, which span both membranes of the nuclear envelope. Import of proteins into the nucleus shares some fundamental features with protein import into other organelles. For example, imported nuclear proteins carry specific targeting sequences known as *nuclear-localization signals (NLSs)*. However, proteins are imported into the nucleus in a folded state, and thus nuclear import differs fundamentally from protein translocation across the membranes of the ER, mitochondrion, and chloroplast, during which proteins are unfolded. In this section, we discuss the main mechanism by which proteins enter and exit the nucleus. We also discuss the process by which mRNAs and other ribonuclear protein complexes are exported from the nucleus, which differs mechanistically from nuclear protein import.

Large and Small Molecules Enter and Leave the Nucleus via Nuclear Pore

Complexes

Numerous nuclear pores perforate the nuclear envelope in all eukaryotic cells. Each nuclear pore is formed from an elaborate structure termed the **nuclear pore complex (NPC)**, which is one of the largest discrete protein assemblages in the cell. The total mass of the pore structure is 60,000–80,000 kDa in vertebrates, which is about 16 times larger than a ribosome. An NPC is made up of many copies of some 30 different proteins called **nucleoporins**. Electron micrographs of nuclear pore complexes reveal a membrane-embedded ring structure that surrounds a largely aqueous pore (Figure 13-32). Eight approximately 100-nm long filaments extend into the nucleoplasm with the distal ends of these filaments joined by a terminal ring, forming a structure called the *nuclear basket*. Cytoplasmic filaments extend from the cytoplasmic side of the NPC into the cytosol.

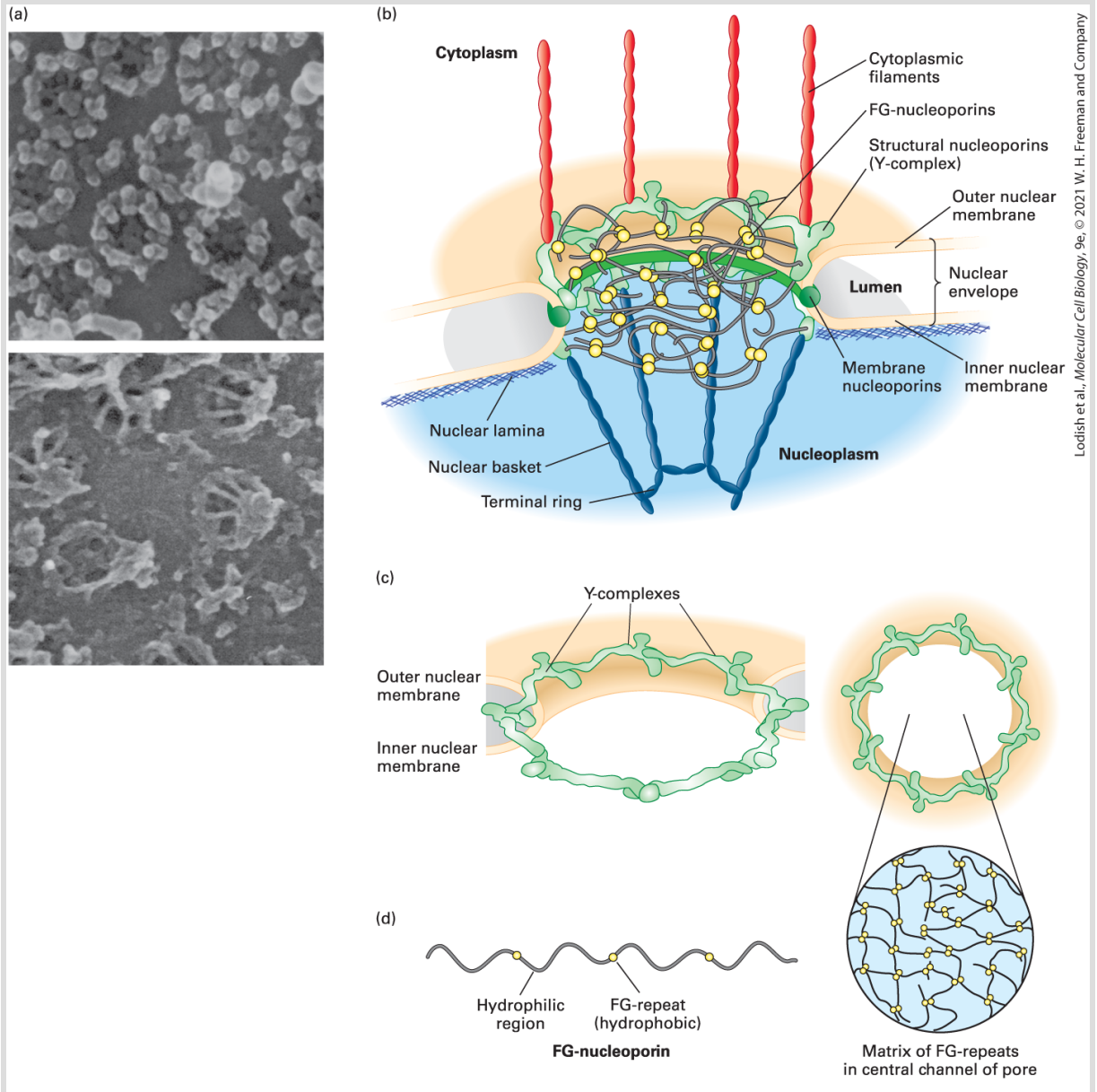


FIGURE 13-32 Nuclear pore complex at different levels of resolution. (a) Nuclear envelopes from the large nuclei of *Xenopus* oocytes, visualized by scanning electron microscopy. *Top:* View of the cytoplasmic face reveals the octagonal shape of the membrane-embedded portion of nuclear pore complexes. *Bottom:* View of the nucleoplasmic face shows the nuclear basket that extends from the membrane-embedded portion. (b) Cutaway model of the nuclear pore complex, showing the major structural features formed by membrane nucleoporins, structural nucleoporins, and FG-nucleoporins. (c) Sixteen copies of the Y-complex form a major part of the structural scaffold of the nuclear pore complex. The three-dimensional structure of the Y-complex is modeled into the pore structure. Note the twofold symmetry across the double membrane of the nucleus (*left*) and the eightfold rotational

symmetry around the axis of the pore (*right*). (d) The FG-nucleoporins have extended disordered structures that are composed of repeats of the sequence Phe–Gly interspersed with hydrophilic regions (*left*). The FG-nucleoporins are most abundant in the central part of the pore, and the FG-repeat sequences are thought to fill the central channel with a gel-like matrix (*right*). See K. Ribbeck and D. Görlich, 2001, *EMBO J.* **20**:1320–1330; and M. P. Rout and J. D. Atchison, 2001, *J. Biol. Chem.* **276**:16593.

[Part (a) Republished with permission from Elsevier, from V. Doye and E. Hurt, 1997, Nucleoporins to Nuclear Pore Complexes, Curr. Opin. Cell Biol. 9(3):401–411; permission conveyed through Copyright Clearance Center, Inc.]

Ions, small metabolites, and globular proteins up to about 40 kDa can diffuse passively through the central aqueous region of the nuclear pore complex. However, large proteins and ribonucleoprotein complexes cannot diffuse in and out of the nucleus. Rather, these macromolecules are actively transported through the NPC with the assistance of soluble transport proteins that bind macromolecules and also interact with nucleoporins. The capacity and efficiency of the NPC for such active transport is remarkable. In one minute, *each* NPC is estimated to import 60,000 protein molecules into the nucleus, while exporting 50–250 mRNA molecules, 10–20 ribosomal subunits, and 1000 tRNAs out of the nucleus.

In general terms, the nucleoporins are of three types: *structural nucleoporins, membrane nucleoporins*, and *FG-nucleoporins*. The structural nucleoporins form the scaffold of the nuclear pore, which is a ring of eightfold rotational symmetry that traverses both membranes of the nuclear envelope, creating an opening in the

middle for movement of macromolecules into and out of the nucleus. The inner and outer membranes of the nuclear envelope are connected at the NPC by a highly curved region of membrane that contains the embedded membrane nucleoporins (Figure 13-32b). A set of seven structural nucleoporins forms a Y-shaped structure about the size of the ribosome, known as the *Y-complex*. Sixteen copies of the Y-complex form the basic structural scaffold of the pore, which has bilateral symmetry across the nuclear envelope and eightfold rotational symmetry in the plane of the envelope (Figure 13-32c). A structural motif repeated several times within the Y-complex is closely related to a structure found in the COPII proteins that drive the formation of coated vesicles within cells (see Chapter 14). This primordial relationship between structural nucleoproteins and vesicle coat proteins suggests that the two types of membrane coat complexes share a common origin. The basic function of this element may be to form a protein lattice that deforms the membrane into a highly curved structure.

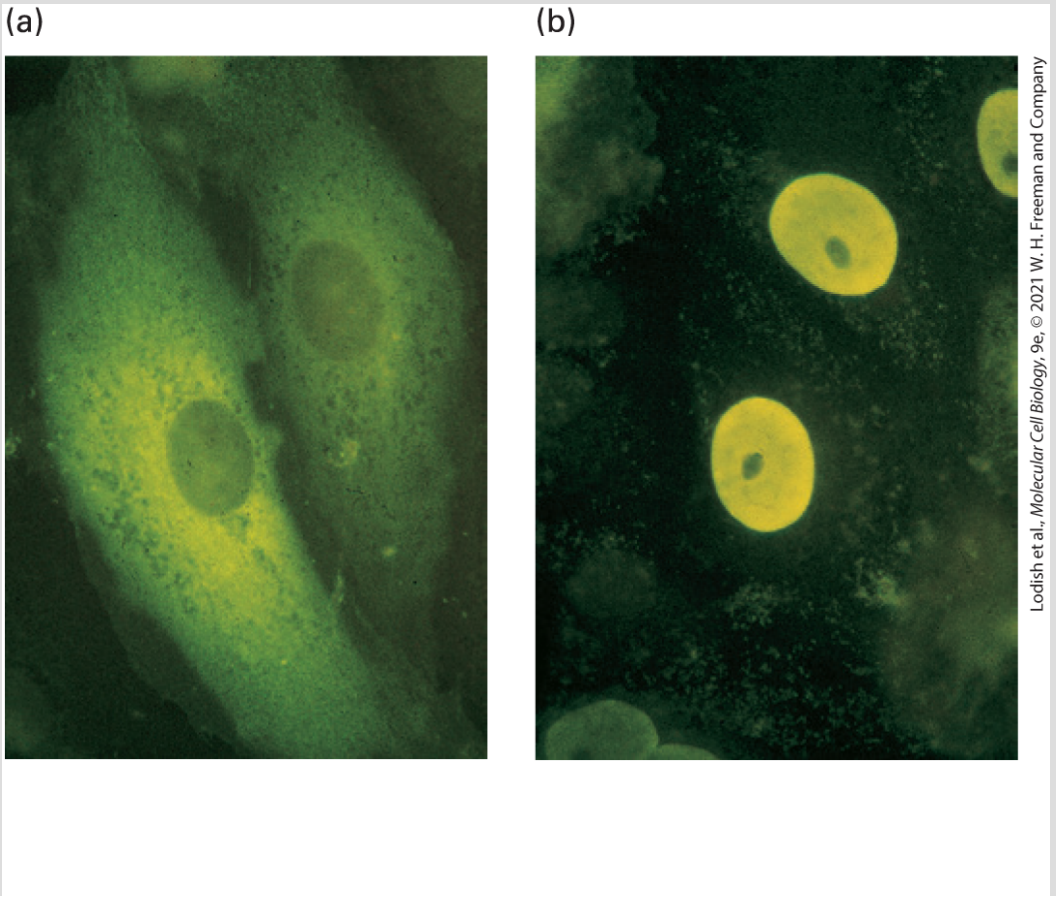
The **FG-nucleoporins**, which line the channel of the nuclear pore complex and are also found associated with the nuclear basket and the cytoplasmic filaments, contain multiple repeats of short hydrophobic sequences that are rich in phenylalanine (F) and glycine (G) residues (FG-repeats). The hydrophobic FG-repeats are thought to occur in regions of extended, otherwise hydrophilic polypeptide chains that fill the central transporter channel. The FG-nucleoporins are essential for the function of the NPC, but their

function appears to depend on a general property of these proteins rather than on a specific structure, since the NPC remains functional even if up to half of the FG-repeats have been deleted. Biophysical studies and dynamic computer modeling show that FG-repeats can have a disordered structure similar to that of an unfolded protein and are highly dynamic. Because of these properties, the central channel of the NPC can be thought of as being filled with a fluidlike phase that allows the diffusion of small molecules while excluding unchaperoned hydrophilic proteins larger than 40 kDa (Figure 13-32d).

Nuclear Transport Receptors Escort Proteins Containing Nuclear-Localization Signals into the Nucleus

All proteins found in the nucleus — such as histones, transcription factors, and DNA and RNA polymerases — are synthesized in the cytoplasm and imported into the nucleus through nuclear pore complexes. Such proteins contain a nuclear-localization signal (NLS) that directs their selective transport into the nucleus. NLSs were first discovered through the analysis of mutations of the gene for large T-antigen encoded by simian virus 40 (SV40). The wild-type form of large T-antigen is localized to the nucleus in virus-infected cells, whereas some mutated forms of large T-antigen accumulate in the cytoplasm. The mutations responsible for this altered cellular

localization all occur within a specific seven-residue sequence rich in basic amino acids near the C-terminus of the protein: Pro-Lys-Lys-Lys-Arg-Lys-Val. Experiments with chimeric proteins in which this sequence was fused to a cytosolic protein demonstrated that it directs transport into the nucleus and consequently functions as an NLS (Figure 13-33). NLS sequences were subsequently identified in numerous other proteins imported into the nucleus. Many of these sequences are similar to the basic NLS in SV40 large T-antigen, whereas others are chemically quite different. For instance, an NLS in the RNA-binding protein hnRNP A1 is hydrophobic. Accordingly, there are multiple mechanisms for the recognition of these very different sequences.



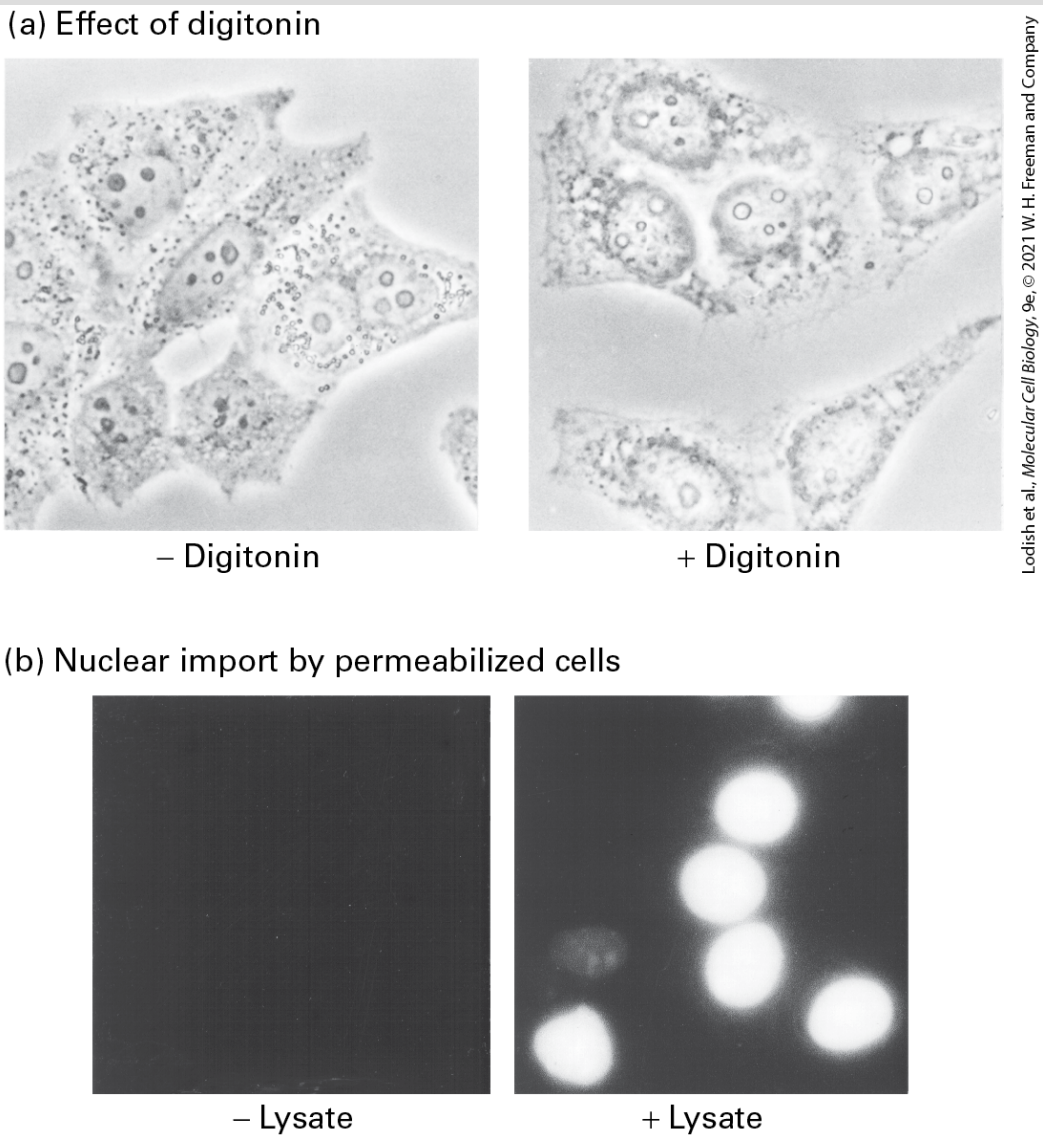
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 13-33 Nuclear-localization signals (NLSs) direct proteins to the cell nucleus. Cytoplasmic proteins can be transported to the nucleus if they are fused to a nuclear-localization signal. (a) Normal pyruvate kinase, here visualized by immunofluorescence after cultured cells were treated with a specific antibody (yellow), is localized to the cytoplasm. This very large cytosolic protein functions in carbohydrate metabolism. (b) When a chimeric pyruvate kinase containing the SV40 NLS at its N-terminus was expressed in cells, it was localized to the nucleus. The chimeric protein was expressed from a transfected engineered gene produced by fusing a viral gene fragment encoding the SV40 NLS to the pyruvate kinase gene.

[Republished with permission from Elsevier, from D. Kalderon et al., 1984, "A Short Amino Acid Sequence Able to Specify Nuclear Location," *Cell* **39**(3 Pt 2):499–509; permission conveyed through Copyright Clearance Center, Inc.]

Early work on the mechanism of nuclear import showed that proteins containing a positively charged NLS, similar to the one in SV40 large T-antigen, were efficiently transported into isolated nuclei only if they were provided with a cytosolic extract (Figure 13-34). Using this assay system, researchers purified two required cytosolic components, a nuclear transport receptor and Ran GTPase. **Nuclear transport receptors** bind to the NLS on a cargo protein to be transported into the nucleus; they also have affinity for FG-repeats on nucleoporins. Because of their affinity for FG-repeats, nuclear transport receptors and NLS-containing cargo proteins to which they are bound readily partition into the fluid phase of FG-repeats in the central channel of the NPC, whereas proteins of similar size that lack this property are excluded from the central channel. Nuclear transport receptors can be monomeric, consisting of a single polypeptide that can bind both to an NLS and to FG-

repeats, or they can be dimeric, with one subunit binding to the NLS and the other binding to FG-repeats.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 13-34 Cytosolic proteins are required for nuclear transport. The failure of nuclear transport to occur in permeabilized cultured cells in the absence of cell lysate demonstrates the involvement of soluble cytosolic components in the process. (a) Phase-contrast micrographs of untreated and digitonin-permeabilized HeLa cells. Treatment of a monolayer of cultured cells with the mild non-ionic detergent digitonin permeabilizes the plasma membrane so that cytosolic constituents leak out, but leaves the

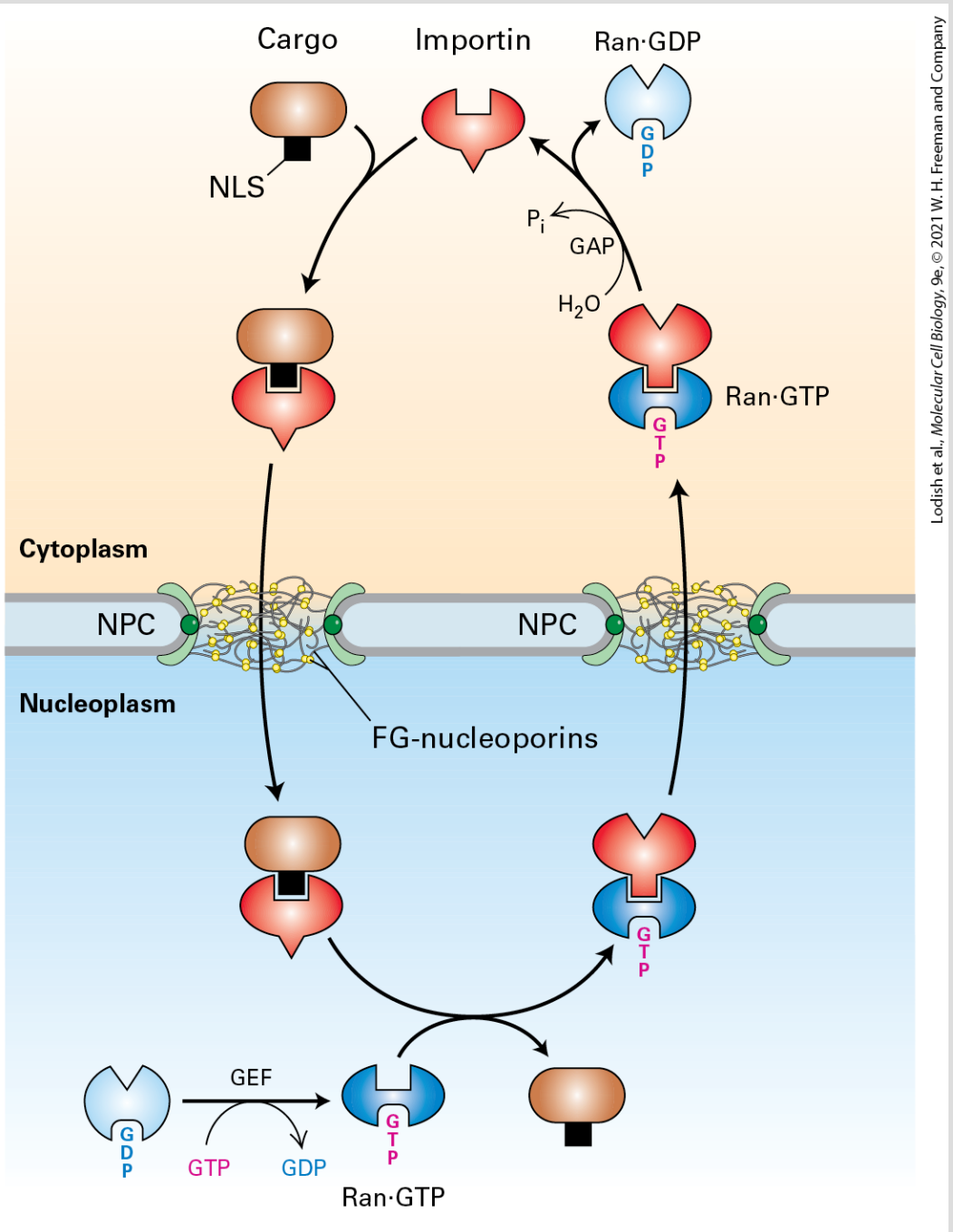
nuclear envelope and NPCs intact. (b) Fluorescence micrographs of digitonin-permeabilized HeLa cells incubated with a fluorescent protein chemically coupled to a synthetic SV40 T-antigen NLS peptide in the presence and absence of cell lysate (cytosol). Accumulation of the transport substrate in the nucleus occurred only when cytosol was present (*right*).

[S. A. Adam, R. S. Marr, and L. Gerace, 1990, *J. Cell Biol.* **111**: 807–816; https://doi.org/10.1083/jcb.111.3.807.]

Ran is a small monomeric G protein that exists in either a GTP-bound or a GDP-bound conformation (see Figure 3-35). It is the cycling of Ran between GTP-bound and GDP-bound conformations, leading to the net hydrolysis of GTP to GDP, that ultimately provides the energy to drive unidirectional transport of macromolecules through the nuclear pore.

The mechanism for the import of cytoplasmic cargo proteins mediated by a nuclear transport receptor known as *importin* is shown in Figure 13-35. Free importin in the cytoplasm binds to its cognate NLS in a cargo protein, forming an importin-cargo complex. The cargo complex then efficiently partitions into the NPC channel as the importin interacts with FG-repeats. Once in the channel, the importin-cargo complex can rapidly reach the nucleoplasmic side of the channel by random diffusion, and there the importin interacts with Ran · GTP, which causes a conformational change in the importin that displaces the NLS, releasing the cargo protein into the nucleoplasm. The importin-Ran · GTP complex then diffuses back through the NPC. Once the importin-Ran · GTP complex reaches

the cytoplasmic side of the NPC, Ran interacts with a specific *GTPase-activating protein (Ran-GAP)* that is a component of the NPC cytoplasmic filaments. This interaction stimulates Ran to hydrolyze its bound GTP to GDP, which causes it to convert to a conformation that has low affinity for the importin, so that the importin is released into the cytoplasm, where it can participate in another cycle of import. Ran · GTP travels back through the pore to the nucleoplasm, where it encounters a specific *guanine nucleotide exchange factor (Ran-GEF)* that causes Ran to release its bound GDP in favor of GTP. The net result of this series of reactions is the coupling of the hydrolysis of GTP to the transfer of an NLS-bearing protein from the cytoplasm to the nuclear interior, thus providing the energy to drive nuclear import.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-35 Mechanism for nuclear import of proteins. In the cytoplasm (*top*), free importin binds to the NLS of a cargo protein, and the importin-cargo complex thus formed diffuses through the NPC by transiently interacting with FG-nucleoporins. In the nucleoplasm, Ran·GTP binds to the importin, causing a conformational change that decreases the importin's affinity for the NLS and releasing the cargo. To support another cycle of import, the importin-Ran·GTP complex is transported back to the cytoplasm. A GTPase-activating protein

(GAP) associated with the cytoplasmic filaments of the NPC stimulates Ran to hydrolyze the bound GTP. This generates a conformational change that causes Ran to dissociate from the nuclear transport receptor, which can then initiate another round of import. Ran·GDP is returned to the nucleoplasm, where a guanine nucleotide exchange factor (GEF) causes release of GDP and rebinding of GTP.

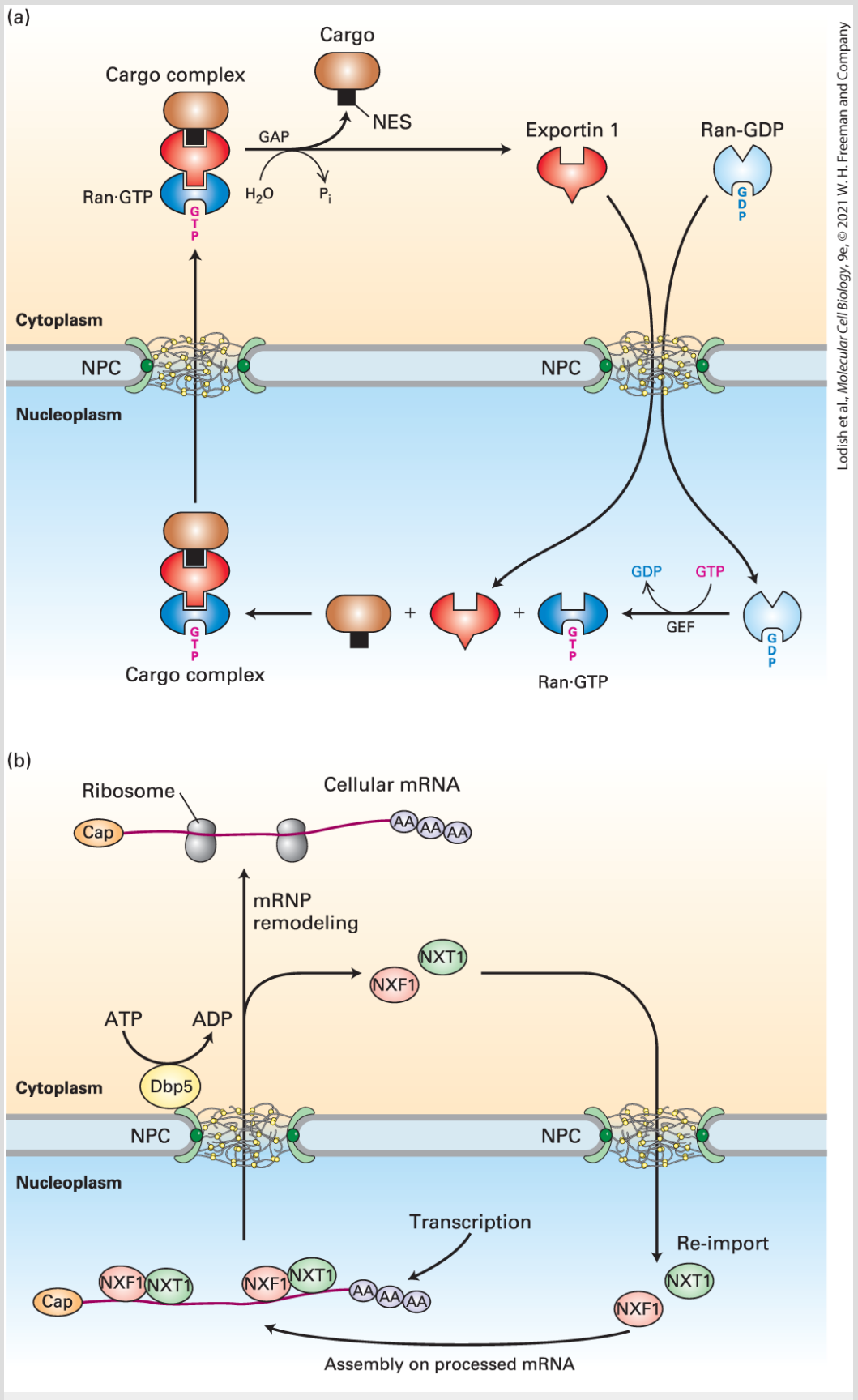
Although the importin-cargo complex travels through the pore by random diffusion, the overall process of transport of cargo into the nucleus is unidirectional. Because of the rapid dissociation of the complex when it reaches the nucleoplasm, there is a concentration gradient of importin-cargo complex across the NPC: high in the cytoplasm, where the complex assembles, and low in the nucleoplasm, where it dissociates. This concentration gradient is responsible for the unidirectional nature of nuclear import. A similar concentration gradient is responsible for driving importin from the nucleus back into the cytoplasm. The concentration of the importin-Ran·GTP complex is higher in the nucleoplasm, where it assembles, than on the cytoplasmic side of the NPC, where it dissociates. Ultimately, the direction of the transport processes depends on the localization of the Ran-GEF predominantly in the nucleoplasm and the Ran-GAP predominantly in the cytoplasm. Ran-GEF in the nucleoplasm maintains Ran in the Ran·GTP state, where it promotes dissociation of the cargo complex. Ran-GAP on the cytoplasmic side of the NPC converts Ran·GTP to Ran·GDP, dissociating the importin-Ran·GTP complex and releasing free importin into the cytosol.

A Second Type of Nuclear Transport Receptor Escorts Proteins Containing Nuclear-Export Signals out of the Nucleus

A mechanism very similar to the one we have just described is used to export proteins, tRNAs, and ribosomal subunits from the nucleus to the cytoplasm. This mechanism was initially elucidated by studies of certain ribonuclear protein complexes that shuttle between the nucleus and the cytoplasm. Such shuttling proteins contain a *nuclear-export signal (NES)* that stimulates their export from the nucleus to the cytoplasm through nuclear pores, in addition to an NLS that results in their uptake into the nucleus. Experiments with engineered hybrid genes encoding a nucleus-restricted protein fused to various segments of a shuttling protein have identified at least three different types of NESs: a leucine-rich sequence found in PKI (an inhibitor of protein kinase A) and in the Rev protein of human immunodeficiency virus (HIV), and two other sequences identified in two different heterogeneous ribonucleoprotein particles (hnRNPs). The precise structural features that determine the recognition of each type of sequence for nuclear export remain poorly understood.

The mechanism whereby shuttling proteins are exported from the nucleus is best understood for those containing a leucine-rich NES.

According to the current model, shown in Figure 13-36a, a specific nuclear transport receptor, called *exportin 1*, first forms a complex with Ran · GTP in the nucleus and then binds the NES in a cargo protein. Binding of exportin 1 to Ran · GTP causes a conformational change in exportin 1 that increases its affinity for the NES, so that a trimolecular cargo complex is formed. Like other nuclear transport receptors, exportin 1 interacts transiently with FG-repeats in FG-nucleoporins and diffuses through the NPC. The cargo complex dissociates when it encounters the Ran-GAP associated with the NPC cytoplasmic filaments, which stimulates Ran to hydrolyze the bound GTP, shifting it into a conformation that has low affinity for exportin 1. After Ran · GDP dissociates from the trimolecular cargo complex, exportin 1 changes its conformation to one that has low affinity for the NES, releasing the cargo into the cytosol. The direction of the export process is driven by this dissociation of the cargo from exportin 1 in the cytoplasm, which causes a concentration gradient of the cargo complex across the NPC that is high in the nucleoplasm and low in the cytoplasm. Exportin 1 and Ran · GDP are then transported back into the nucleus through the NPC.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 13-36 Ran-dependent and Ran-independent nuclear export. (a) Ran-dependent mechanism for nuclear export of cargo proteins containing a leucine-rich nuclear-export signal (NES). In the nucleoplasm (*bottom*), the protein exportin 1 binds cooperatively to the NES of the cargo protein to be transported and to $\mathrm{Ran} \cdot \mathrm{GTP}$. After the resulting cargo complex diffuses through an NPC via transient interactions with FG-repeats in FG-nucleoporins, the GAP associated with the NPC cytoplasmic filaments stimulates GTP hydrolysis, converting $\mathrm{Ran} \cdot \mathrm{GTP}$ to $\mathrm{Ran} \cdot \mathrm{GDP}$. The accompanying conformational change in Ran leads to dissociation of the complex. The NES-containing cargo protein is released into the cytosol, whereas exportin 1 and $\mathrm{Ran} \cdot \mathrm{GDP}$ are transported back into the nucleus through an NPC. Ran-GEF in the nucleoplasm then stimulates conversion of $\mathrm{Ran} \cdot \mathrm{GDP}$ to $\mathrm{Ran} \cdot \mathrm{GTP}$. (b) Ran-independent nuclear export of mRNAs. The heterodimeric $\mathrm{NXF1/NXT1}$ complex binds to mRNA-protein complexes (mRNPs) in the nucleus. $\mathrm{NXF1/NXT1}$ acts as a nuclear transport receptor and directs the associated mRNP to the central channel of the NPC by transiently interacting with FG-nucleoporins. An RNA helicase (Dbp5) located on the cytoplasmic side of the NPC removes NXF1 and NXT1 from the mRNA in a reaction that is powered by ATP hydrolysis. Free NXF1 and NXT1 proteins are recycled back into the nucleus by the Ran-dependent import process depicted in Figure 13-35.

By comparing this model for nuclear export with that in Figure 13-35 for nuclear import, we can see one obvious difference: $\mathrm{Ran} \cdot \mathrm{GTP}$ is part of the cargo complex during export, but not during import. Apart from this difference, the two transport processes are remarkably similar. In both processes, association of a nuclear transport receptor with $\mathrm{Ran} \cdot \mathrm{GTP}$ in the nucleoplasm causes a conformational change that affects its affinity for the transport signal. During import, the interaction causes release of the cargo, whereas during export, the interaction promotes association with the cargo. In both export and import, stimulation of $\mathrm{Ran} \cdot \mathrm{GTP}$

hydrolysis in the cytoplasm by Ran-GAP produces a conformational change in Ran that releases the nuclear transport receptor. During nuclear export, the cargo is also released. Localization of Ran-GAP and Ran-GEF to the cytoplasm and nucleus, respectively, is the basis for the unidirectional import and export of cargo proteins across the NPC.

In keeping with their similarity in function, the two types of nuclear transport receptors — importins and exportins — are highly homologous in sequence and structure. The family of nuclear transport receptors has 14 members in yeast and more than 20 in mammalian cells. The NESs or NLSs to which they bind have been determined for only a fraction of them. Some individual nuclear transport receptors function in both import and export.

A similar shuttling mechanism has been shown to export other cargoes from the nucleus. For example, exportin-t functions to export tRNAs. Exportin-t binds fully processed tRNAs in a complex with Ran · GTP that diffuses through NPCs and dissociates when it interacts with Ran-GAP in the NPC cytoplasmic filaments, releasing the tRNA into the cytosol. A Ran-dependent process is also required for the nuclear export of ribosomal subunits through NPCs once the protein and RNA components have been properly assembled in the nucleolus. Likewise, certain specific mRNAs that associate with particular hnRNP proteins can be exported by a Ran-dependent mechanism.

Most mRNAs Are Exported from the Nucleus by a Ran-Independent Mechanism

Once the processing of an mRNA is completed in the nucleus, it remains associated with specific proteins in a *messenger ribonuclear protein complex*, or *mRNP*. The principal transporter of mRNPs out of the nucleus is the **mRNP exporter**, a heterodimeric protein composed of a large subunit called *nuclear export factor* 1 *(NXF1)* and a small subunit called *nuclear export transporter* 1 *(NXT1)*. Multiple NXF1/NXT1 dimers bind to nuclear mRNPs through cooperative interactions with the RNA and other mRNP adapter proteins that associate with nascent pre-mRNAs during transcriptional elongation and pre-mRNA processing. This process adds a degree of quality control to mRNA export since NXF1/NXT1 dimers do not assemble onto the mRNP until it has all of the attributes of a properly and completely processed mRNA. Once incorporated into an mRNP, the subunits of NXF1/NXT1 in many respects act like a nuclear transport receptor that binds to an NLS or NES, since both NXF1 and NXT1 interact with the FG-repeats of FG-nucleoporins, and this interaction allows them to diffuse through the central channel of the NPC.

The process of mRNP export does not require Ran, and thus the unidirectional transport of mRNA out of the nucleus requires a

source of energy other than GTP hydrolysis by Ran. Once the mRNP-NXF1/NXT1 complex reaches the cytoplasmic side of the NPC, NXF1, and NXT1 dissociate from the mRNP with the help of the RNA helicase Dbp5, which is associated with cytoplasmic NPC filaments. Recall that RNA helicases use the energy derived from hydrolysis of ATP to move along RNA molecules, separating double-stranded RNA chains and dissociating RNA-protein complexes (see Chapter 5). This leads to the simple idea that Dpb5, which is associated with the cytoplasmic side of the nuclear pore complex, acts as an ATP-driven motor to remove NXF1/NXT1 from the mRNP complexes as they emerge on the cytoplasmic side of the NPC. The portions of an mRNP that have emerged into the cytosol and have been stripped of NXF1/NXT1 lack the ability to interact with FG-repeat nucleoporins. Thus an extended mRNP that transits the NPC cannot slide back into the nucleus, and the process of energy dependent removal of NXF1/NXT1 by Dbp5 helicase causes the mRNP to be progressively ratcheted into the cytoplasm. After being removed from the mRNP, the free NXF1 and NXT1 subunits are imported back into the nucleus by a process that depends on Ran and a nuclear transport receptor (Figure 13-36b).

KEY CONCEPTS OF SECTION 13.6

Transport into and out of the Nucleus

- The nuclear envelope contains numerous nuclear pore complexes (NPCs), which are large, complex structures composed of multiple copies of 30 different proteins called *nucleoporins* (see Figure 13-32).

- FG-nucleoporins, which contain multiple repeats of a short hydrophobic sequence (FG-repeats), fill the central transporter channel with a gel-like matrix that allows small molecules to pass but excludes macromolecules larger than 40 kDa.
- Active transport plays a role in the transport of all large proteins and RNAs through nuclear pores. This process requires the assistance of nuclear transport receptors that interact with both the transported molecule and FG-repeats of FG-nucleoporins.
- Proteins imported into or exported from the nucleus contain a specific amino acid sequence that functions as a nuclear-localization signal (NLS) or a nuclear-export signal (NES). Nucleus-restricted proteins contain an NLS but not an NES, whereas proteins that shuttle between the nucleus and cytoplasm contain both signals.
- Several different types of NESs and NLSs have been identified. Each type of nuclear-transport signal is thought to interact with a specific nuclear transport receptor belonging to a family of homologous proteins.
- A cargo protein bearing an NES or NLS translocates through nuclear pores bound to its cognate nuclear transport receptor. The transient interactions between nuclear transport receptors and FG-repeats allow very rapid diffusion of nuclear transport receptor–cargo complexes through the central channel of the NPC, which is filled with a hydrophobic matrix of FG-repeats.
- The unidirectional nature of protein export and import through nuclear pores results from the participation of Ran, a monomeric G protein that exists in different conformations when bound to GTP or GDP. Localization of the Ran guanine nucleotide exchange factor (Ran-GEF) in the nucleus and of the Ran GTPase-activating protein (Ran-GAP) in the cytoplasm creates a gradient with high concentrations of $\mathrm{Ran} \cdot \mathrm{GTP}$ in the nucleoplasm and of $\mathrm{Ran} \cdot \mathrm{GDP}$ in the cytoplasm. The interaction of a cargo complex with $\mathrm{Ran} \cdot \mathrm{GTP}$ in the nucleoplasm causes dissociation of the complex, releasing the cargo into the nucleoplasm (see Figure 13-35), whereas the assembly of an export cargo complex is stimulated by interaction with $\mathrm{Ran} \cdot \mathrm{GTP}$ in the nucleoplasm (see Figure 13-36).
- Most mRNPs are exported from the nucleus by binding to a heterodimeric mRNP exporter in the nucleoplasm that interacts with FG-repeats. The direction of transport (nucleus to cytoplasm) results from the action of an RNA helicase associated with the cytoplasmic filaments of the NPC that removes the heterodimeric mRNP exporter once the transport complex has reached the cytoplasm.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the concepts in this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional tools including videos, animations, and quizzes

Key Terms

AAA ATPase family
cotranslational translocation
dislocation
disulfide bonds (–S–S–)
dolichol phosphate
ERAD (ER-associated protein degradation)
FG-nucleoporin
general import pore
glycoprotein
hydropathy profile
microsome
mRNP exporter

multipass membrane protein
N-linked oligosaccharide
nuclear pore complex (NPC)
nuclear transport receptor
nucleoporins
O-linked oligosaccharide
post-translational translocation
protein disulfide isomerase (PDI)
Ran
rough ER
secretory pathway
signal-anchor sequence (SA)
signal-based targeting
signal recognition particle (SRP)
single-pass membrane protein
SRP receptor
stop-transfer anchor sequence (STA)
targeting sequence
thylakoids
topogenic sequence
topology
translocon
unfolded-protein response
vesicle-based trafficking

Review the Concepts

1. The following results were obtained in early studies on the translation of secretory proteins. Based on what we now know of this process, explain the reason why each result was observed.
   a. An in vitro translation system consisting only of mRNA and ribosomes resulted in secretory proteins that were larger than the identical protein when translated in a cell.
   b. A similar system that also included microsomes produced secretory proteins that were identical in size to those found in a cell.
   c. When the microsomes were added after in vitro translation, the synthesized proteins were again larger than those made in a cell.
2. Describe the source or sources of energy needed for unidirectional translocation across the membrane in (a) cotranslational translocation into the endoplasmic reticulum (ER); (b) post-translational translocation into the ER; and (c) translocation into the mitochondrial matrix.
3. Translocation into most organelles usually requires the activity of one or more cytosolic proteins. Describe the basic functions of three different cytosolic factors required for translocation into the ER, mitochondria, and peroxisomes, respectively.
4. Describe the typical principles used to identify topogenic sequences within proteins and how these principles can be used to develop computer algorithms. How does the identification of topogenic sequences lead to prediction of

the membrane arrangement of a multipass protein? What is the importance of the arrangement of positive charges relative to the membrane orientation of a signal-anchor sequence?

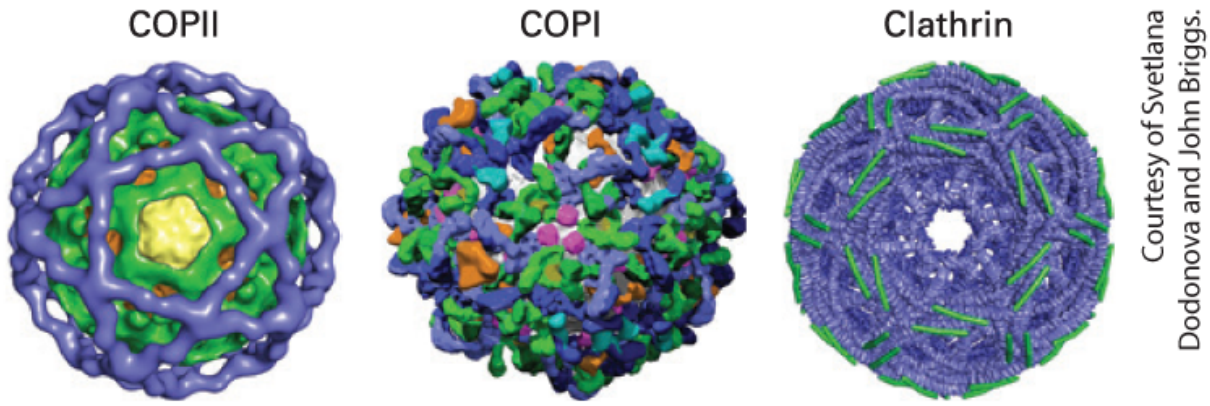
5. An abundance of misfolded proteins in the ER can result in the activation of the unfolded-protein response (UPR) and ER-associated degradation (ERAD) pathways. UPR decreases the abundance of unfolded proteins by altering gene expression of what type of genes? What is one manner in which ERAD may identify misfolded proteins? Why is dislocation of these misfolded proteins to the cytoplasm necessary?
6. Temperature-sensitive yeast mutants have been isolated that block each of the enzymatic steps in the synthesis of the dolichol-linked oligosaccharide precursor for *N*-linked glycosylation. Propose an explanation for why mutations that block synthesis of the intermediate with the structure dolichol-PP-$(\mathrm{GlcNAc})_2\mathrm{Man}_5$ completely prevent addition of *N*-linked oligosaccharide chains to secretory proteins, whereas mutations that block conversion of this intermediate into the completed precursor — dolichol-PP-$(\mathrm{GlcNAc})_2\mathrm{Man}_9\mathrm{Glc}_3$ — allow the addition of *N*-linked oligosaccharide chains to secretory glycoproteins.
7. Name four different proteins that facilitate the modification or folding of secretory proteins within the lumen of the ER. Indicate which of these proteins covalently modifies substrate proteins and which brings about only conformational changes in substrate proteins.

8. Describe what would happen to the precursor of a mitochondrial matrix protein in the following types of mitochondrial mutants: (a) a mutation in the Tom22 signal receptor; (b) a mutation in the Tom70 signal receptor; (c) a mutation in the matrix Hsp70; and (d) a mutation in the matrix signal peptidase.
9. Describe the similarities and differences between the mechanism of import into the mitochondrial matrix and the chloroplast stroma.
10. Design a set of experiments using chimeric proteins, composed of a mitochondrial precursor protein fused to dihydrofolate reductase (DHFR), that could be used to determine how much of the precursor protein must protrude into the mitochondrial matrix in order for the matrix-targeting sequence to be cleaved by the matrix-processing protease.
11. Peroxisomes contain enzymes that use molecular oxygen to oxidize various substrates, but in the process, hydrogen peroxide — a compound that can damage DNA and proteins — is formed. What is the name of the enzyme responsible for the breakdown of hydrogen peroxide to water? What is the mechanism of the import of this protein into the peroxisome, and what other proteins are involved?
12. Suppose that you have identified a new mutant cell line that lacks functional peroxisomes. Describe how you could determine experimentally whether the mutant is primarily defective for insertion/assembly of peroxisomal membrane proteins or matrix proteins.

13. The nuclear import of proteins larger than 40 kDa requires the presence of what amino acid sequence? Describe the mechanism of nuclear import. How are nuclear transport receptors able to get through the nuclear pore complex?
14. Why is localization of Ran-GAP in the nucleus and Ran-GEF in the cytoplasm necessary for unidirectional transport of cargo proteins containing an NES?

CHAPTER 14

Vesicular Traffic, Secretion, and Endocytosis



Structural models of the coat protein complexes for three types of transport vesicles. The coat complexes are responsible for sculpting the curvature of the vesicle membrane and for selecting the types of cargo proteins incorporated into the vesicle.

[Data from A. J. Noble and S. M. Stagg, 2015, *Science* **349**(6244):142–143; https://doi.org/10.1126/science.aac6537.]

OUTLINE

14.1 Techniques for Studying the Secretory Pathway
14.2 Molecular Mechanisms of Vesicle Budding and Fusion
14.3 Early Stages of the Secretory Pathway
14.4 Later Stages of the Secretory Pathway
14.5 Receptor-Mediated Endocytosis
14.6 Directing Membrane Proteins and Cytosolic Materials to the Lysosome for Degradation

In the previous chapter, we explored how proteins are targeted to and translocated across the membranes of different intracellular organelles, including the endoplasmic reticulum (ER), mitochondria, chloroplasts, peroxisomes, and the nucleus. In this chapter, we turn our attention to the **secretory pathway** and the mechanisms of vesicular traffic that allow proteins to be secreted from the cell or delivered to the plasma membrane and the lysosomes. We also discuss the related processes of endocytosis, which delivers proteins and small molecules from outside the cell, or autophagy, which delivers organelles, proteins, and small molecules from the interior of the cell to the lysosome for degradation.

The secretory pathway is so named because it was initially studied in dedicated secretory cells that produce and secrete large quantities of proteins such as insulin or digestive enzymes to the outside of the cell. It was later discovered that the same pathway used for extracellular secretion of proteins delivers almost all newly synthesized membrane lipids and membrane proteins from the endoplasmic reticulum (ER) to the cell surface. Newly synthesized membrane proteins delivered to the plasma membrane include cell-surface receptors, transporters for nutrient uptake, and ion channels that maintain the proper ionic and electrochemical balance across the plasma membrane. Soluble secreted proteins also follow the secretory pathway to the cell surface, but instead of remaining embedded in the membrane, they are released into the aqueous extracellular environment. Examples of secreted proteins

include digestive enzymes, peptide hormones, serum proteins, and collagen.

The secretory pathway is also responsible for the distribution of particular sets of soluble and membrane proteins to the organelles that lie along the secretory pathway including the ER itself, the Golgi, the endosome, and the lysosome. As we saw in Chapter 13, the ER is the site of synthesis and folding of all proteins that enter the secretory pathway. In this chapter, we describe the principal function of the Golgi — remodeling of carbohydrate modifications of secretory proteins. The Golgi along with the endosome are two major hubs for distributing proteins to their final destination at the plasma membrane or the lysosome. The lysosome, as described in Chapter 4, is the organelle with an acidic interior that is used for degradation of proteins and storage of small molecules such as amino acids. Accordingly, the types of proteins delivered to the lysosomal membrane include subunits of the V-class proton pump that pumps H^+ from the cytosol into the acidic lumen of the lysosome, as well as transporters that release small molecules stored in the lysosome into the cytoplasm. Soluble proteins delivered by this pathway include lysosomal digestive enzymes such as proteases, glycosidases, phosphatases, and lipases.

For all of its apparent complexity, the organization and operation of the secretory pathway can be understood as the consequence of repeated variations of just two basic mechanistic themes. The first

theme is that transport of membrane and soluble proteins from one membrane-bounded compartment to another is mediated by coated **transport vesicles** that collect *cargo proteins* in buds arising from the membrane of one compartment and then deliver these cargo proteins to the next compartment by fusing with the membrane of that compartment. As we will see, different types of vesicles are responsible for each step in the secretory pathway, but the underlying mechanisms of budding of a coated vesicle from a donor membrane and the specific fusion with the appropriate target membrane are remarkably conserved, regardless of vesicle type. Importantly, as transport vesicles bud from one membrane and fuse with the next, the same face of the membrane remains oriented toward the cytosol. Therefore, once a protein has been inserted into the membrane or the lumen of the ER, that protein can be carried along the secretory pathway, moving from one organelle to the next without being translocated across another membrane or altering its orientation within the membrane.

The second repeated theme is that each organelle along the secretory pathway both receives new proteins delivered by targeted transport vesicle fusion and donates proteins that are removed from the organelle by transport vesicle budding. Because transport vesicles package and deliver only specific subsets of proteins, the outcome of vesicular trafficking to and from each organelle establishes the composition and thus the identity of each organelle.

The composition of the plasma membrane is determined by two pathways: vesicle transport to the plasma membrane by the secretory pathway, which is counterbalanced by removal of proteins and membrane by the **endocytic pathway** (Figure 14-1). Both selective delivery by secretory vesicles and selective removal of proteins by endocytic vesicles can be regulated, allowing the protein composition of the plasma membrane to be regulated in response to environmental or developmental signals. In addition, the endocytic pathway is used to ingest certain nutrients that are too large to be transported across the plasma membrane by one of the transport mechanisms discussed in Chapter 11. For example, the endocytic pathway is used in the uptake of cholesterol carried in LDL particles. In addition, the endocytic pathway can be used to remove receptor proteins from the cell surface as a way to down-regulate their activity.

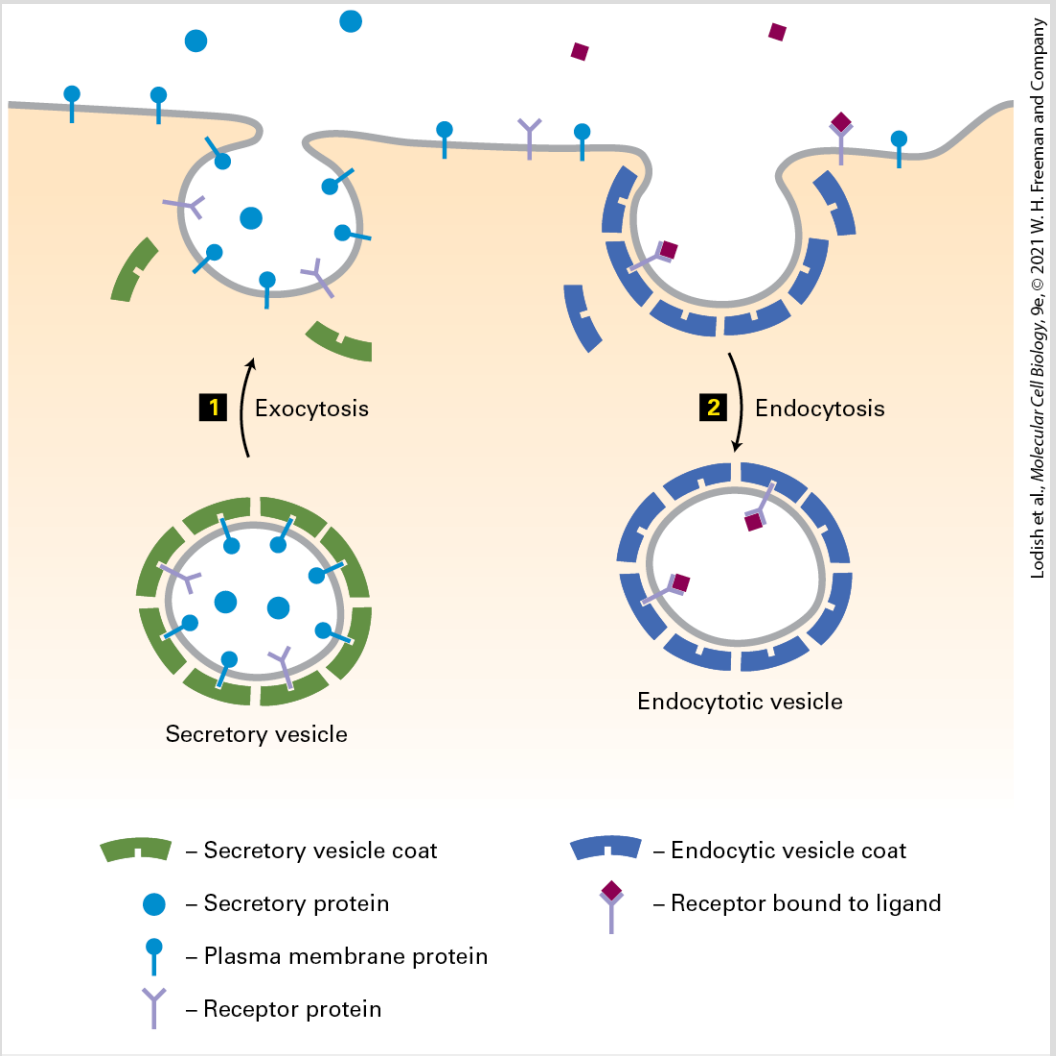
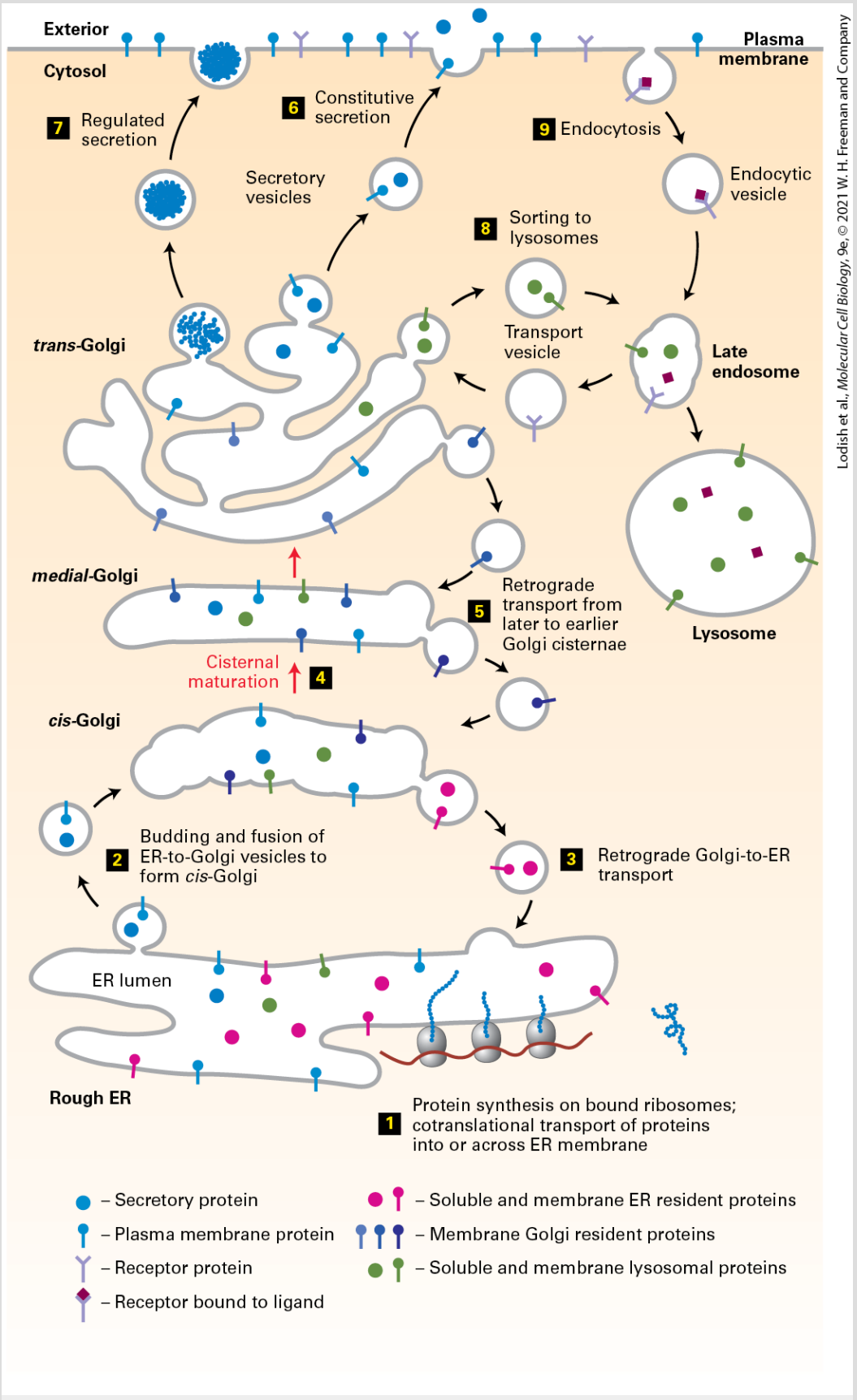


FIGURE 14-1 Basic principles of vesicle trafficking govern the composition of the plasma membrane. 1 : Membranes and proteins are delivered to the plasma membrane by exocytosis of coated vesicles that bud from the Golgi complex. The types of proteins incorporated into these vesicles are determined by interactions with vesicle coat proteins. Note that the membrane proteins preserve their orientation in the membrane such that the domain facing the vesicle interior will ultimately face the exoplasmic face of the plasma membrane. Correspondingly, soluble proteins within the vesicle interior will be released into the extracellular space. 2 : Proteins and membrane are removed from the plasma membrane by the process of endocytosis by which coated vesicles bud from the plasma membrane into the cytoplasm. In this example, a membrane receptor protein is only incorporated into endocytic vesicles when bound to ligand. The overall composition of the plasma membrane is

controlled by a balance between the proteins that are selectively delivered by exocytic vesicles and the proteins that are removed by endocytic vesicles.

The overall organization of the individual vesicle trafficking steps to form the intersecting secretory and endocytic pathways is outlined in Figure 14-2. The first stage of the secretory pathway takes place in the rough endoplasmic reticulum (ER) (Figure 14-2, step **1**). As described in Chapter 13, newly synthesized soluble and membrane proteins are translocated into the ER, where they fold into their proper conformation and receive covalent modifications such as *N*-linked and *O*-linked carbohydrates and disulfide bonds. Once correctly folded and modified in the ER lumen, the newly synthesized proteins progress to the second stage of the secretory pathway: transport to and through the Golgi complex.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-2 Overview of the secretory and endocytic pathways of protein sorting. *Secretory pathway:* Synthesis of proteins bearing an ER signal sequence is completed on the rough ER (step **1**), and the newly made polypeptide chains are inserted into the ER membrane or cross it into the ER lumen (see Chapter 13). ER resident proteins such as BiP and PDI remain within the ER. The remainder are packaged into transport vesicles (step **2**) that bud from the ER and fuse to form new *cis*-Golgi cisternae. Missorted ER-resident proteins and vesicle membrane proteins that need to be reused are retrieved to the ER by vesicles (step **3**) that bud from the *cis*-Golgi and fuse with the ER. Each *cis*-Golgi cisterna, with its protein content, physically moves from the *cis* to the *trans* face of the Golgi complex (step **4**) by a nonvesicular process called *cisternal maturation*. Retrograde transport vesicles (step **5**) move Golgi-resident proteins to the proper Golgi compartments. In all cells, certain soluble proteins move to the cell surface in transport vesicles (step **6**) and are secreted continuously (constitutive secretion). In certain cell types, some soluble proteins are stored in secretory vesicles (step **7**) and are released only after the cell receives an appropriate neuronal or hormonal signal (regulated secretion). Lysosome-destined membrane and soluble proteins, which are transported in vesicles that bud from the *trans*-Golgi (step **8**), first move to the late endosome and then to the lysosome. *Endocytic pathway:* Plasma membrane proteins such as receptor proteins bound to their ligands can be taken up in vesicles that bud from the plasma membrane (step **9**) and move to the endosome from where they can either be recycled to the plasma membrane or be delivered to the lysosome for degradation.

The transport of cargo proteins from the ER to the Golgi occurs via **anterograde** (forward-moving) transport vesicles (Figure 14-2, step **2**). These vesicles fuse with one another to form a flattened membrane-bounded compartment known as the *cis-Golgi network* or *cis*-Golgi **cisterna** (a "cistern" is a container for holding water or other liquid). Certain proteins, mainly proteins that function in the ER, can be retrieved from the *cis*-Golgi cisterna and returned to the

ER via a different set of **retrograde** (backward-moving) transport vesicles (step **3**). In a manner reminiscent of an assembly line, the new *cis*-Golgi cisterna, with its cargo of proteins, physically moves from the *cis* position (nearest the ER) to the *trans* position (farthest from the ER), successively becoming first a *medial*-Golgi cisterna and then a *trans*-Golgi cisterna (step **4**). This process, known as **cisternal maturation**, primarily involves retrograde transport vesicles (step **5**), which retrieve enzymes and other Golgi-resident proteins from later to earlier Golgi cisternae, thereby maturing the *cis*-Golgi cisternae to *medial*-Golgi cisternae, and *medial*-Golgi cisternae to *trans*-Golgi cisternae. As secretory proteins move through the Golgi, their linked carbohydrates may be further modified by specific glycosyltransferases that are housed in the different Golgi compartments.

Proteins in the Golgi are eventually delivered to a complex network of membranes and vesicles termed the ***trans*-Golgi network**. The *trans*-Golgi network is a major branch point in the secretory pathway. It is at this stage that proteins are loaded into different kinds of vesicles and thereby trafficked to different destinations. Depending on which kind of vesicle the protein is loaded into, it will be transported to the plasma membrane and secreted immediately, stored for later release, or shipped to the lysosome (steps **6**–**8**). The process by which a vesicle moves to and fuses with the plasma membrane and releases its contents is known as **exocytosis**. In all cell types, at least some proteins are secreted continuously (a

process commonly called **constitutive secretion**), while others are stored inside the cell in specialized vesicles known as *secretory granules* until a signal for exocytosis causes them to be released (**regulated secretion**). Secretory proteins destined for lysosomes are first transported by vesicles from the *trans*-Golgi network to a compartment usually called the **late endosome**; the proteins are then transferred to the lysosome by direct fusion of the late endosome with the lysosomal membrane.

In this chapter, we first discuss the experimental techniques that have contributed to our knowledge of the secretory pathway and endocytosis. Then we focus on the general mechanisms of membrane budding and fusion. We will see that although different kinds of transport vesicles use distinct sets of proteins for their formation and fusion, all vesicles use the same general mechanism for budding, selection of particular sets of cargo molecules, and fusion with the appropriate target membrane. In the remaining sections of the chapter, we discuss both the early and late stages of the secretory pathway, including how specificity of targeting to different destinations is achieved. We conclude with a discussion of how proteins are transported to the lysosome by the endocytic pathway.

14.1 Techniques for Studying the Secretory Pathway

The key to understanding how proteins are transported through the organelles of the secretory pathway has been to develop a basic description of the function of transport vesicles and the mechanism by which they incorporate particular cargo molecules. Many components required for the formation and fusion of transport vesicles have been identified by a remarkable convergence of the genetic and biochemical approaches described in this section. All studies of intracellular protein trafficking employ some method for assaying the transport of a given protein from one compartment to another. We begin by describing how intracellular protein transport can be followed in live cells and then consider genetic and in vitro systems that have proved useful in elucidating the secretory pathway.

Transport of a Protein Through the Secretory Pathway Can Be Assayed in Live Cells

The classic studies of George Palade and his colleagues in the 1960s first established the order in which proteins move from one organelle to the next in the secretory pathway (see Classic

Experiment 14-1). These early studies also showed that secretory proteins are not released into the cytosol — the first indication that transported proteins are always associated with some type of membrane-bounded intermediate. In these experiments, radioactively labeled amino acids were injected into the pancreas of hamsters and the newly synthesized proteins that incorporated the labeled amino acids could be tracked as they made their way to the cell surface. At different times after injection, the animals were sacrificed and the pancreatic cells were immediately fixed with glutaraldehyde, sectioned, and subjected to autoradiography to visualize the locations of the radiolabeled proteins. Because the radioactive amino acids were administered in a short pulse, only those proteins synthesized immediately after injection were labeled, forming a distinct cohort of labeled proteins whose transport could be followed. In addition, because pancreatic acinar cells are dedicated secretory cells, almost all of the labeled amino acids in these cells were incorporated into secretory proteins, facilitating the observation of transported proteins.

Although autoradiography is rarely used today to localize proteins within cells, these early experiments illustrate the two basic requirements for any assay of intercompartmental transport. First, it is necessary to label a cohort of proteins in an early compartment so that their subsequent transfer to later compartments can be followed over time. Second, it is necessary to have a way to identify the compartment in which a labeled protein resides. Here we describe two modern experimental procedures for observing the

intracellular trafficking of a secretory protein in almost any type of cell.

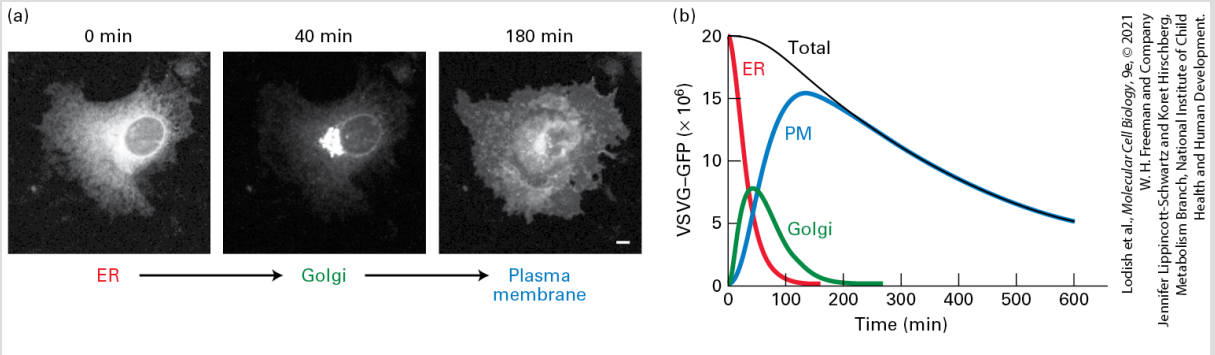
In both procedures, a gene encoding an abundant membrane glycoprotein (G protein) from vesicular stomatitis virus (VSV) is introduced into cultured mammalian cells either by transfection or simply by infecting the cells with the virus. The transfected cells efficiently produce the VSV G protein, which is then inserted into the membrane of the ER. The use of a mutant gene encoding a temperature-sensitive VSV G protein allows researchers to turn subsequent transport of this protein on and off. At the restrictive temperature of 40 °C, newly made VSV G protein is misfolded and therefore retained within the ER by the quality-control mechanisms discussed in Chapter 13, whereas at the permissive temperature of 32 °C, the protein is correctly folded and transported through the secretory pathway to the cell surface. Importantly, the misfolding of the temperature-sensitive VSV G protein is reversible; thus when cells synthesizing mutant VSV G protein are grown at 40 °C and then shifted to 32 °C, the misfolded mutant VSV G protein that had accumulated in the ER will refold and be transported normally. This clever use of a temperature-sensitive mutation in effect defines a protein cohort whose subsequent transport can be followed.

In two variations of this basic procedure, transport of VSV G protein is monitored by different techniques. Studies using both of these modern trafficking assays came to the same conclusion as Palade's early experiments: in mammalian cells, vesicle-mediated transport

of a protein molecule from its site of synthesis on the rough ER to its arrival at the plasma membrane takes from 30 to 60 minutes.

Microscopy of GFP-Labeled VSV G Protein

One approach for observing the transport of VSV G protein employs a hybrid gene in which the viral gene is fused to the gene encoding green fluorescent protein (GFP), a naturally occurring fluorescent protein (see Chapter 4). The hybrid gene is transfected into cultured cells by techniques described in Chapter 6. When cells expressing the temperature-sensitive form of the hybrid protein (VSVG-GFP) are grown at the restrictive temperature, VSVG-GFP accumulates in the ER, which appears as a lacy network of membranes when the cells are observed in a fluorescent microscope. When the cells are subsequently shifted to a permissive temperature, the VSVG-GFP can be seen to move first to the membranes of the Golgi complex, which are densely concentrated at the edge of the nucleus, and then to the cell surface (Figure 14-3a). By observing the distribution of VSVG-GFP at different times after shifting cells to the permissive temperature, researchers have determined how long VSVG-GFP resides in each organelle of the secretory pathway (Figure 14-3b).

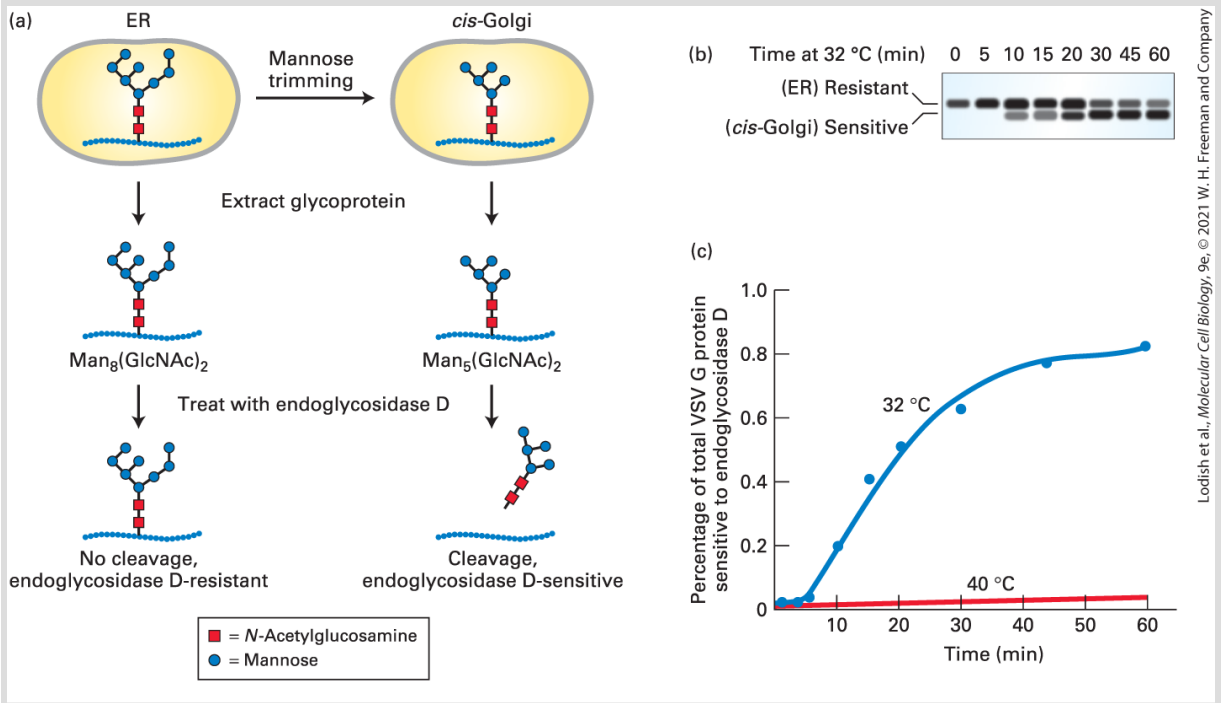


EXPERIMENTAL FIGURE 14-3 Protein transport through the secretory pathway can be visualized by fluorescence microscopy of cells producing a GFP-tagged membrane protein. Cultured cells were transfected with a hybrid gene encoding the viral membrane glycoprotein VSV G linked to the gene for green fluorescent protein (GFP). A temperature-sensitive mutant version of the viral gene was used so that newly made hybrid protein (VSVG-GFP) was retained in the ER at 40 °C but was released for transport at 32 °C. (a) Fluorescence micrographs of cells just before and at two times after they were shifted to the lower temperature. Movement of VSVG-GFP from the ER to the Golgi and finally to the cell surface occurred within 180 minutes. The scale bar is 5 μm. (b) Plot of the amount of VSVG-GFP in the endoplasmic reticulum (ER), Golgi, and plasma membrane (PM) at different times after the shift to the permissive temperature. The kinetics of transport from one organelle to another can be reconstructed from computer analysis of these data. The decrease in total fluorescence that occurs at later times probably results from slow inactivation of GFP fluorescence.

Detection of Compartment-Specific Oligosaccharide Modifications

A second way to follow the transport of secretory proteins takes advantage of modifications to their carbohydrate side chains that occur at different stages of the secretory pathway. To understand this approach, recall that many secretory proteins leaving the ER are carrying one or more copies of the *N*-linked oligosaccharide

$Man_8(GlcNAc)_2$, which are synthesized and attached to secretory proteins in the ER (see Figure 13-18). As a protein moves through the Golgi complex, different enzymes localized to the *cis-, medial-*, and *trans*-Golgi cisternae catalyze an ordered series of modifications to these core $Man_8(GlcNAc)_2$ chains, as discussed in a later section of this chapter. For instance, glycosidases that reside specifically in the *cis*-Golgi compartment sequentially trim mannose residues off the core oligosaccharide to yield a trimmed form, $Man_5(GlcNAc)_2$. By monitoring the trimming of the *N*-linked oligosaccharide, scientists can distinguish glycosylated VSV G protein that has entered the *cis*-Golgi from VSV G protein that remains in the ER. A specialized carbohydrate-cleaving enzyme known as *endoglycosidase D* enables the trimming of the *N*-linked oligosaccharide to be readily detected by gel electrophoresis; endoglycosidase D will cleave trimmed *cis*-Golgi-specific oligosaccharides from proteins, but will not cleave the core (untrimmed) oligosaccharide chains on secretory proteins within the ER (Figure 14-4a). Because a deglycosylated VSV G protein produced by endoglycosidase D digestion moves faster on an SDS gel than the corresponding glycosylated protein, these proteins can be readily distinguished (Figure 14-4b).



EXPERIMENTAL FIGURE 14-4 Transport of a membrane glycoprotein from the ER to the Golgi can be assayed based on sensitivity to cleavage by endoglycosidase D. Cells expressing a temperature-sensitive VSV G protein were labeled with a pulse of radioactive amino acids at the nonpermissive temperature so that the labeled protein was retained in the ER. At periodic times after a return to the permissive temperature of 32 °C, VSV G protein was extracted from cells and digested with endoglycosidase D. (a) As proteins move to the *cis*-Golgi from the ER, the core oligosaccharide $\mathrm{Man_8(GlcNAc)_2}$ is trimmed to $\mathrm{Man_5(GlcNAc)_2}$ by enzymes that reside in the *cis*-Golgi compartment. Endoglycosidase D cleaves the oligosaccharide chains from proteins processed in the *cis*-Golgi but not from proteins in the ER. (b) SDS-polyacrylamide gel electrophoresis of the digestion mixtures resolves the resistant, uncleaved (slower migrating) and sensitive, cleaved (faster migrating) forms of labeled VSV G protein. Initially, as this gel shows, all of the VSV G protein was resistant to digestion, but over time, an increasing fraction was sensitive to digestion, reflecting transport of the protein from the ER to the Golgi and its processing there. In control cells kept at 40 °C, only slow-moving, digestion-resistant VSV G protein was detected after 60 minutes (not shown). (c) A plot of the percentage of VSV G protein that is sensitive to digestion, derived from electrophoretic data, reveals the time course of ER-to-Golgi transport.

[Part (b) Data from C. J. Beckers et al., 1987, “Semi-Intact Cells Permeable to Macromolecules: Use in Reconstitution of Protein Transport from the Endoplasmic

Reticulum to the Golgi Complex," *Cell* **50**(4):523–34.]

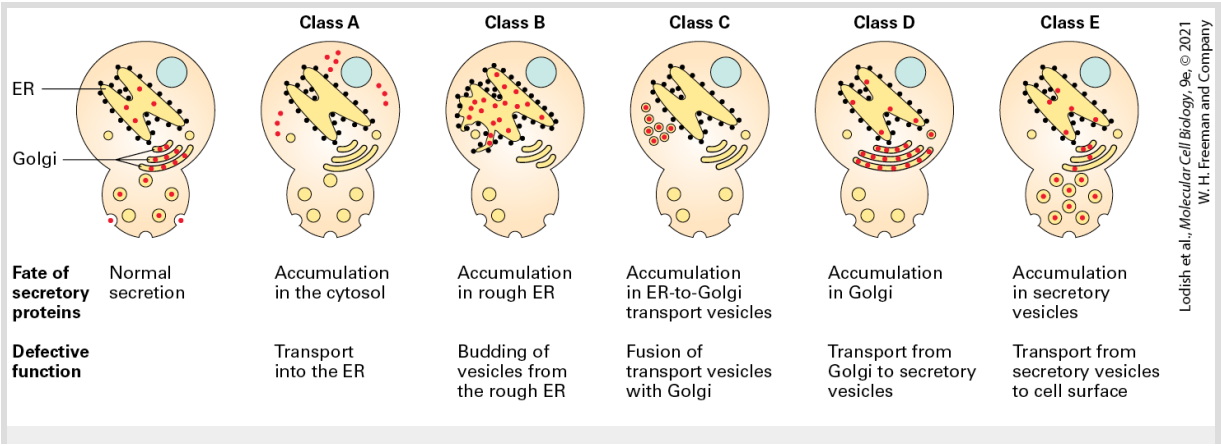
This type of assay can be used to track movement of VSV G protein from the ER to the *cis*-Golgi in virus-infected cells pulse-labeled with radioactive amino acids. Immediately after labeling, all the labeled VSV G protein is still in the ER and, upon extraction, is resistant to digestion by endoglycosidase D. Over time, however, the fraction of the extracted glycoprotein that has received carbohydrate trimming in the *cis*-Golgi and is thus sensitive to digestion by endoglycosidase D increases. Note that transport of VSV G protein from the ER to the Golgi takes about 30 minutes, as measured either by the assay based on oligosaccharide processing or by fluorescence microscopy of VSVG-GFP (Figure 14-4c). A variety of assays based on specific carbohydrate modifications that occur in later Golgi compartments have been developed to measure progression of VSV G protein through each compartment of the Golgi complex.

Yeast Mutants Define Major Stages and Components of Vesicular Transport

The general organization of the secretory pathway and many of the molecular components required for vesicle trafficking are similar in all eukaryotic cells. Because of this conservation, genetic studies with yeast have been useful in identifying many of the proteins that participate in vesicular traffic. For yeast cells, as for all cells, the

secretory pathway is essential for transport and delivery of new protein and membrane to the cell surface. Thus genes encoding important components of the secretory pathway are essential for cell growth and can be studied only as conditional temperature-sensitive mutants, as described in Chapter 8.

A large number of yeast mutants were initially identified by their inability to secrete proteins at a nonpermissive temperature. When these temperature-sensitive ***secretion (sec)* mutants** are transferred from the lower permissive temperature to the higher nonpermissive temperature, they accumulate secretory proteins at the point in the secretory pathway blocked by the mutation. Analysis of such mutants identified five classes (A–E) characterized by protein accumulation in the cytosol, rough ER, small vesicles taking proteins from the ER to the Golgi complex, Golgi cisternae, or constitutive secretory vesicles (Figure 14-5). Subsequent characterization of *sec* mutants in these various classes has helped elucidate the fundamental components and molecular mechanisms of vesicle trafficking that we discuss in later sections.



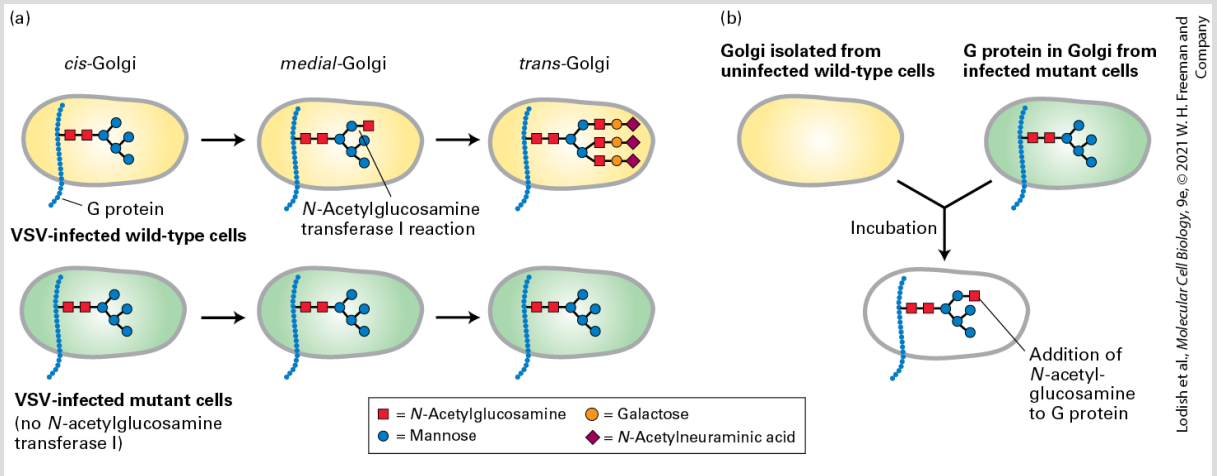
EXPERIMENTAL FIGURE 14-5 Phenotypes of yeast *sec* mutants identified five stages in the secretory pathway. These temperature-sensitive mutants can be grouped into five classes (A–E) based on the site where newly made secretory proteins (red dots) accumulate when cells are shifted from the permissive temperature to the higher, nonpermissive one. Analysis of double mutants permitted the sequential order of the steps to be determined. Note that organelle structures before the stage blocked by a mutant are exaggerated because the *sec* mutants block movement of membranes as well as cargo. See P. Novick et al., 1981, *Cell* **25**:461; and C. A. Kaiser and R. Schekman, 1990, *Cell* **61**:723.

To determine the order of the steps in the pathway, researchers analyzed double *sec* mutants. For instance, when yeast cells contain mutations in both class B and class D functions, proteins accumulate in the rough ER, not in the Golgi cisternae. Because proteins accumulate at the earliest blocked step, this finding shows that class B mutations must act at an earlier point in the secretory pathway than class D mutations do. These studies confirmed that as a secreted protein is synthesized and processed, it moves sequentially from the cytosol to the rough ER, to ER-to-Golgi transport vesicles, to Golgi cisternae, to secretory vesicles, and finally is exocytosed.

Most importantly, the yeast *sec* mutants define many of the genes and encoded proteins that are required for vesicle budding and fusion at each of the major steps of the secretory pathway. Each of the individual steps in the secretory pathway is currently being studied in mechanistic detail, and biochemical assays and protein structural studies are being used to understand each of these steps in terms of the structure and function of individual protein molecules.

Cell-Free Transport Assays Allow Dissection of Individual Steps in Vesicular Transport

In vitro assays for intercompartmental transport are powerful complementary approaches to studies with yeast *sec* mutants for identifying and analyzing the cellular components responsible for vesicular trafficking. In one application of this approach, cultured mutant cells lacking one of the enzymes that modify *N*-linked oligosaccharide chains in the Golgi are infected with vesicular stomatitis virus, and the fate of the VSV G protein is followed. For example, if infected cells lack *N*-acetylglucosamine transferase I, they produce abundant amounts of VSV G protein but cannot add *N*-acetylglucosamine residues to the oligosaccharide chains in the *medial*-Golgi as wild-type cells do (Figure 14-6a). When Golgi membranes isolated from such mutant cells are mixed with Golgi membranes from wild-type, uninfected cells, the addition of *N*-acetylglucosamine to VSV G protein is restored (Figure 14-6b). This modification is the consequence of vesicular transport of *N*-acetylglucosamine transferase I from the wild-type *medial*-Golgi to the *cis*-Golgi isolated from virally infected mutant cells.



EXPERIMENTAL FIGURE 14-6 A cell-free assay demonstrates protein transport from one Golgi cisterna to another. (a) A mutant line of cultured fibroblasts is essential in this type of assay. In this example, the cells lack the enzyme *N*-acetylglucosamine transferase I (see step **2** in Figure 14-15). In wild-type cells, this enzyme is localized to the *medial*-Golgi and modifies *N*-linked oligosaccharides by the addition of one *N*-acetylglucosamine. In VSV-infected wild-type cells, the oligosaccharide on the viral G protein is modified to a typical complex oligosaccharide, as shown in the *trans*-Golgi panel. In infected mutant cells, however, the G protein reaches the cell surface with a simpler high-mannose oligosaccharide containing only two *N*-acetylglucosamine and five mannose residues. (b) When Golgi cisternae isolated from infected mutant cells are incubated with Golgi cisternae from normal, uninfected cells, the VSV G protein produced in vitro contains the additional *N*-acetylglucosamine. This modification is carried out by transferase enzyme that is moved by transport vesicles from the wild-type *medial*-Golgi cisternae to the mutant *cis*-Golgi cisternae in the reaction mixture. See W. E. Balch et al., 1984, *Cell* **39**:405 and 525; W. A. Braell et al., 1984, *Cell* **39**:511; and J. E. Rothman and T. Söllner, 1997, *Science* **276**:1212.

Proteins that play a role in Golgi transport have been identified by fractionation methods. Under appropriate conditions, a uniform population of the transport vesicles that move *N*-acetylglucosamine transferase I from the *medial*- to *cis*-Golgi can be separated from the donor wild-type Golgi membranes by centrifugation. By examining the proteins that are enriched in these vesicles, scientists have been

able to identify many of the integral membrane proteins and peripheral vesicle coat proteins that are the structural components of this type of vesicle. Moreover, fractionation of the cytosolic extract required for transport in cell-free reaction mixtures has permitted isolation of the various proteins required for formation of transport vesicles and of proteins required for the targeting and fusion of vesicles with appropriate acceptor membranes. In vitro assays similar in general design to the one shown in Figure 14-6 have been used to study the proteins that play a role at various transport steps in the secretory pathway.

KEY CONCEPTS OF SECTION 14.1

Techniques for Studying the Secretory Pathway

- All assays for following the trafficking of proteins through the secretory pathway require a way to label a cohort of secretory proteins and a way to identify the compartments where the labeled proteins are subsequently located.
- Transport of a fluorescently labeled protein along the secretory pathway can be observed by microscopy (see Figure 14-3). A temperature-sensitive mutant protein that is retained in the ER due to misfolding at the nonpermissive temperature will be released as a cohort for transport when cells are shifted to the permissive temperature.
- Alternatively, pulse labeling with radioactive amino acids can specifically label a cohort of newly made proteins in the ER and the transport of a radiolabeled protein can be tracked by following compartment-specific covalent modifications to the protein.
- Many of the components required for intracellular protein trafficking have been identified in yeast by analysis of temperature-sensitive *sec* mutants defective for the secretion of proteins at the nonpermissive temperature (see Figure 14-5).
- Cell-free assays for intercompartmental protein transport have allowed the biochemical dissection of individual steps of the secretory pathway. Such in vitro

reactions can be used to produce pure transport vesicles and to test the biochemical function of individual transport proteins.

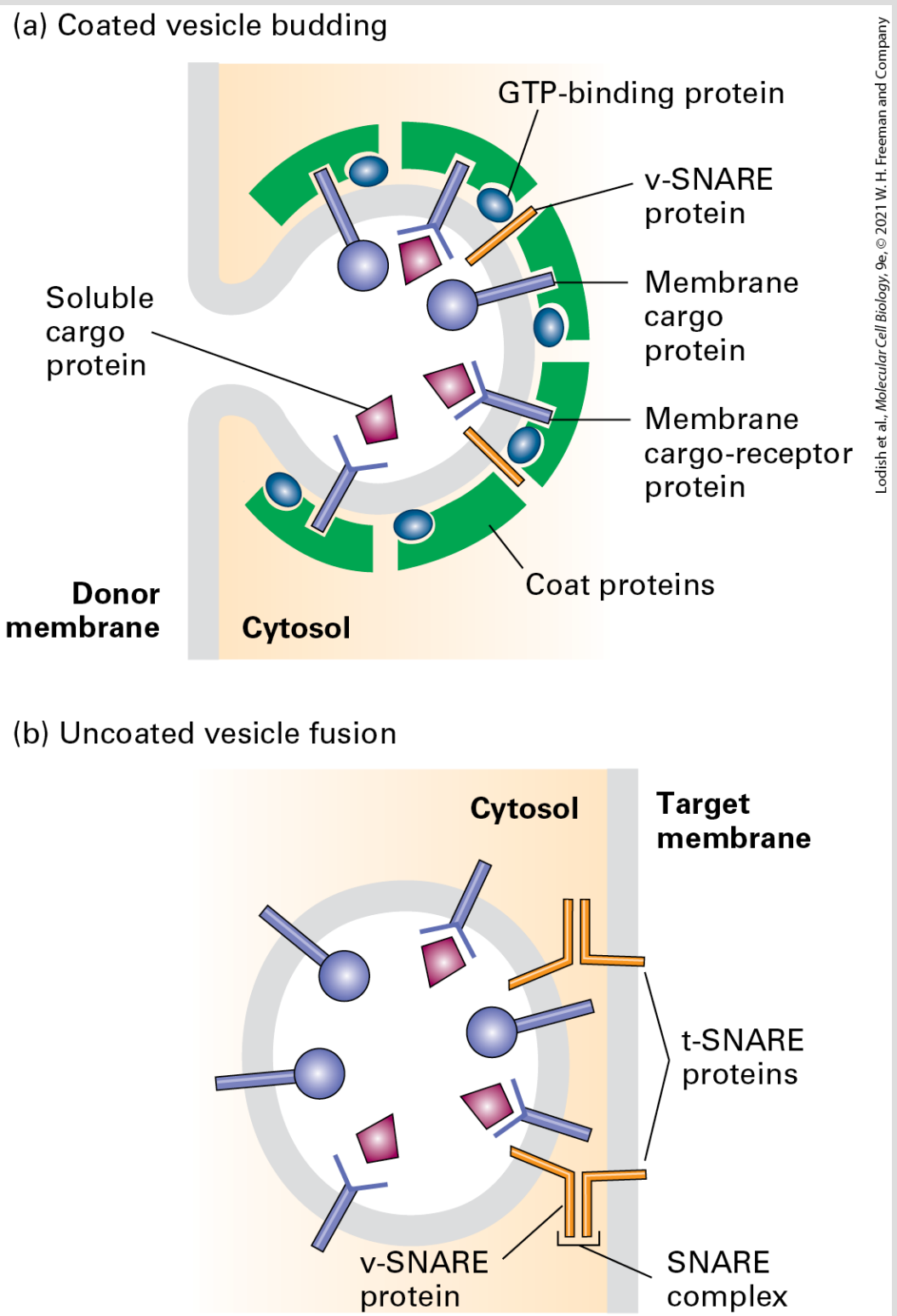
14.2 Molecular Mechanisms of Vesicle Budding and Fusion

Small membrane-bounded vesicles that transport proteins from one organelle to another are the fundamental functional elements in the secretory and endocytic pathways (see Figure 14-1). These vesicles bud from the membrane of a particular *parent (donor) organelle* and fuse with the membrane of a particular *target (destination) organelle*. Although each step in the secretory and endocytic pathways employs a different type of vesicle, studies employing genetic and biochemical techniques have revealed that each of the different vesicular transport steps is simply a variation on a common theme. In this section, we explore the basic mechanisms underlying vesicle budding and fusion that all vesicle types have in common, before discussing the details unique to each step in the pathway.

Assembly of a Protein Coat Drives Vesicle Formation and Selection of Cargo Molecules

The budding of a vesicle from its parent membrane is triggered by the activation of a GTP-binding protein followed by the polymerization of soluble protein complexes on the membrane to form a proteinaceous vesicle coat (Figure 14-7a). Interactions

between the cytosolic portions of integral membrane proteins and the vesicle coat gather the appropriate cargo proteins into the forming vesicle. Thus the coat gives curvature to the membrane to form a vesicle and acts as the filter to determine which proteins are admitted into the vesicle.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-7 Overview of vesicle budding and fusion with a target membrane. (a) Budding is initiated by recruitment of a small GTP-binding protein to a patch

of donor membrane. Complexes of coat proteins in the cytosol then bind to the cytosolic domain of membrane cargo proteins, some of which also act as receptors that bind soluble proteins in the lumen, thereby recruiting luminal cargo proteins into the budding vesicle. (b) After being released and shedding its coat, a vesicle fuses with its target membrane in a process that involves interaction of cognate SNARE proteins.

Proteins responsible for the eventual fusion of a vesicle with the target membrane, known as **v-SNAREs**, are incorporated into the vesicle membrane during assembly of the vesicle coat. Once the coat is at least partially shed from a completed vesicle, v-SNARE proteins embedded in the vesicle membrane become accessible to join with cognate **t-SNAREs** in the target membrane to which the vesicle is docked. This joining brings the membranes into close apposition, allowing the two bilayers to fuse (Figure 14-7b). Regardless of target organelle, all transport vesicles use v-SNAREs and t-SNAREs to fuse.

Three major types of coated vesicles have been characterized, each with a different type of protein coat and each formed by reversible polymerization of a distinct set of protein subunits (Table 14-1). Each type of vesicle, named for its primary coat proteins, transports cargo proteins from particular parent organelles to particular target organelles:

TABLE 14-1 • Coated Vesicles Involved in Protein Trafficking

Vesicle	Transport Step	Coat Proteins	Associated

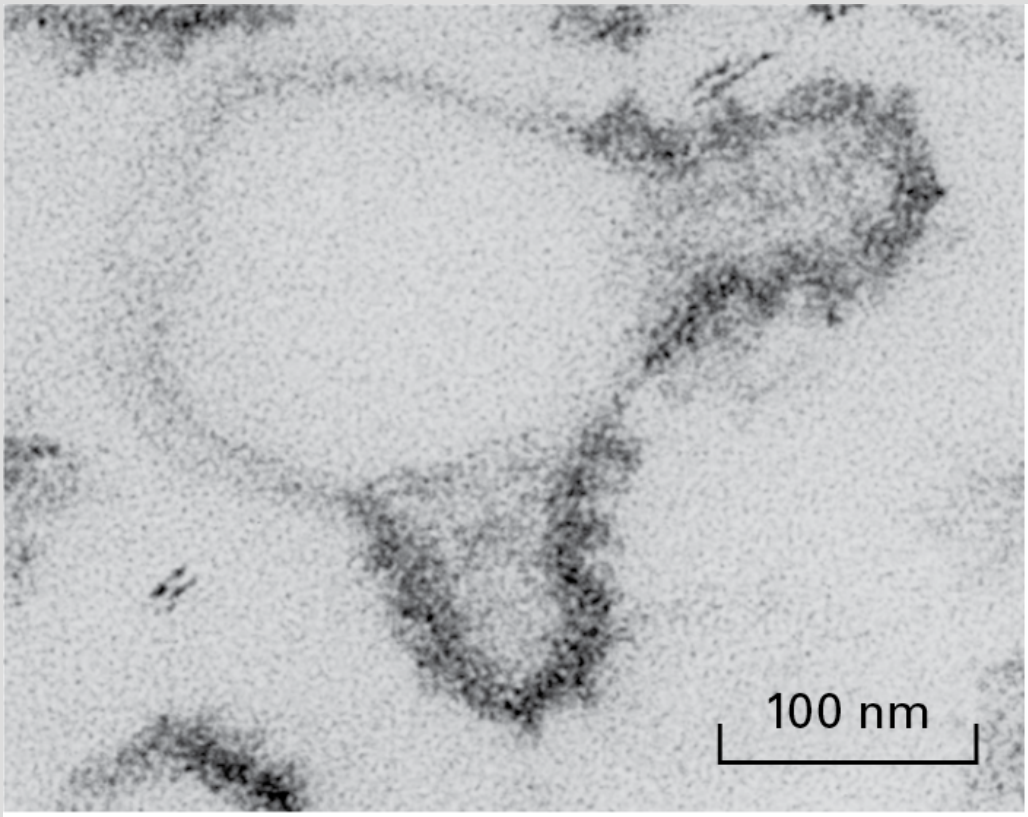
Type	Mediated		GTPase
COPII	ER to *cis*-Golgi	Sec23/Sec24 and Sec13/Sec31 complexes, Sec16	Sar1
COPI	*cis*-Golgi to ER later to earlier Golgi cisternae	Coatomers containing seven different COP subunits	ARF
Clathrin and adapter proteins[i]	*trans*-Golgi to endosome	$\text{Clathrin} + \text{AP1 complexes}$	ARF
	trans-Golgi to endosome	$\text{Clathrin} + \text{GGA}$	ARF
	Plasma membrane to endosome	$\text{Clathrin} + \text{AP2 complexes}$	ARF
	Golgi to lysosome, melanosome, or platelet vesicles	AP3 complexes	ARF

[i]Each type of AP complex consists of four different subunits. It is not known whether the coat of AP3 vesicles contains clathrin.

- **COPII** vesicles transport proteins from the ER to the Golgi.
- **COPI** vesicles mainly transport proteins in the retrograde direction between Golgi cisternae and from the *cis*-Golgi back to the ER.
- **Clathrin**-coated vesicles transport proteins from the plasma membrane (cell surface) and the *trans*-Golgi network to late endosomes.

Every vesicle-mediated trafficking step is thought to use some kind of vesicle coat; however, a specific coat protein complex has not been identified for every type of vesicle. For example, vesicles that move proteins from the *trans*-Golgi to the plasma membrane during either constitutive or regulated secretion exhibit a uniform size and morphology, which suggests that their formation is driven by assembly of a regular coat structure, yet researchers have not identified specific coat proteins surrounding these vesicles.

The general scheme of vesicle budding shown in Figure 14-7a applies to all three known types of coated vesicles. Experiments with isolated or artificial membranes and purified coat proteins have shown that polymerization of the coat proteins on the cytosolic face of the parent membrane is necessary to produce the high curvature of the membrane that is typical of a transport vesicle that is about 50 nm in diameter. Electron micrographs of in vitro budding reactions often reveal structures that exhibit discrete regions of the parent membrane bearing a dense coat accompanied by the curvature characteristic of a completed vesicle (Figure 14-8). Such structures, usually called *vesicle buds*, appear to be intermediates that are visible after the coat has begun to polymerize but before the completed vesicle pinches off from the parent membrane. The polymerized coat proteins form a curved lattice that drives the formation of a vesicle bud by adhering to the cytosolic face of the membrane.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 14-8 Vesicle buds can be visualized during in vitro budding reactions. When purified COPII coat components are incubated with isolated ER vesicles or artificial phospholipid vesicles (liposomes), polymerization of the coat proteins on the vesicle surface induces emergence of highly curved buds. In this electron micrograph of an in vitro budding reaction, note the distinct membrane coat, visible as a dark protein layer, present on the vesicle buds.

[Republished with permission from Elsevier, from K. Matsuoka et al., 1998, "COPII-Coated Vesicle Formation Reconstituted with Purified Coat Proteins and Chemically Defined Liposomes," *Cell* **93**(2):263–275; permission conveyed through Copyright Clearance Center, Inc.]

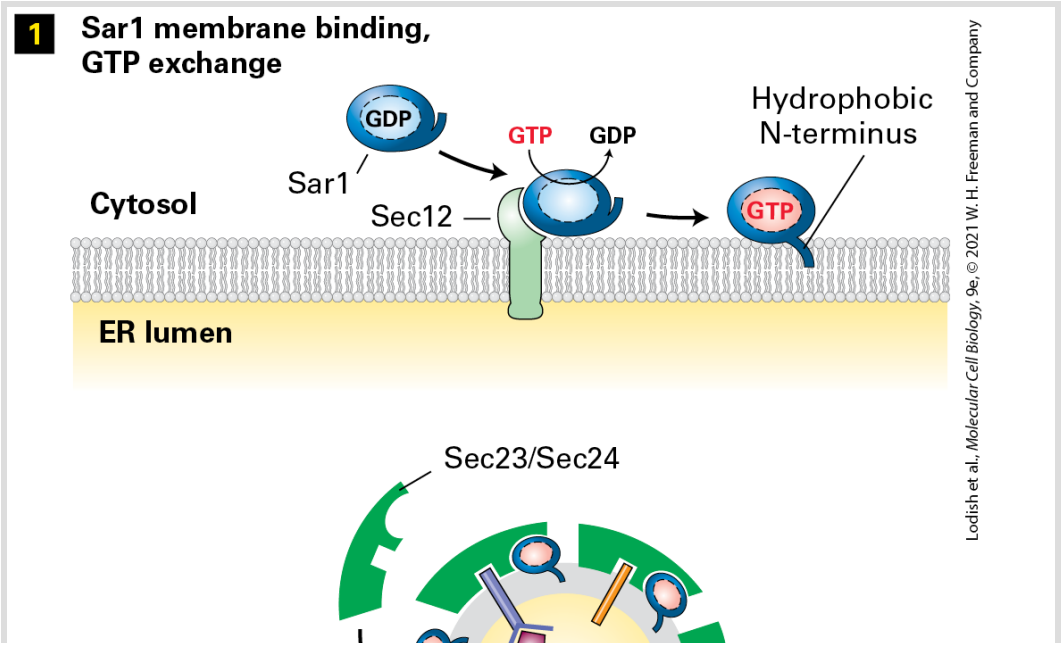
A Conserved Set of GTPase Switch Proteins Controls the Assembly of

Different Vesicle Coats

Using in vitro vesicle-budding reactions among isolated membranes and purified coat proteins, scientists have determined the minimum set of coat components required to form each of the three major types of vesicles. Although most of the coat proteins differ considerably from one type of vesicle to another, the coats of all three vesicles contain a small GTP-binding protein that acts as a regulatory subunit to control coat assembly (see Figure 14-7a). A GTP-binding protein known as **ARF protein** plays this role in COPI and clathrin-coated vesicles. A different but related GTP-binding protein known as *Sar1 protein* is present in the coat of COPII vesicles. Both ARF and Sar1 are monomeric proteins with a structure generally similar to that of Ras, a key intracellular signal-transducing protein (see Figure 16-12). ARF and Sar1 proteins, like Ras, belong to the GTPase superfamily of switch proteins that cycle between GDP-bound and GTP-bound forms (see Figure 3-9 to review the mechanism of GTPase switch proteins).

As depicted in Figure 14-9, activation of Sar1 by binding to GTP is the initial event that triggers assembly of the COPII vesicle coat. Note that Sar1 is the only GTPase switch protein activated at the ER membrane and thus only COPII-coated vesicles bud from the ER membrane. This initiating event begins with an ER membrane protein known as *Sec12*, which acts as a guanine nucleotide exchange factor (GEF) for Sar1 by catalyzing the release of GDP from

cytosolic Sar1·GDP and the binding of GTP. Sec12 apparently receives and integrates multiple as yet unknown signals, probably including the presence of cargo proteins that are ready to be transported at the ER membrane. Binding of GTP causes a conformational change in Sar1; this exposes its amphipathic N-terminus, which then becomes embedded in the phospholipid bilayer and tethers Sar1·GTP to the ER membrane (Figure 14-9, step **1**). The membrane-attached Sar1·GTP drives the polymerization of cytosolic complexes of COPII subunits on the membrane, eventually leading to formation of vesicle buds (step **2**). Once COPII vesicles are released from the donor membrane, the Sar1 GTPase activity hydrolyzes Sar1·GTP in the vesicle membrane to Sar1·GDP with the assistance of one of the coat subunits (step **3**). This hydrolysis triggers disassembly of the COPII coat (step **4**). Thus Sar1 couples a cycle of GTP binding and hydrolysis to the formation and then disassembly of the COPII coat.



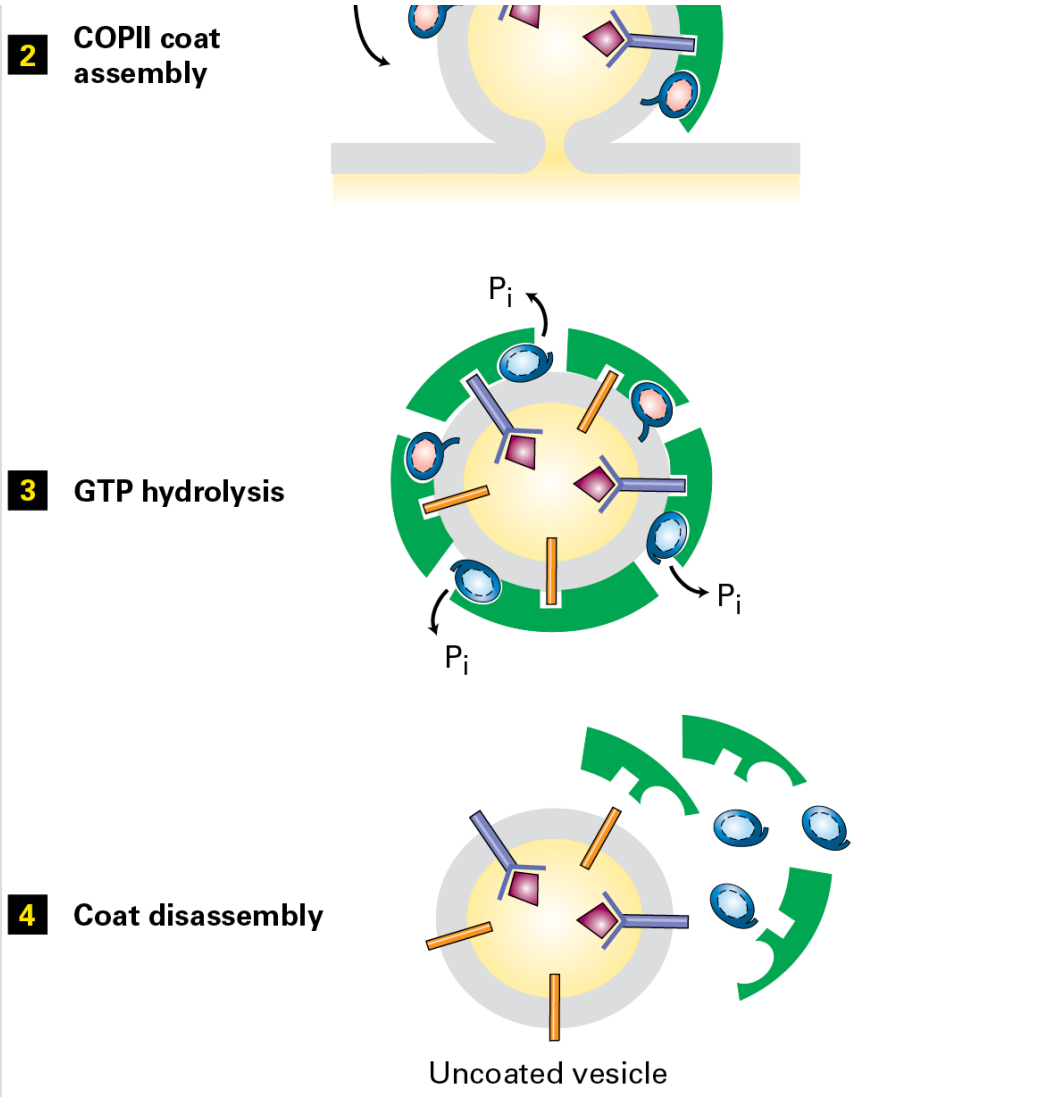


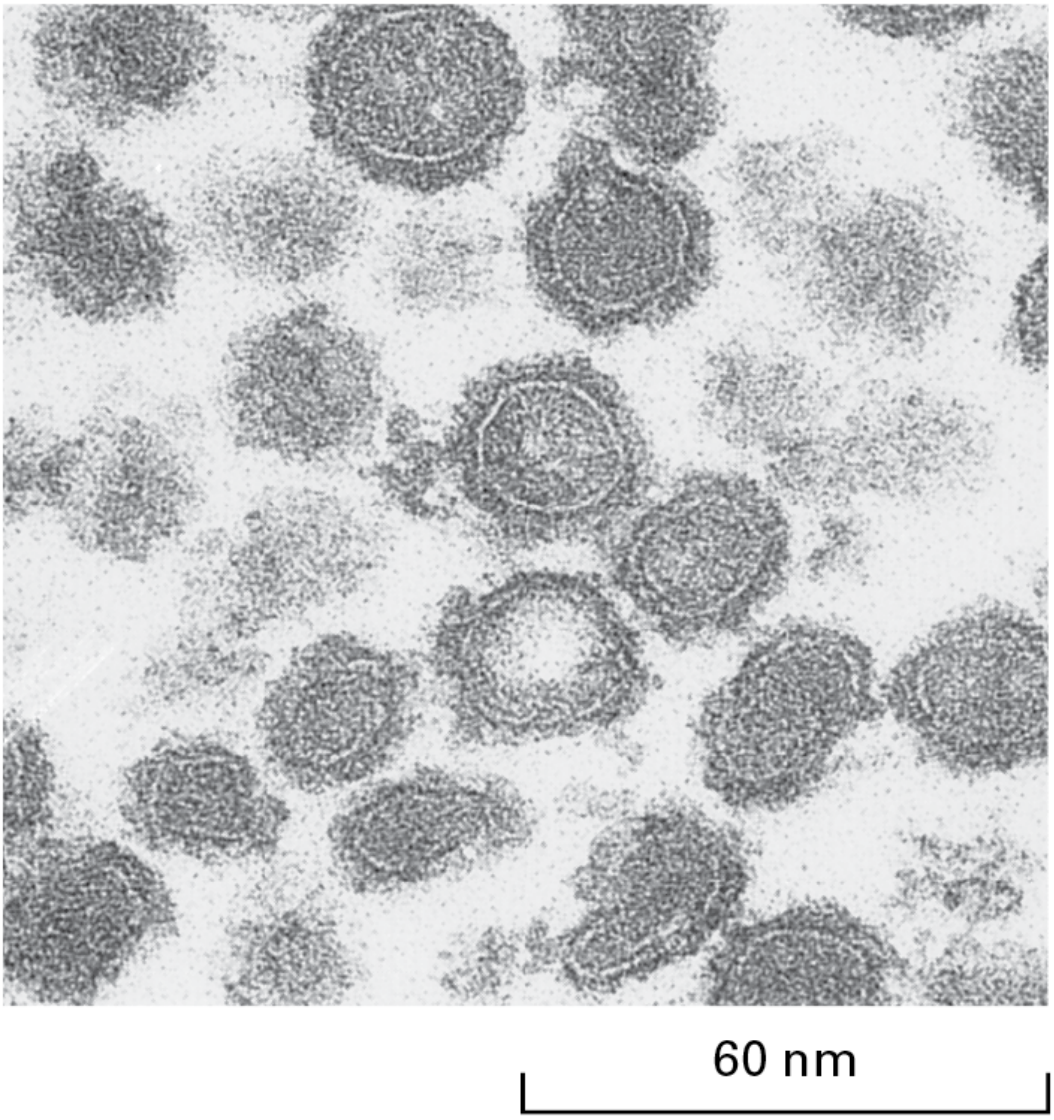
FIGURE 14-9 Model for the role of Sar1 in the assembly and disassembly of the COPII coat. Step 1: Interaction of soluble GDP-bound Sar1 with the GEF Sec12, an ER integral membrane protein, catalyzes exchange of GTP for GDP on Sar1. The hydrophobic N-terminus of the GTP-bound form of Sar1 extends outward from the protein's surface and anchors Sar1 to the ER membrane. Step 2: Once attached to the membrane, Sar1 serves as a binding site for the Sec23/Sec24 coat protein complex. Membrane cargo proteins are recruited to the forming vesicle bud by binding of specific short sequences (sorting signals) in their cytosolic regions to sites on the Sec23/Sec24 complex. Some membrane cargo proteins also act as receptors that bind soluble proteins in the lumen. The coat is completed by assembly of a second type of coat complex composed of Sec13 and Sec31 (not shown). Step 3: After the vesicle coat is

complete, the Sec23 coat subunit promotes GTP hydrolysis by Sar1. Step **4**: Release of Sar1·GDP from the vesicle membrane causes disassembly of the coat. See S. Springer et al., 1999, *Cell* **97**:145.

ARF protein undergoes a similar cycle of nucleotide exchange and hydrolysis coupled to the assembly of vesicle coats composed either of COPI or of clathrin and other coat proteins (AP complexes), discussed later. A covalent protein modification known as a *myristate anchor* on the N-terminus of the ARF protein weakly tethers ARF·GDP to the Golgi membrane. When GTP is exchanged for the bound GDP by a GEF attached to the Golgi membrane, the resulting conformational change in ARF allows hydrophobic residues in its N-terminal segment to insert into the membrane bilayer. The resulting tight association of ARF·GTP with the membrane serves as the foundation for further coat assembly.

Drawing on the structural similarities of Sar1 and ARF to other small GTPase switch proteins, researchers have constructed genes encoding mutant versions of the two proteins that have predictable effects on vesicular traffic when transfected into cultured cells. For example, in cells expressing mutant versions of Sar1 or ARF that cannot hydrolyze GTP, vesicle coats form and vesicle buds pinch off. However, because the mutant proteins cannot trigger disassembly of the coat, all available coat subunits eventually become permanently assembled into coated vesicles that are unable to fuse with target membranes. Addition of a nonhydrolyzable GTP analog to in vitro vesicle-budding reactions causes a similar blocking of coat

disassembly. The vesicles that form in such reactions have coats that never dissociate, allowing their composition and structure to be more readily analyzed. The purified COPI vesicles shown in Figure 14-10 were produced in such a budding reaction.



Courtesy of L. Orci (University of Geneva, Switzerland).

EXPERIMENTAL FIGURE 14-10 Coated vesicles accumulate during in vitro budding reactions in the presence of a nonhydrolyzable analog of GTP. When isolated Golgi membranes are incubated with a cytosolic extract containing COPI coat proteins, vesicles form and bud off from the membranes. Inclusion of a nonhydrolyzable

analog of GTP in the budding reaction prevents disassembly of the coat after vesicle release. This micrograph shows COPI vesicles generated in such a reaction and separated from membranes by centrifugation. Coated vesicles prepared in this way can be analyzed to determine their components and properties.

A second general function of small GTPases in vesicle formation is in the pinching off of a completed vesicle from the parent membrane. In vitro budding experiments show that the Sar1 GTPase that accumulates at the neck of the budding vesicle is required for the pinching off of COPII vesicles and that the ARF GTPase similarly drives the pinching off of COPI vesicles. The mechanism by which these small GTPases convert the energy from GTP hydrolysis to a mechanical force to complete the pinching off of the membrane is not understood. As we will see in Section 14.4, a large polymeric GTPase known as *dynamin* plays this role in clathrin-coated vesicles.

Targeting Sequences on Cargo Proteins Make Specific Molecular Contacts with Coat Proteins

In order for transport vesicles to move specific proteins from one compartment to the next, vesicle buds must be able to discriminate among potential membrane and soluble cargo proteins, accepting only those cargo proteins that should advance to the next compartment and excluding those that should remain as residents in the donor compartment. In addition to sculpting the curvature of

a donor membrane, the vesicle coat functions in selecting specific proteins as cargo. Membrane cargo proteins and soluble cargo proteins carry two different kinds of **sorting signals.** The sorting signals for membrane cargo proteins usually lie in the cytosolic portion of the protein and bind to one or another of the vesicle coat proteins (see Figure 14-7a). The polymerized coat thus acts as an affinity matrix to cluster selected membrane cargo proteins into forming vesicle buds. Because soluble proteins within the lumen of the parent organelle cannot contact the coat directly, they require a different kind of sorting signal. Soluble luminal proteins often contain what can be thought of as *luminal sorting signals*, which bind to the luminal domains of certain membrane cargo proteins. The properties of some of the known sorting signals in membrane and soluble proteins are summarized in Table 14-2. We describe the role of these signals in more detail in later sections.

TABLE 14-2 • Known Sorting Signals That Direct Proteins to Specific Transport Vesicles

Signal Sequence[ii]	Signal-Bearing Protein	Proteins with Signal	Signal Receptor
Cytoplasmic Sorting Signals			
Lys-Lys-X-X (KKXX)	ER-resident membrane proteins	COPI α and β subunits	COPI
Di-arginine (X-Arg-Arg-X)	ER-resident membrane proteins	COPI α and β subunits	COPI

Di-acidic (e.g., Asp-X-Glu)	Cargo membrane proteins in ER	COPII Sec24 subunit	COPII
Asn-Pro-X-Tyr (NPXY)	LDL receptor in plasma membrane	AP2 complex	Clathrin/AP2
Tyr-X-X-Φ (YXXΦ)	Membrane proteins in *trans*-Golgi	AP1 (μ1 subunit)	Clathrin/AP1
	Plasma membrane proteins	AP2 (μ2 subunit)	Clathrin/AP2
Leu-Leu (LL)	Plasma membrane proteins	AP2 complexes	Clathrin/AP2
Luminal Sorting Signals			
Lys-Asp-Glu-Leu (KDEL)	ER-resident soluble proteins	KDEL receptor in *cis*-Golgi membrane	COPI
Mannose 6-phosphate (M6P)	Soluble lysosomal enzymes after processing in *cis*-Golgi	M6P receptor in *trans*-Golgi membrane	Clathrin/AP1
	Secreted lysosomal enzymes	M6P receptor in plasma membrane	Clathrin/AP2

[ii] $\mathrm{X} = \text{any amino acid};\ \Phi = \text{hydrophobic amino acid}$. Single-letter amino acid abbreviations are in parentheses.

Rab GTPases Control Docking of Vesicles on Target Membranes

A second set of small GTP-binding proteins, known as **Rab proteins**, associate with transport vesicles and act as key regulators of vesicle trafficking to and fusion with the appropriate target membrane. Like

Sar1 and ARF, Rab proteins belong to the GTPase superfamily of switch proteins. Rab proteins also contain an isoprenoid anchor that allows them to become tethered to the vesicle membrane. Association of an activated Rab protein with a specific vesicle type is generally a two-step process. In the first step, cytosolic Rab·GDP is targeted to the appropriate vesicle, becoming attached there by insertion of its isoprenoid anchor into the vesicle membrane. Often this attachment step is facilitated by a protein that can associate with Rab·GDP along with its isoprenoid anchor, usually known as a *guanine nucleotide dissociation inhibitor (GDI)*. In the second step, a specific GEF located in the vesicle membrane converts membrane-bound Rab·GDP to Rab·GTP. Once localized and activated in this way, Rab·GTP is enabled to bind to a variety of different proteins, known as *Rab effectors*. Binding of Rab·GTP to a Rab effector can ultimately lead to docking of the vesicle on an appropriate target membrane (Figure 14-11a, step **1**). After vesicle fusion occurs, the GTP bound to the Rab protein is hydrolyzed to GDP, triggering the release of Rab·GDP, which can then undergo another cycle of GDP-GTP exchange, binding, and hydrolysis.

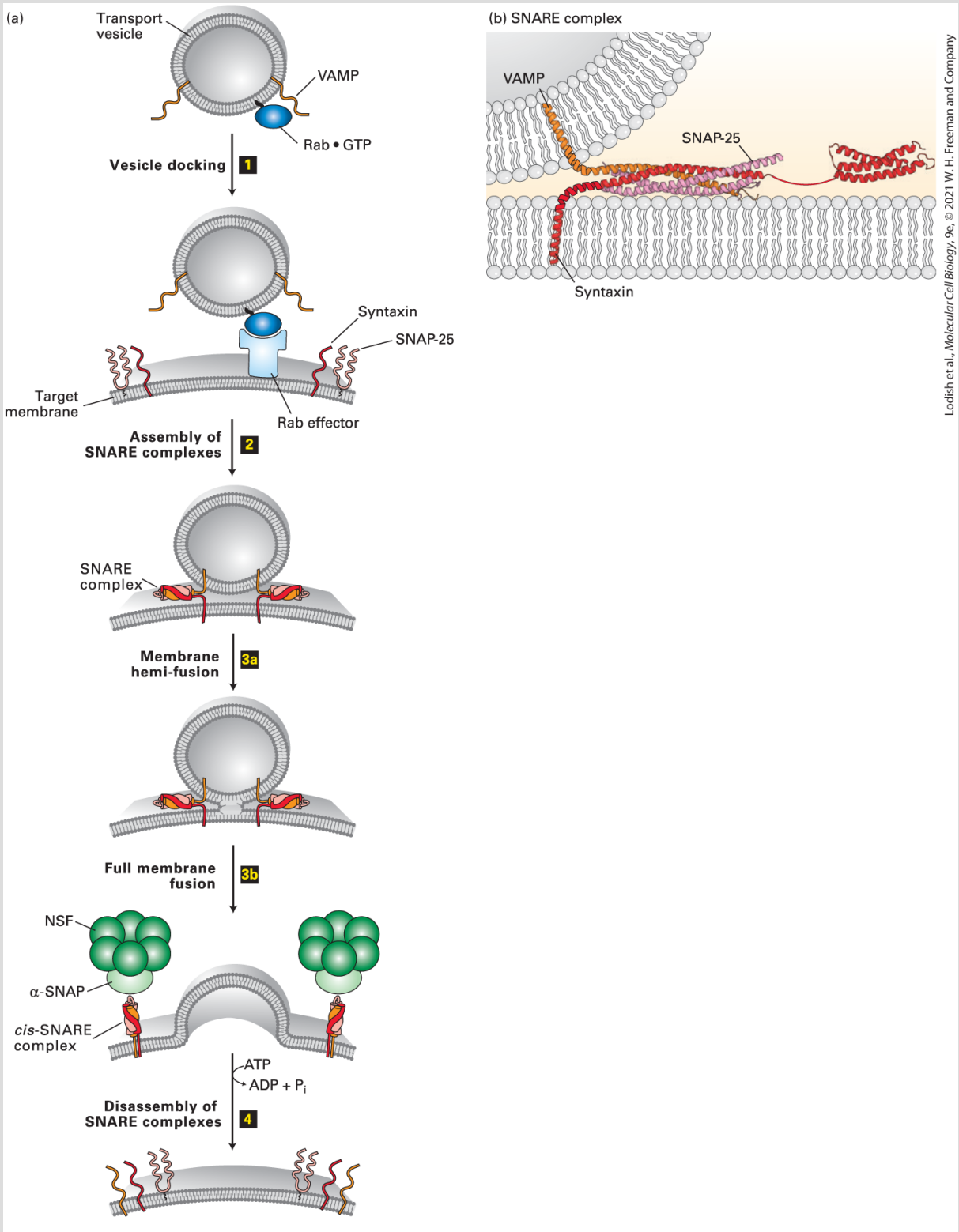


FIGURE 14-11 Model for docking and fusion of transport vesicles with their target membranes. (a) The proteins shown in this example participate in fusion of secretory vesicles with the plasma membrane, but similar proteins mediate all vesicle-fusion events.

Step **1**: A Rab protein tethered via a lipid anchor to a secretory vesicle binds to an effector protein complex on the plasma membrane, thereby docking the transport vesicle on the appropriate target membrane. Step **2**: A v-SNARE protein (in this case, VAMP) interacts with the cytosolic domains of the cognate t-SNAREs (in this case, syntaxin and SNAP-25). The very stable coiled-coil SNARE complexes that are formed hold the vesicle close to the target membrane. Step **3**: Fusion of the two membranes immediately follows formation of SNARE complexes, first by fusion of one leaflet of the bilayer to form a hemi-fused state (**3a**), followed by fusion of the second leaflet (**3b**). Step **4**: Following membrane fusion, NSF, in conjunction with α-SNAP, binds to the SNARE complexes. The NSF-catalyzed hydrolysis of ATP then drives dissociation of the SNARE complexes, freeing the SNARE proteins for another round of vesicle fusion. Also at this time, Rab·GTP is hydrolyzed to Rab·GDP and dissociates from the Rab effector (not shown). (b) The SNARE complex. Numerous noncovalent interactions between four long α helices, two from SNAP-25 and one each from syntaxin and VAMP, stabilize the coiled-coil structure. See J. E. Rothman and T. Söllner, 1997, *Science* **276**:1212; Y. A. Chen and R. H. Scheller, 2001, *Nat. Rev. Mol. Cell Biol.* **2**:98; and W. Weis and R. Scheller, 1998, *Nature* **395**:328.

[Part (b) Data from I. Fernandez et al., 1998, *Cell* **94**:841–849, PDB ID 1br0; and R. B. Sutton et al., 1998, *Nature* **395**:347–353, PDB ID 1sfc.]

A well-understood example of a Rab protein that enables vesicle fusion with the correct target membrane is the Sec4 protein of yeast, which specifically tags secretory vesicles, enabling them to fuse with the plasma membrane. In experiments, yeast cells expressing mutant Sec4 proteins accumulate secretory vesicles that are unable to fuse with the plasma membrane (class E mutants in Figure 14-5). Sec4·GDP binds to secretory vesicles, where it is activated to Sec4·GTP by its cognate GEF, which is itself located on secretory vesicles. Sec4·GTP, in turn, binds to its effector, a large tethering complex composed of eight subunits, known as the *exocyst*.

Tethering of secretory vesicles to the exocyst by binding of Sec4·GTP ultimately leads to vesicle fusion with the plasma membrane.

In mammalian cells, Rab5 protein is localized to endocytic vesicles, also known as *early endosomes*. These uncoated vesicles form from clathrin-coated vesicles just after they bud from the plasma membrane during endocytosis (see Figure 14-2, step **9**). The fusion of early endosomes with one another in cell-free systems requires the presence of Rab5, and addition of Rab5 and GTP to cell-free extracts accelerates the rate at which these vesicles fuse with one another. A long coiled protein known as *EEA1 (e*arly *e*ndosome *a*ntigen *1*), which resides on the membrane of the early endosome, functions as the effector for Rab5. In this case, Rab5·GTP on one endocytic vesicle is thought to bind specifically to EEA1 on the membrane of another endocytic vesicle, setting the stage for fusion of the two vesicles.

Every type of transport vesicle appears to be labeled with one or more specific Rab proteins. These Rab proteins, through their specific association with effectors that are membrane tethers and molecular motors, ensure that the vesicles are directed to the correct target membrane address (Table 14-3). The tracks on which the molecular motors run to deliver the vesicles to their destinations — microfilaments — will be discussed in Chapter 17.

TABLE 14-3 • Rab Proteins Involved in Vesicle Tethering and Vesicle

Movement

Rab Protein	Transport Step	Effector Proteins
Rab1/Ypt1	ER to *cis*-Golgi	TRAPP complex and GM130 complex in *cis*-Golgi
Sec4	*trans*-Golgi to plasma membrane	Exocyst on plasma membrane and Type V Myosin
Rab5	Plasma membrane to endosome	EEA1 tether on endosome
Rab6	Golgi movement along microtubules	Kinesin
Rab27	Melanosome movement to cell periphery	Type V Myosin

Paired Sets of SNARE Proteins Mediate Fusion of Vesicles with Target Membranes

As noted previously, shortly after a vesicle buds off from the donor membrane, the vesicle coat disassembles to uncover a vesicle-specific membrane protein, a v-SNARE (see Figure 14-7b). Likewise, each type of target membrane in a cell contains t-SNARE membrane proteins, which interact specifically with v-SNAREs. After Rab-mediated docking of a vesicle on its target membrane, the

interaction of cognate SNAREs brings the two membranes close enough together that they can fuse.

One of the best understood examples of SNARE-mediated fusion occurs during exocytosis of secreted proteins (see Figure 14-11a, steps **1** and **3**). In this case, the v-SNARE, known as *VAMP* (*v*esicle-*a*ssociated *m*embrane *p*rotein), is incorporated into secretory vesicles as they bud from the *trans*-Golgi network. The t-SNAREs are *syntaxin*, an integral membrane protein in the plasma membrane, and *SNAP-25*, which is attached to the plasma membrane by a hydrophobic lipid anchor in the middle of the protein. The cytosolic region in each of these three SNARE proteins contains a repeating heptad sequence that allows four α helices — one from VAMP, one from syntaxin, and two from SNAP-25 — to coil around one another to form a four-helix bundle (Figure 14-11b). The unusual stability of this bundled SNARE complex is conferred by the arrangement of hydrophobic and charged amino acid residues in the heptad repeats. The hydrophobic amino acids are buried in the central core of the bundle, and amino acids of opposite charge are aligned to form favorable electrostatic interactions between helices. As multiple four-helix bundles form, the embedded transmembrane domains of VAMP and syntaxin pull the vesicle and target membranes together into very close apposition. The energetically favorable formation of four-helix bundles can overcome the electrostatic repulsion of the generally negatively charged phospholipid head groups in the vesicle and target membranes,

allowing the hydrophobic interiors of the two membranes to come into contact, creating an opening between the two membranes, and ultimately causing the vesicle membrane to fuse with the target membrane.

In vitro experiments have shown that when liposomes containing purified VAMP are incubated with other liposomes containing syntaxin and SNAP-25, the two classes of membranes fuse, albeit slowly. This finding is strong evidence that the close apposition of membranes resulting from formation of SNARE complexes is sufficient to bring about membrane fusion. Fusion of a vesicle and target membrane occurs more rapidly and efficiently in the cell than it does in liposome experiments in which fusion is catalyzed only by SNARE proteins. The likely explanation for this difference is that in the cell, other proteins, such as Rab proteins and their effectors, are involved in targeting vesicles to the correct membrane.

Yeast cells, like all eukaryotic cells, express more than 20 different related v-SNARE and t-SNARE proteins. Analyses of yeast mutants defective in each of the SNARE genes have identified specific membrane-fusion events in which each SNARE protein participates. For all fusion events that have been examined, the SNAREs form four-helix bundled complexes similar to the VAMP/syntaxin/SNAP-25 complexes that mediate fusion of secretory vesicles with the plasma membrane. However, in other fusion events (e.g., fusion of COPII vesicles with the *cis*-Golgi network), each participating SNARE

protein contributes only one α helix to the bundle (unlike SNAP-25, which contributes two helices); in these cases, the SNARE complexes comprise one v-SNARE and three t-SNARE molecules.

Using the in vitro liposome fusion assay, researchers have tested the ability of various combinations of individual v-SNARE and t-SNARE proteins to mediate fusion of donor and target membranes. Of the very large number of different combinations tested, only a small number could efficiently mediate membrane fusion. To a remarkable degree, the functional combinations of v-SNAREs and t-SNAREs revealed in these in vitro experiments correspond to the actual SNARE protein interactions that mediate known membrane-fusion events in the yeast cell. Thus together with the specificity of interaction between Rab and Rab effector proteins, the specificity of the interaction between SNARE proteins can account for most, if not all, of the specificity of fusion between a particular vesicle type and its target membrane.

Dissociation of SNARE Complexes After Membrane Fusion Is Driven by ATP Hydrolysis

After a vesicle and its target membrane have fused, the SNARE complexes must dissociate to make the individual SNARE proteins available for additional fusion events. Because of the stability of

SNARE complexes, which are held together by numerous noncovalent intermolecular interactions, their dissociation depends on additional proteins and the input of energy.

The first clue that dissociation of SNARE complexes required the assistance of other proteins came from in vitro transport reactions depleted of certain cytosolic proteins. The observed accumulation of vesicles in these reactions indicated that vesicles could form under these conditions, but were unable to fuse with a target membrane. Eventually two proteins, designated *NSF* and *α-SNAP*, were found to be required for ongoing vesicle fusion in the in vitro transport reaction.

Yeast mutants have also contributed to our understanding of SNARE function. Among the class C yeast *sec* mutants are strains that lack functional Sec18 or Sec17, the yeast counterparts of mammalian NSF and α-SNAP, respectively. When these class C mutants are kept at the nonpermissive temperature, they accumulate ER-to-Golgi transport vesicles; when the cells are shifted to the lower, permissive temperature, the accumulated vesicles are able to fuse with the *cis*-Golgi.

Subsequent to the initial biochemical and genetic studies that identified NSF and α-SNAP, more sophisticated in vitro transport assays were developed. Using these newer assays, researchers have shown that NSF and α-SNAP proteins are not necessary for actual

membrane fusion, but rather are required for regeneration of free SNARE proteins. NSF, a hexamer of identical subunits, associates with a SNARE complex with the aid of α-SNAP (*s*oluble *N*SF *a*ttachment *p*rotein). The bound NSF then hydrolyzes ATP, releasing sufficient energy to dissociate the SNARE complex (see Figure 14-11a, step **4**). Evidently, the defects in vesicle fusion observed in the earlier in vitro fusion assays and in the yeast mutants after a loss of Sec17 or Sec18 were a consequence of free SNARE proteins rapidly becoming sequestered in undissociated SNARE complexes and thus being unavailable to mediate membrane fusion.

KEY CONCEPTS OF SECTION 14.2

Molecular Mechanisms of Vesicle Budding and Fusion

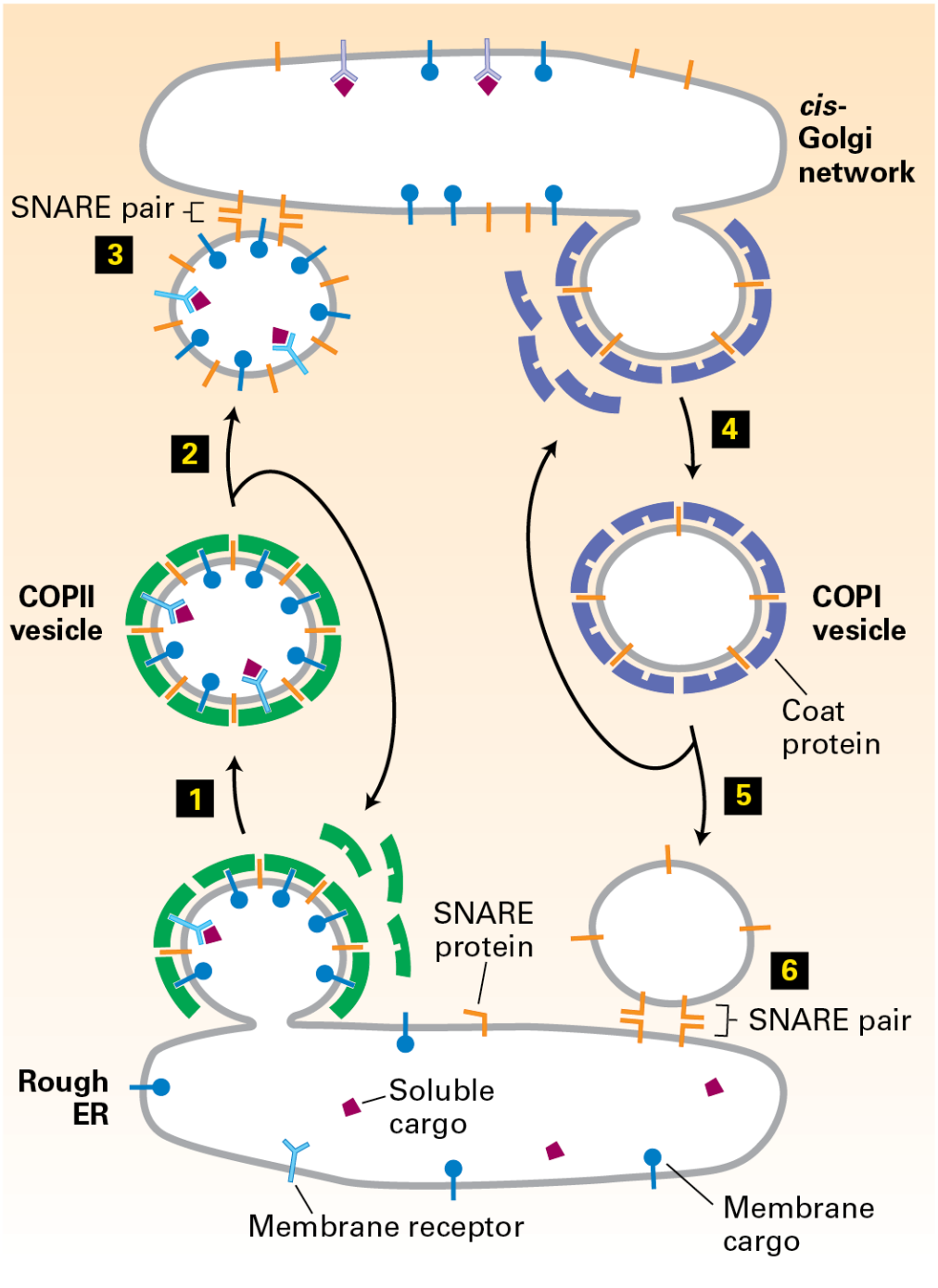
- The three well-characterized types of transport vesicles — COPI, COPII, and clathrin-coated vesicles — are distinguished by the proteins that form their coats and the transport routes they mediate (see Table 14-1).
- All types of coated vesicles are formed by polymerization of cytosolic coat proteins on a parent (donor) membrane to form vesicle buds that eventually pinch off from the membrane to release a complete vesicle. Shortly after vesicle release, the coat is shed, exposing proteins required for fusion with the target membrane (see Figure 14-7).
- Small GTP-binding proteins (ARF or Sar1) belonging to the GTPase superfamily control polymerization of coat proteins, the initial step in vesicle budding (see Figure 14-9). After vesicles are released from the donor membrane, hydrolysis of GTP bound to ARF or Sar1 triggers disassembly of the vesicle coats.
- Specific sorting signals in membrane and luminal proteins in donor organelles interact with coat proteins during vesicle budding, thereby recruiting cargo proteins to vesicles (see Table 14-2).
- A second set of GTP-binding proteins, the Rab proteins, label specific vesicle types and mediate their targeting to the appropriate membrane (see Table 14-3). Activated Rab·GTP in a vesicle can bind to a specific type of effector protein. One type of effector

is a peripheral membrane protein complex that tethers the vesicle to the correct target membrane.

- Each v-SNARE in a vesicular membrane specifically binds to a complex of cognate t-SNARE proteins in the target membrane, inducing fusion of the two membranes. After fusion is completed, the SNARE complex is disassembled in an ATP-dependent reaction mediated by other cytosolic proteins (see Figure 14-11).

14.3 Early Stages of the Secretory Pathway

In this section, we take a closer look at vesicular traffic between the ER and the Golgi and at some of the evidence supporting the general mechanisms discussed in the previous section. Recall that anterograde transport from the ER to the Golgi, the first vesicle trafficking step in the secretory pathway, is mediated by COPII vesicles. These vesicles contain newly synthesized proteins destined for the Golgi, cell surface, or lysosomes, as well as vesicle components such as v-SNAREs that are required to target vesicles to the *cis*-Golgi membrane. Proper sorting of proteins between the ER and Golgi also requires retrograde transport from the *cis*-Golgi to the ER, which is mediated by COPI vesicles (Figure 14-12). This retrograde vesicle transport serves to retrieve v-SNARE proteins and components of the membrane itself to provide the necessary material for additional rounds of vesicle budding from the ER. COPI-mediated retrograde transport also retrieves missorted ER-resident proteins from the *cis*-Golgi.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-12 Vesicle-mediated protein trafficking between the ER and *cis*-Golgi. Steps 1 – 3: Forward (anterograde) transport is mediated by COPII vesicles, which are formed by polymerization of soluble COPII coat protein complexes (green) on the ER membrane. v-SNAREs (orange) and other cargo proteins (blue) in the ER membrane are incorporated into the vesicle by interacting with coat proteins. Soluble cargo proteins (magenta) are recruited

by binding to appropriate receptors in the membrane of budding vesicles. Dissociation of the coat recycles free coat complexes and exposes v-SNARE proteins on the vesicle surface. After the uncoated vesicle becomes tethered to the *cis*-Golgi membrane in a Rab-mediated process, pairing between the exposed v-SNAREs and cognate t-SNAREs in the Golgi membrane allows membrane fusion, releasing the contents of the vesicle into the *cis*-Golgi compartment (see Figure 14-10). Steps **4** – **6**: Reverse (retrograde) transport, mediated by vesicles coated with COPI proteins (purple), recycles the membrane bilayer and certain proteins, such as v-SNAREs and missorted ER-resident proteins (not shown), from the *cis*-Golgi to the ER. All SNARE proteins are shown in orange, although v-SNAREs and t-SNAREs are distinct proteins.

We also discuss in this section the process by which proteins that have been correctly delivered to the Golgi advance through successive compartments of the Golgi, from the *cis*- to the *trans*-Golgi network. This process of cisternal maturation involves budding and fusion of retrograde rather than anterograde transport vesicles.

COPII Vesicles Mediate Transport from the ER to the Golgi

COPII vesicles were first recognized when cell-free extracts of yeast rough ER membranes were incubated with cytosol and a nonhydrolyzable analog of GTP. The vesicles that formed from the ER membranes had a distinct coat similar to that on COPI vesicles, but composed of different proteins, designated *COPII proteins*. The genes encoding COPII proteins were identified by the analysis of

yeast cells with class B *sec* mutations that accumulate proteins in the rough ER (see Figure 14-5).

As described previously, formation of COPII vesicles is triggered when Sec12, a GEF in the ER membrane, catalyzes the exchange of bound GDP for GTP on cytosolic Sar1. This exchange induces binding of Sar1 to the ER membrane, followed by binding of a complex of Sec23 and Sec24 proteins (see Figure 14-9). The resulting ternary complex formed between Sar1·GTP, Sec23, and Sec24 is shown in Figure 14-13. This core coat protein complex then provides binding sites for the recruitment of a second complex of Sec13 and Sec31 proteins to complete the coat structure. Since pure Sec13 and Sec31 proteins can spontaneously assemble into cagelike lattices, it is thought that Sec13 and Sec31 form the structural scaffold for COPII vesicles (see the COPII vesicle model in the chapter opener figure). Finally, a large fibrous protein called *Sec16*, which is bound to the cytosolic surface of the ER, interacts with Sar1·GTP and the Sec13/31 and Sec23/24 complexes to organize the other coat proteins, increasing the efficiency of coat polymerization.

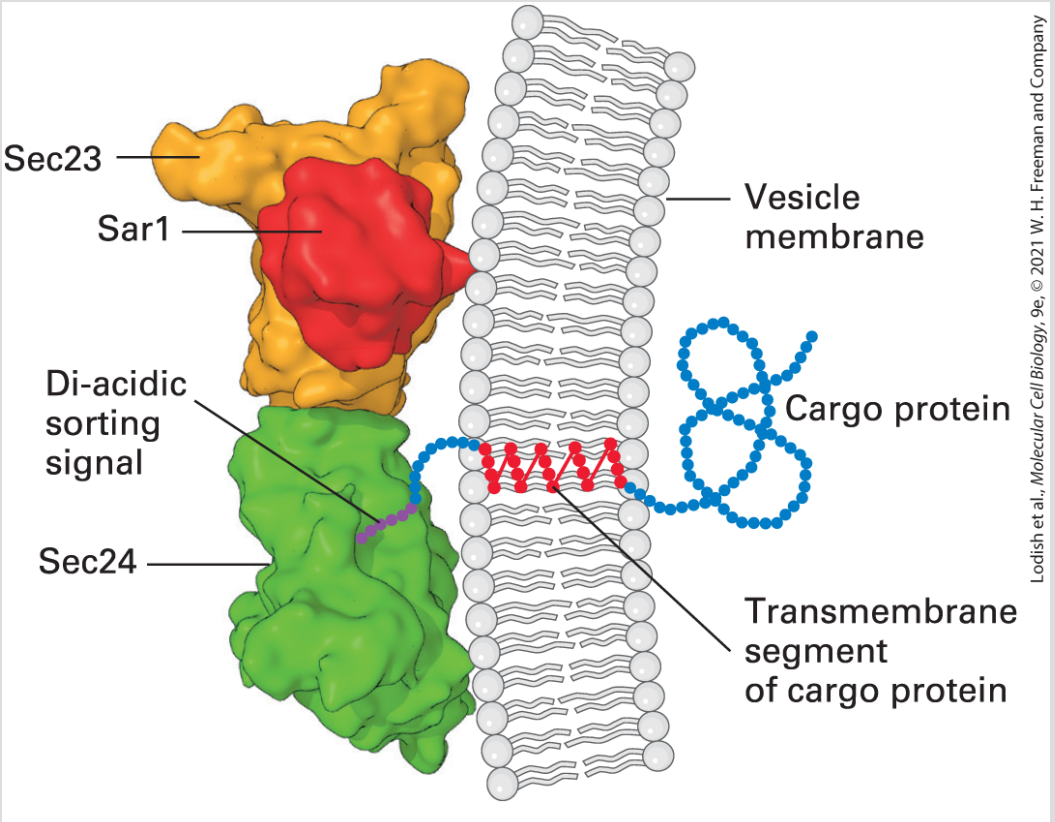


FIGURE 14-13 Three-dimensional structure of the ternary complex comprising the COPII coat proteins Sec23, Sec24, and Sar1·GTP. Early in the formation of the COPII coat, Sec23 (orange)/Sec24 (green) complexes are recruited to the ER membrane by Sar1 (red) in its GTP-bound state. A cargo protein in the ER membrane can be recruited to COPII vesicles by the interaction of a tripeptide di-acidic sorting signal (purple) in the cargo protein's cytosolic domain with Sec24. The likely positions of the COPII vesicle membrane and the transmembrane segment of the cargo protein are indicated. The N-terminal segment of Sar1 that tethers it to the membrane is not shown.

[Republished with permission from *Nature*, from X. Bi et al., 2002, "Structure of the Sec23/24-Sar1 Prebudding Complex of the COPII Vesicle Coat," *Nature* **419**(6904):271–277; permission conveyed through the Copyright Clearance Center, Inc.]

Certain integral ER membrane proteins are specifically recruited into COPII vesicles for transport to the Golgi. The cytosolic segments

of many of these proteins contain a *di-acidic sorting signal* (the key residues in this sequence are Asp-X-Glu, or DXE in the one-letter code) (see Table 14-2). This sorting signal, which binds to the Sec24 subunit of the COPII coat, is essential for the selective export of certain membrane proteins from the ER (see Figure 14-12). Biochemical and genetic studies have identified additional signals that help direct membrane cargo proteins into COPII vesicles. All of the known sorting signals bind to one or another site on the Sec24 subunit of COPII. Ongoing studies seek to determine how soluble cargo proteins are selectively loaded into COPII vesicles.

The inherited disease cystic fibrosis is characterized by an imbalance in chloride and sodium ion transport in the epithelial cells of the lungs, leading to fluid buildup and difficulty breathing. Cystic fibrosis is caused by mutations in a protein known as *CFTR*, which is synthesized as an integral membrane protein in the ER and is transported to the Golgi before being transported to the plasma membranes of epithelial cells, where it functions as a chloride channel. Researchers have recently shown that the CFTR protein contains a di-acidic sorting signal that binds to the Sec24 subunit of the COPII vesicle coat and is necessary for transport of the CFTR protein out of the ER. The most common CFTR mutation is a deletion of a phenylalanine at position 508 in the protein sequence (known as $\Delta F508$). This mutation prevents normal transport of CFTR to the plasma membrane by blocking its packaging into COPII vesicles budding from the ER. While the CFTR protein with the

ΔF508 mutation triggers the ER quality-control machinery, a folded CFTR with this mutation can still function properly as a normal chloride channel. However, it never reaches the membrane; the disease state is therefore caused by the absence of the channel, rather than by a defective channel. ■

COPI Vesicles Mediate Retrograde Transport Within the Golgi and from the Golgi to the ER

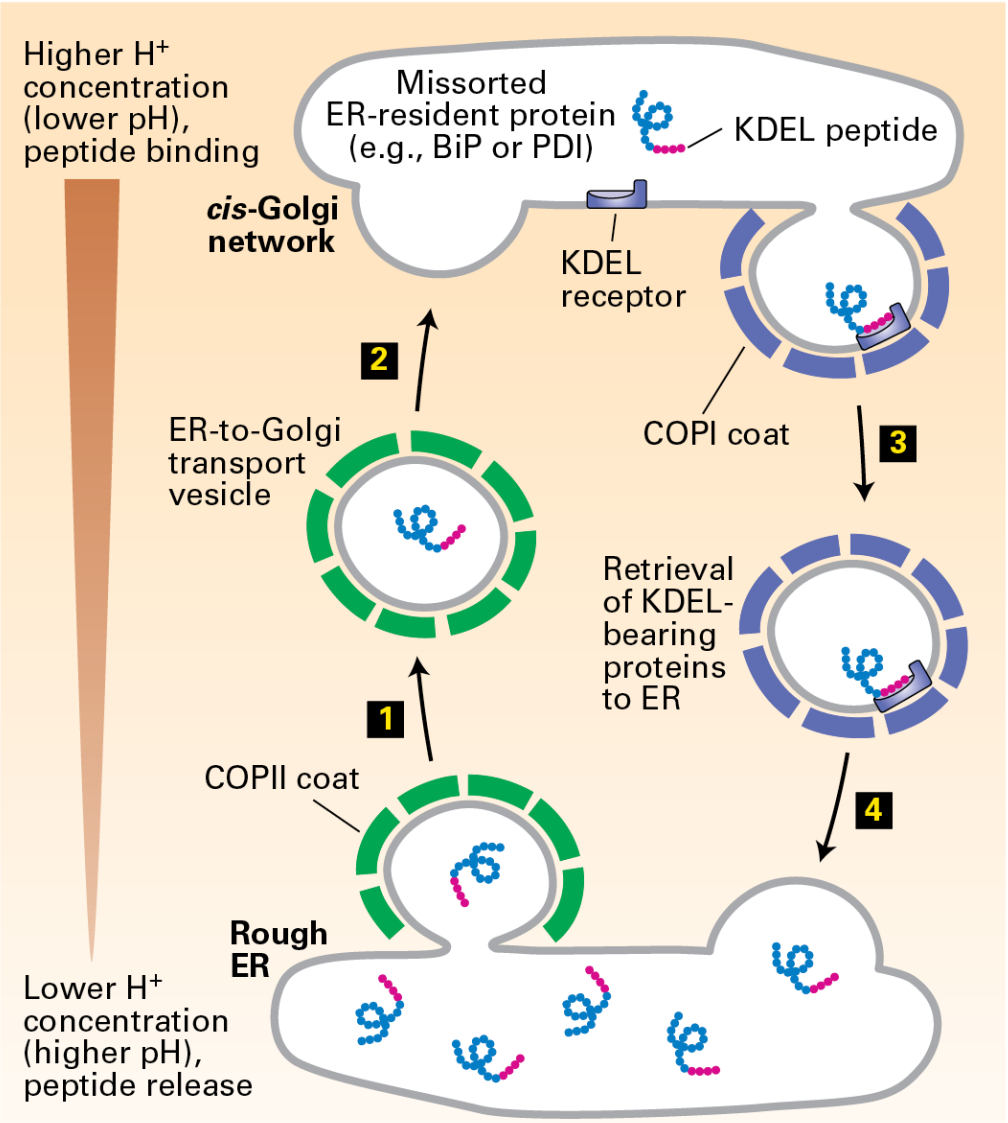
COPI vesicles were first discovered when isolated Golgi fractions were incubated in a solution containing cytosol and a nonhydrolyzable analog of GTP (see Figure 14-10). Subsequent analysis of these vesicles showed that the coat is formed from large cytosolic complexes, called *coatomers*, composed of seven polypeptide subunits. Yeast cells containing temperature-sensitive mutations in COPI proteins accumulate proteins in the rough ER at the nonpermissive temperature and thus are categorized as class B *sec* mutants (see Figure 14-5). Although the discovery of these mutants initially suggested that COPI vesicles mediate ER-to-Golgi transport, subsequent experiments showed that their main function is retrograde transport, both between Golgi cisternae and from the *cis*-Golgi to the rough ER (see Figure 14-12, *right*). Because COPI mutants cannot recycle key membrane proteins back to the rough ER, the ER gradually becomes depleted of ER proteins, such as v-

SNAREs, that are necessary for COPII vesicle function. Eventually, vesicle formation from the rough ER grinds to a halt; secretory proteins continue to be synthesized but accumulate in the ER — the defining characteristic of class B *sec* mutants. The general ability of *sec* mutations involved in either COPI or COPII vesicle function to eventually block both anterograde and retrograde transport illustrates the fundamental interdependence of these two transport processes.

As discussed in Chapter 13, the ER contains several soluble proteins dedicated to the folding and modification of newly synthesized secretory proteins. They include the chaperone BiP and the enzyme protein disulfide isomerase, which are necessary for the ER to carry out its functions. Although such ER-resident luminal proteins are not specifically selected by COPII vesicles, their sheer abundance causes them to be continuously loaded passively into vesicles destined for the *cis*-Golgi. The transport of these soluble proteins back to the ER, mediated by COPI vesicles, prevents their eventual depletion.

Most soluble ER-resident proteins carry a Lys-Asp-Glu-Leu (KDEL in the one-letter code) sequence at their C-terminus (see Table 14-2). Several experiments have demonstrated that this *KDEL sorting signal* is both necessary and sufficient to cause a protein bearing this sequence to be located in the ER. For instance, when a mutant protein disulfide isomerase lacking these four residues is

synthesized in cultured fibroblasts, the protein is secreted. Moreover, if a protein that is normally secreted is altered so that it contains the KDEL sorting signal at its C-terminus, the protein is located in the ER. The KDEL sorting signal is recognized and bound by the *KDEL receptor*, a transmembrane protein found primarily on small transport vesicles shuttling between the ER and the *cis*-Golgi and on the *cis*-Golgi membrane. In addition, soluble ER-resident proteins that carry the KDEL signal have oligosaccharide chains bearing modifications that are catalyzed by enzymes found only in the *cis*-Golgi or *cis*-Golgi network; thus at some time these proteins must have left the ER and been transported at least as far as the *cis*-Golgi network. These findings indicate that the KDEL receptor acts mainly to retrieve soluble proteins containing the KDEL sorting signal that have escaped to the *cis*-Golgi network and return them to the ER (Figure 14-14). The KDEL receptor binds more tightly to its ligand at low pH, and it is thought that the receptor is able to bind KDEL peptides in the *cis*-Golgi and release these peptides in the ER because the pH of the Golgi is slightly lower than that of the ER.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-14 Role of the KDEL receptor in retrieval of ER-resident luminal proteins from the Golgi. ER luminal proteins, especially those present at high concentrations, can be passively incorporated into COPII vesicles and transported to the Golgi (steps 1 and 2). Many such proteins bear a C-terminal KDEL (Lys-Asp-Glu-Leu) sequence (red) that allows them to be retrieved. The KDEL receptor, located mainly in the *cis*-Golgi network and in both COPII and COPI vesicles, binds proteins bearing the KDEL sorting signal and returns them to the ER (steps 3 and 4). This retrieval system prevents

depletion of ER luminal proteins such as those needed for proper folding of newly made secretory proteins. The binding affinity of the KDEL receptor is very sensitive to pH. The small difference between the pH of the ER and that of the Golgi favors binding of KDEL-bearing proteins to the receptor in Golgi-derived vesicles and their release in the ER. See J. Semenza et al., 1990, *Cell* **61**:1349.

The KDEL receptor and other membrane proteins that are transported back to the ER from the Golgi contain a Lys-Lys-X-X sequence at the very end of their C-terminal segment, which faces the cytosol (see Table 14-2). This *KKXX sorting signal*, which binds to a complex of the COPI α and β subunits (two of the seven polypeptide subunits in the COPI coatomer), is both necessary and sufficient to incorporate membrane proteins into COPI vesicles for retrograde transport to the ER. Temperature-sensitive yeast mutants lacking COPIα or COPIβ are not only unable to bind the KKXX signal, but are also unable to transport proteins bearing this signal back to the ER, indicating that COPI vesicles mediate retrograde Golgi-to-ER transport.

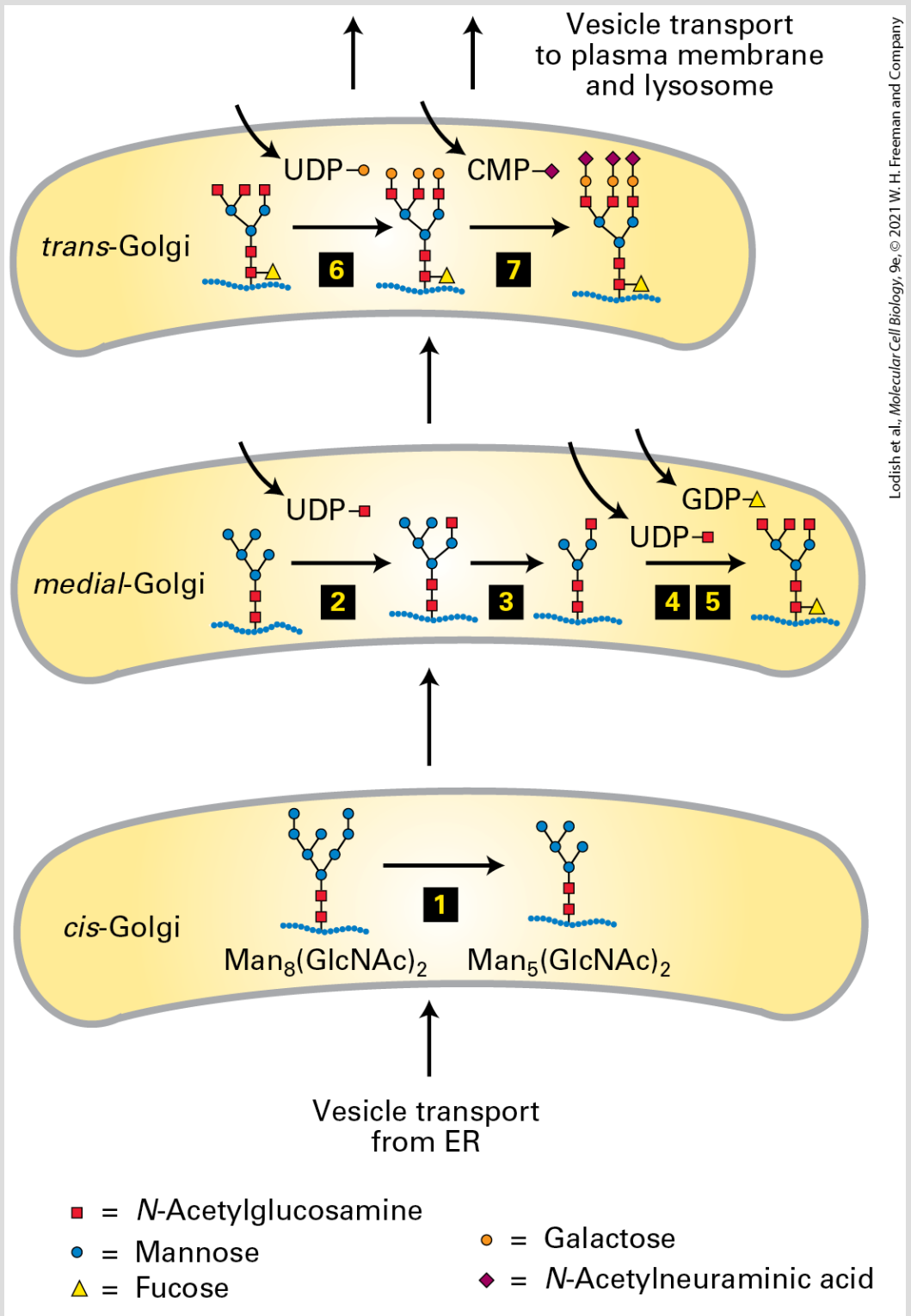
A second sorting signal that targets proteins to COPI vesicles and thus enables recycling from the Golgi to the ER is a di-arginine sequence. Unlike the KKXX sorting signal, which must be located at the cytoplasmically oriented C-terminus of a protein, the di-arginine sorting signal can reside in any segment of a membrane protein that is on the cytoplasmic face of the membrane.

The partitioning of proteins between the ER and Golgi complex is a highly dynamic process depending on both COPII (anterograde) and COPI (retrograde) vesicles, with each type of vesicle responsible for recycling the components necessary for the function of the other type of vesicle. The organization of this partitioning process raises an interesting puzzle: How do vesicles preferentially use the v-SNAREs that specify fusion with the correct target membrane instead of the v-SNAREs that are being recycled and specify fusion with the donor membrane?

This basic question concerning correct membrane partitioning has recently been answered for COPII vesicles. After these vesicles form, the COPII coat proteins remain assembled long enough for the Sec23/Sec24 complex to interact with a specific tethering factor attached to the *cis*-Golgi membrane. Vesicle uncoating to expose the v-SNAREs is completed only after the COPII vesicle is already closely associated with the *cis*-Golgi membrane and the COPII v-SNAREs are in position to form complexes with their cognate t-SNAREs. Although COPII vesicles also carry COPI-specific v-SNARE proteins, which are being recycled back to the *cis*-Golgi, these COPI v-SNARE proteins never have the opportunity to form SNARE complexes with cognate ER-localized t-SNARE proteins.

Anterograde Transport Through the Golgi Occurs by Cisternal Maturation

The Golgi complex is organized into three compartments, often arranged in a stacked set of flattened sacs, called *cisternae*. The compartments of the Golgi differ from one another according to the enzymes they contain. Many of the enzymes are glycosidases and glycosyltransferases that are involved in modifying the *N*-linked or *O*-linked carbohydrates attached to secretory proteins as they transit through the Golgi complex. On the whole, the Golgi complex operates much like an assembly line, where each compartment serves as an assembly station in which the carbohydrate chains modified in the previous compartment are substrates for the modifying enzymes of the next compartment. (Figure 14-15 shows a representative sequence of modification steps.)



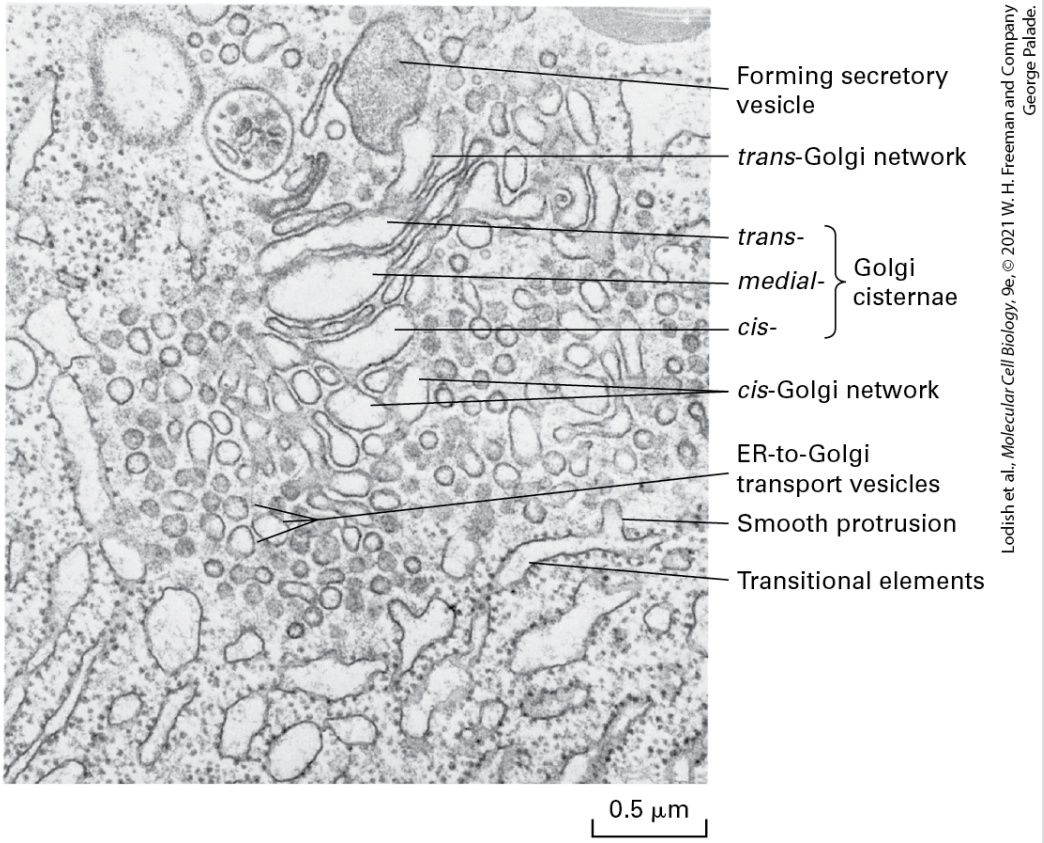
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-15 Processing of *N*-linked oligosaccharide chains on glycoproteins within *cis*-, *medial*-, and *trans*-Golgi cisternae in vertebrate cells. After removal of three mannose residues in the *cis*-Golgi (step 1), the protein

moves by cisternal maturation to the *medial*-Golgi. Here three *N*-acetylglucosamine (GlcNAc) residues are added (steps 2 and 4), two more mannose residues are removed (step 3), and a single fucose is added (step 5). Processing is completed in the *trans*-Golgi by addition of three galactose residues (step 6) and finally by linkage of an *N*-acetylneuraminic acid residue to each of the galactose residues (step 7). Specific transferase enzymes add sugars to the oligosaccharide, one at a time, from sugar nucleotide precursors imported from the cytosol. The enzymes catalyzing each step are localized to the indicated compartments. This pathway represents the Golgi processing events for a typical mammalian glycoprotein. Variations in the structure of *N*-linked oligosaccharides can result from differences in processing steps in the Golgi. See R. Kornfeld and S. Kornfeld, 1985, *Annu. Rev. Biochem.* **45**:631.

For many years, it was thought that the Golgi complex was an essentially static set of compartments with small transport vesicles carrying secretory proteins forward, from the *cis*- to the *medial*-Golgi and from the *medial*- to the *trans*-Golgi. Indeed, electron microscopy reveals many small vesicles associated with the Golgi complex that appear to move proteins from one Golgi compartment to another (Figure 14-16). However, electron microscopy only provides an image of structures that is frozen in time and cannot show the direction of movement of the transport vesicles. We now know that the small vesicles in the vicinity of the Golgi complex are actually moving in the retrograde direction, retrieving ER or Golgi enzymes from a later compartment and transporting them to an earlier compartment in the secretory pathway. Thus the Golgi appears to have a highly dynamic organization, continually forming transport vesicles, though only in the retrograde direction. To see the effect of this retrograde transport on the organization of the Golgi, consider

the net effect on the *medial*-Golgi compartment as enzymes from the *trans*-Golgi move to the *medial*-Golgi while enzymes from the *medial*-Golgi are transported to the *cis*-Golgi. As this process continues, the *medial*-Golgi acquires enzymes from the *trans*-Golgi while losing *medial*-Golgi enzymes to the *cis*-Golgi and thus progressively becomes a new *trans*-Golgi compartment. In this way, secretory cargo proteins acquire carbohydrate modifications in the proper sequential order without being moved from one cisterna to another via anterograde vesicle transport.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
George Palade.

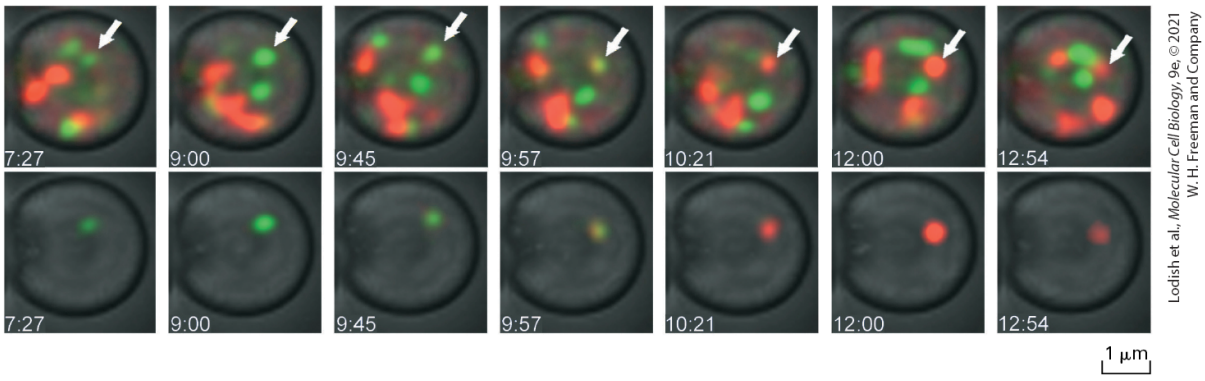
EXPERIMENTAL FIGURE 14-16 Anterograde-moving COPII transport vesicles and retrograde-moving COPI transport vesicles are difficult to distinguish in an electron micrograph of the Golgi complex in a pancreatic acinar cell.

Elements of the rough ER are on the bottom and left in this micrograph. Adjacent to the rough ER are transitional elements from which smooth protrusions appear to be budding. These buds form the small COPII vesicles that transport secretory proteins in the anterograde direction from the rough ER to the Golgi complex. Interspersed among the Golgi cisternae are other small vesicles that are similar in appearance to the COPII vesicles, but these transport vesicles, which bear COPI coats, are now known to carry resident Golgi enzymes between Golgi compartments in the retrograde direction. A large anterograde secretory vesicle can be seen forming from the *trans*-Golgi network.

The first evidence that the forward transport of cargo proteins from the *cis*- to the *trans*-Golgi occurs by this progressive mechanism of cisternal maturation came from careful microscopic analysis of the synthesis of large macromolecular complexes that follow the secretory pathway to form scales on the surface of certain types of unicellular algae. These algal scales are assembled from glycoproteins in the *cis*-Golgi into large dense complexes visible in the electron microscope. Like other secretory proteins, newly made scales move from the *cis*- to the *trans*-Golgi, but they were never observed in the transport vesicles that bud from Golgi cisternae. In further experiments, the observation that the movement of large cargo molecules through the Golgi can occur without entry of cargo into transport vesicles was shown also to be true of collagen synthesized in animal cells. In fibroblasts, large aggregates of the procollagen precursor often form in the lumen of the *cis*-Golgi. The procollagen aggregates are too large to be incorporated into small transport vesicles, and investigators could never find such

aggregates in transport vesicles. These observations show that the forward movement of these, and perhaps all, secretory proteins from one Golgi compartment to another does *not* occur via small vesicles.

A particularly elegant demonstration of cisternal maturation in yeast takes advantage of different-colored fluorescent labels to image two different Golgi proteins simultaneously. Figure 14-17 shows how a *cis*-Golgi resident protein labeled with a green fluorescent protein and a *trans*-Golgi resident protein labeled with a red fluorescent protein behave in the same yeast cell. At any given moment, individual Golgi cisternae appear to have a distinct compartmental identity, in the sense that they contain either the *cis*-Golgi protein or the *trans*-Golgi protein, but only rarely contain both proteins. However, over time, an individual cisterna labeled with the *cis*-Golgi protein can be seen to progressively lose this protein and acquire the *trans*-Golgi protein. This behavior is exactly that predicted by the cisternal maturation model, in which the composition of an individual cisterna changes as Golgi resident proteins move from later to earlier Golgi compartments.



EXPERIMENTAL FIGURE 14-17 Fluorescence-tagged fusion proteins demonstrate Golgi cisternal maturation in a live yeast cell. Yeast cells expressing the early Golgi protein Vrg4 fused to GFP (green fluorescence) and the late Golgi protein Sec7 fused to DsRed (red fluorescence) were imaged by time-lapse microscopy. The top series of images, taken approximately 1 minute apart, shows a collection of Golgi cisternae, which at any one time are labeled with either Vrg4 or Sec7. The bottom series of images are digitally processed to show just the Golgi cisterna, identified by the white arrows in the top set of images. At first, only Vrg4-GFP is located in the isolated cisterna, and later only Sec7-DsRed is located in the isolated cisterna, following a brief period in which both proteins are co-localized in this compartment. This experiment is a direct demonstration of the cisternal maturation hypothesis, showing that the composition of individual cisternae follows a process of maturation characterized by loss of early Golgi proteins and gain of late Golgi proteins.

[Republished with permission from *Nature*, from E. Losev et al., 2006, "Golgi Maturation Visualized in Living Yeast," *Nature* **441**:1002–1006; permission conveyed through Copyright Clearance Center, Inc.]

Although most protein traffic moves through the Golgi complex by a cisternal maturation mechanism, there is evidence that at least some cargo proteins can be detected in COPI transport vesicles that bud from Golgi membranes. This observation suggests that there may be a subset of COPI transport vesicles that move cargo in an anterograde direction.

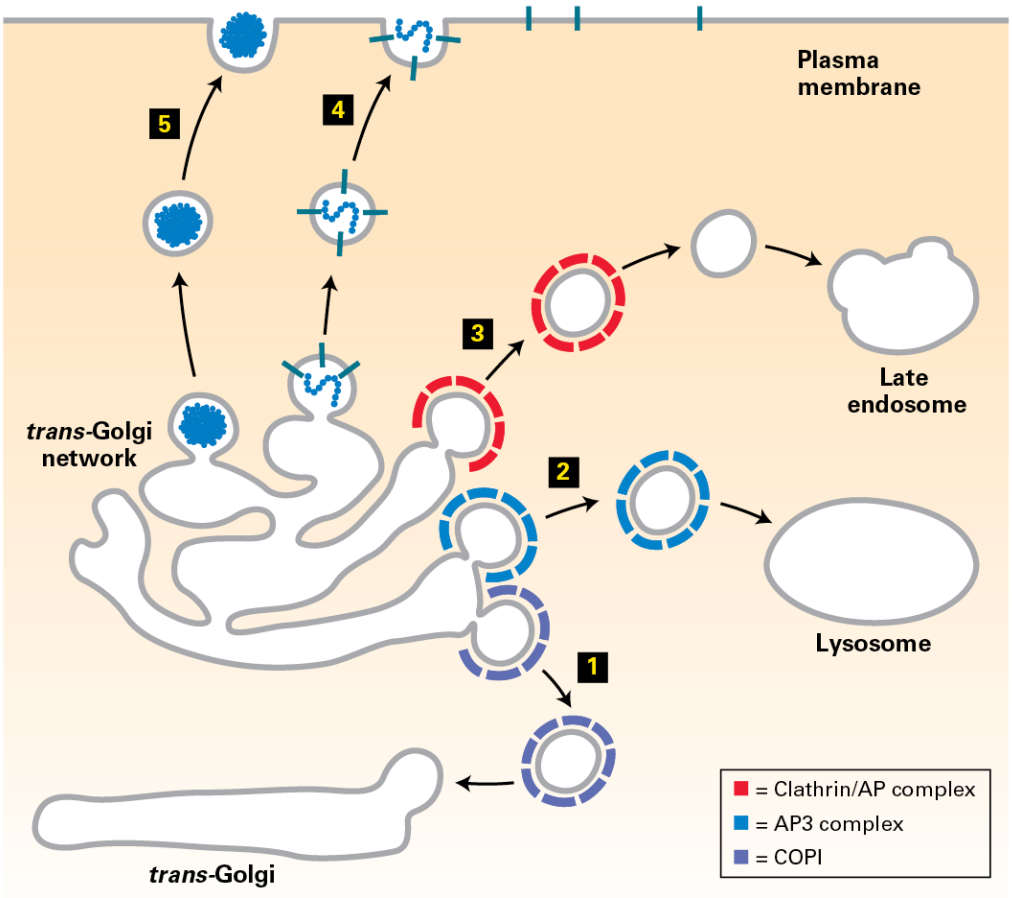
KEY CONCEPTS OF SECTION 14.3

Early Stages of the Secretory Pathway

- COPII vesicles transport proteins from the rough ER to the *cis*-Golgi; COPI vesicles transport proteins in the reverse direction (see Figure 14-12).
- COPII coats comprise three major components: the small GTP-binding protein Sar1, a Sec23/Sec24 complex, and a Sec13/Sec31 complex.
- Components of the COPII coat bind to membrane cargo proteins containing a di-acidic or other sorting signal in their cytosolic regions (see Figure 14-13). Soluble cargo proteins are probably targeted to COPII vesicles by binding to a membrane protein receptor.
- Many soluble ER-resident proteins contain a KDEL sorting signal. Binding of this retrieval sequence to a specific receptor protein in the *cis*-Golgi membrane recruits missorted ER proteins into retrograde COPI vesicles (see Figure 14-14).
- Membrane proteins needed to form COPII vesicles can be retrieved from the *cis*-Golgi by COPI vesicles. One of the sorting signals that directs membrane proteins into COPI vesicles is a KKXX sequence, which binds to subunits of the COPI coat. A distinct di-arginine sorting signal operates by a similar mechanism.
- COPI vesicles also carry Golgi-resident proteins from later to earlier compartments in the Golgi complex.
- Soluble and membrane proteins advance through the Golgi complex by cisternal maturation, a process of anterograde transport that depends on resident Golgi enzymes moving by COPI vesicular transport in a retrograde direction.

14.4 Later Stages of the Secretory Pathway

As cargo proteins move from the *cis*- to the *trans*-Golgi by cisternal maturation, modifications to their oligosaccharide chains are carried out by Golgi-resident enzymes. The retrograde trafficking of COPI vesicles from later to earlier Golgi compartments maintains sufficient levels of these carbohydrate-modifying enzymes in the appropriate compartments. Eventually, properly processed cargo proteins reach the *trans*-Golgi network, the most distal Golgi compartment. Here they are sorted into a number of different kinds of vesicles for delivery to their final destination. Each of the target destinations, such as the plasma membrane, endosomes, and lysosomes, has a unique composition of lipids and membrane proteins, and it is primarily the sorting in the *trans*-Golgi network that gives each of these organelles its unique identity. In this section, we discuss the different kinds of vesicles that bud from the *trans*-Golgi network, the mechanisms that segregate cargo proteins among them, and key processing events that occur late in the secretory pathway. The various types of vesicles that bud from the *trans*-Golgi and will be described in this section are summarized in Figure 14-18.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-18 Vesicle-mediated protein trafficking from the *trans*-Golgi network. COPI (purple) vesicles mediate retrograde transport within the Golgi (1). Proteins that function in the lumen or in the membrane of the lysosome are first transported from the *trans*-Golgi network via AP complex and clathrin-coated (red) vesicles (2); after uncoating, these vesicles fuse with late endosomes, which deliver their contents to the lysosome. Some vesicles from the *trans*-Golgi carrying cargo destined for the lysosome fuse with the lysosome directly (3), bypassing the late endosome. These vesicles are coated with AP3 complex (blue); it is unknown whether these vesicles also contain clathrin. The coat proteins surrounding constitutive (4) and regulated (5) secretory vesicles have not yet been characterized; these vesicles carry secreted proteins and plasma-membrane proteins from the *trans*-Golgi network to the cell surface.

Vesicles Coated with Clathrin and Adapter Proteins Mediate Transport from the *trans*-Golgi

The best characterized vesicles that bud from the *trans*-Golgi network have a two-layered coat: an outer layer composed of the fibrous protein clathrin and an inner layer composed of **adapter protein (AP) complexes**. Purified clathrin molecules, which have a three-limbed shape, are called *triskelions*, from the Greek for "three-legged" (Figure 14-19a). Each limb contains one clathrin heavy chain (180 kDa) and one clathrin light chain (~35–40 kDa). Triskelions polymerize to form a polygonal lattice with an intrinsic curvature (Figure 14-19b). When clathrin polymerizes on a donor membrane, it does so in association with AP complexes, which fill the space between the clathrin lattice and the membrane. Each AP complex (340 kDa) contains one copy each of four different adapter subunit proteins. A specific association between the globular domain at the end of each clathrin heavy chain in a triskelion and one subunit of the AP complex both promotes the co-assembly of clathrin triskelions with AP complexes and adds to the stability of the completed vesicle coat.

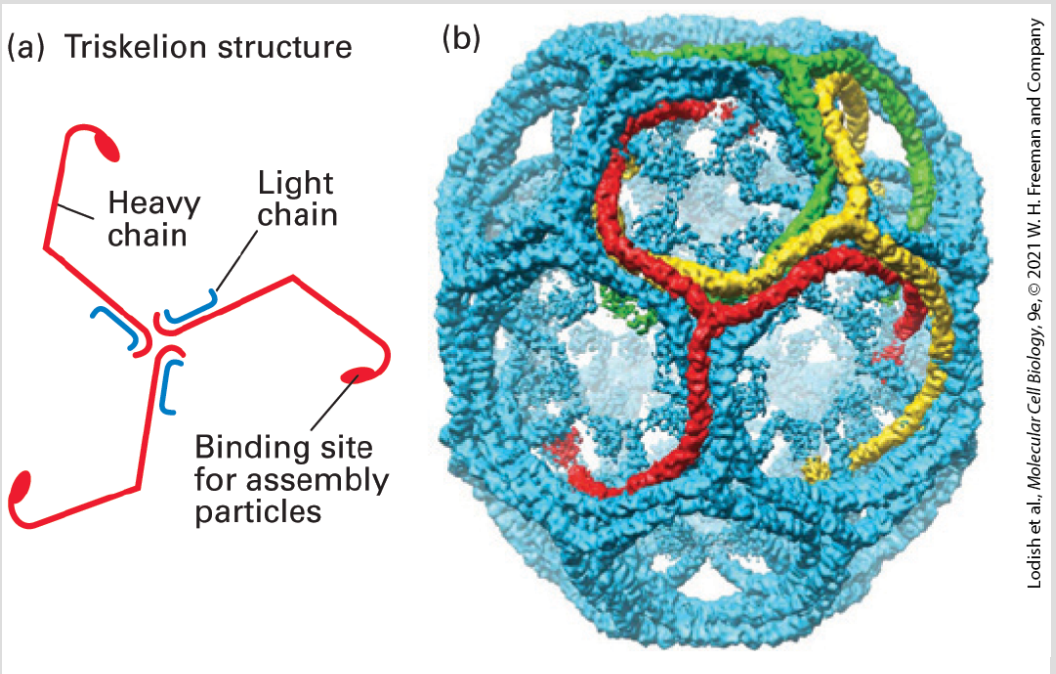


FIGURE 14-19 Structure of clathrin coats. (a) A clathrin molecule, called a *triskelion*, is composed of three heavy and three light chains. It has an intrinsic curvature due to the bend in the heavy chains. (b) Clathrin coats were formed in vitro by mixing purified clathrin heavy and light chains with AP2 complexes in the absence of membranes. Cryoelectron micrographs of more than 1000 assembled hexagonal clathrin barrel particles were analyzed by digital image processing to generate an average structural representation. The processed image shows only the clathrin heavy chains in a structure composed of 36 triskelions. Three representative triskelions are highlighted in red, yellow, and green. Some of the AP2 complexes packed into the interior of the clathrin cage are also visible in this representation. See B. Pishvaee and G. Payne, 1998, *Cell* **95**:443.

[Part (b) Republished with permission from *Nature*, from A. Fotin et al., 2004, "Molecular Model for a Complete Clathrin Lattice from Electron Cryomicroscopy," *Nature* **432**:573–579; permission conveyed through Copyright Clearance Center, Inc.]

Adapter proteins determine which cargo proteins are specifically included in a budding transport vesicle by binding to the cytosolic

face of membrane proteins. Thus each of the multiple adapter protein complexes that operate in the late secretory pathway can be classified based on the types of cargo proteins that they select as well as the specific transport step that they mediate (see Table 14-1). Three different AP complexes are known (AP1, AP2, and AP3), each with four subunits of different, though related, proteins. A second general type of adapter protein, known as *GGA*, consists of a single 70 kDa polypeptide. This monomeric adapter protein has been shown to contain both clathrin- and cargo-binding elements similar to those found in the much larger hetero-tetrameric AP complexes. All vesicles whose coats contain one of these complexes use ARF to initiate coat assembly on the donor membrane. As discussed previously, ARF also initiates assembly of COPI coats. The additional membrane features or protein factors that determine which type of coat will assemble after ARF attachment are not well understood at this time.

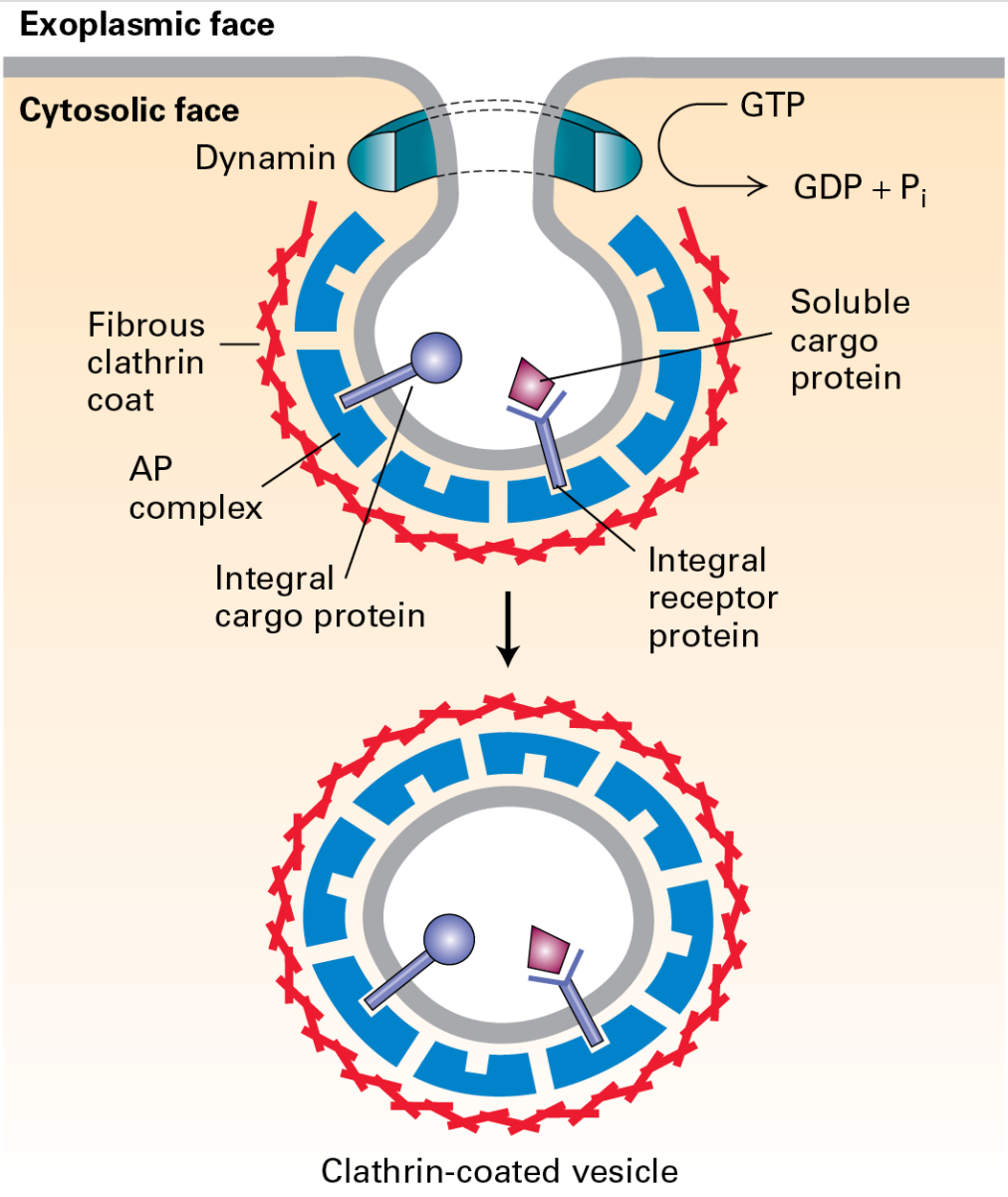
Vesicles that bud from the *trans*-Golgi network en route to the lysosome by way of the late endosome (see Figure 14-18, **3**) have clathrin coats associated with either AP1 or GGA. One of the AP1 subunits in the vesicle coat binds to membrane proteins containing a Tyr-X-X-Φ sequence. Proteins bearing this *YXXΦ sorting signal* are recruited into clathrin/AP1-coated vesicles budding from the *trans*-Golgi network. As we discuss in the next section, vesicles with clathrin/AP2 coats, which bud from the plasma membrane during endocytosis, can also recognize the YXXΦ sorting signal. Vesicles

coated with GGA proteins and clathrin bind cargo molecules with different kinds of sorting sequences including Asp-X-Leu-Leu and Asp-Phe-Gly-X-Φ sequences.

Some vesicles that bud from the *trans*-Golgi network have coats composed of the AP3 complex. Although the AP3 complex does contain a binding site for clathrin similar to those in the AP1 and AP2 complexes, it is not clear whether clathrin is necessary for the functioning of AP3-containing vesicles because mutant versions of AP3 that lack the clathrin binding site appear to be fully functional. AP3-coated vesicles mediate trafficking to the lysosome, but they appear to bypass the late endosome and fuse directly with the lysosomal membrane (see Figure 14-18, **2**). In certain types of cells, such AP3 vesicles mediate protein transport to specialized storage compartments related to the lysosome. For example, AP3 is required for delivery of proteins to melanosomes, which contain the black pigment melanin in skin cells, and to platelet storage vesicles in megakaryocytes, large cells that fragment into dozens of platelets. Mice with mutations in either of two different subunits of AP3 not only have abnormal skin pigmentation but also exhibit bleeding disorders. The latter occur because platelets require normal storage vesicles in order to repair tears in blood vessels.

Dynamin Is Required for Pinching Off of Clathrin-Coated Vesicles

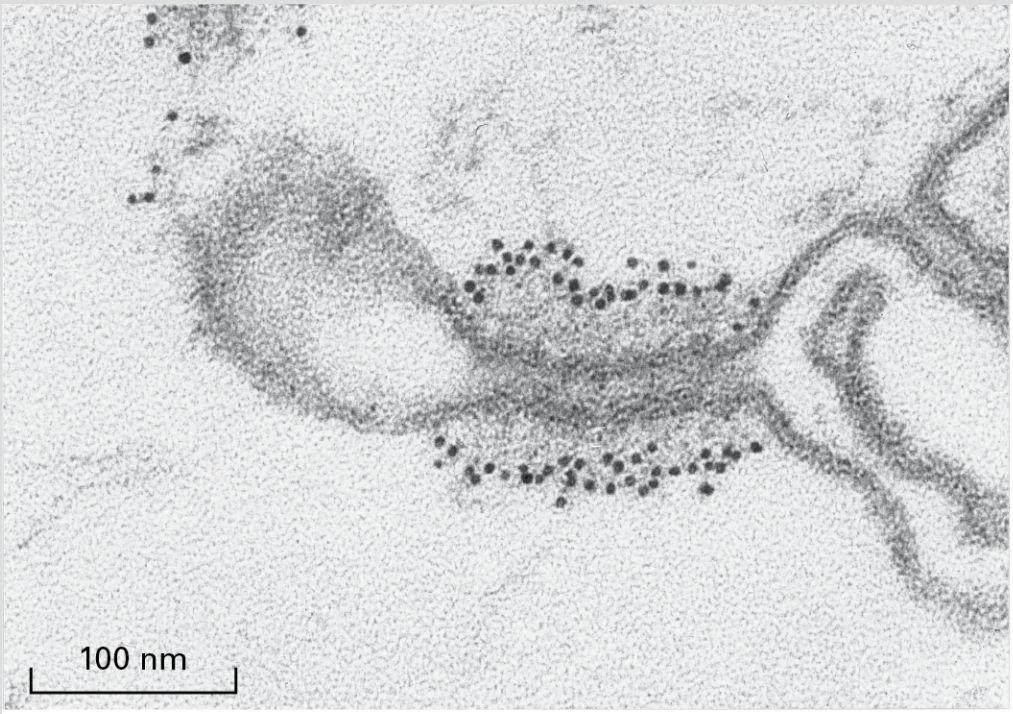
A fundamental step in the formation of a transport vesicle that we have not yet considered is how a vesicle bud is pinched off from the donor membrane. In the case of clathrin-coated vesicles, a cytosolic protein called **dynamin** is essential for the release of complete vesicles. At the later stages of bud formation, dynamin polymerizes around the neck portion of the bud and then hydrolyzes GTP. The energy derived from GTP hydrolysis is thought to drive a conformational change in dynamin that stretches the neck until the vesicle pinches off (Figure 14-20).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-20 Model for dynamin-mediated pinching off of clathrin-coated vesicles. After a vesicle bud forms, dynamin polymerizes over the neck. By a mechanism that is not well understood, dynamin-catalyzed hydrolysis of GTP leads to release of the vesicle from the donor membrane. Note that membrane proteins in the donor membrane are incorporated into vesicles by interacting with AP complexes in the coat. See K. Takel et al., 1995, *Nature* **374**:186.

Incubation of cell extracts with a nonhydrolyzable derivative of GTP provides dramatic evidence for the importance of dynamin in the pinching off of clathrin/AP2-coated vesicles during endocytosis. Such treatment leads to accumulation of clathrin-coated vesicle buds with excessively long necks that are surrounded by polymeric dynamin but do not pinch off (Figure 14-21). Likewise, cells expressing mutant forms of dynamin that cannot bind GTP do not form clathrin-coated vesicles and instead accumulate similar long-necked vesicle buds encased with polymerized dynamin.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 14-21 GTP hydrolysis by dynamin is required for the pinching off of clathrin-coated vesicles in cell-free extracts. A preparation of nerve terminals, which undergo extensive endocytosis, was lysed by treatment with distilled water and incubated with GTP-γ-S, a nonhydrolyzable derivative of GTP. After sectioning, the preparation was treated with gold-tagged anti-dynamin antibody and viewed in the electron microscope. This image, which shows a long-necked clathrin/AP-coated bud with polymerized dynamin lining

the neck, reveals that buds can form in the absence of GTP hydrolysis, but vesicles cannot pinch off. The extensive polymerization of dynamin that occurs in the presence of GTP-γ-S probably does not occur during the normal budding process.

[Republished by permission from *Nature*, from K. Takei et al., 1995, "Tubular Membrane Invaginations Coated by Dynamin Rings Are Induced by GTP-Gamma S in Nerve Terminals," *Nature* **374**:(6518):186–190; permission conveyed through Copyright Clearance Center, Inc.]

Like COPI and COPII vesicles, clathrin-coated vesicles normally lose their coat soon after their formation. Cytosolic Hsp70, a constitutive chaperone protein found in all eukaryotic cells, is thought to use energy derived from the hydrolysis of ATP to drive depolymerization of the clathrin coat into triskelions. In the case of endocytic vesicles, uncoating not only releases triskelions for reuse in the formation of additional vesicles but also exposes v-SNAREs for use in fusion with target membranes. Vesicle uncoating by cytosolic Hsp70 appears to be activated by a co-chaperone, auxillin, that contains a domain that stimulates the ATP hydrolysis by Hsp70. Conformational changes that occur when ARF switches from the GTP-bound to the GDP-bound state are thought to regulate the timing of clathrin coat depolymerization, but how the action of Hsp70 and auxillin is coupled to ARF switching is not well understood.

Mannose 6-Phosphate Residues Target Resident Enzymes to Lysosomes

As we have seen, many of the sorting signals that direct cargo-protein trafficking in the secretory pathway are short amino acid sequences in the targeted protein. In contrast, the sorting signal that directs soluble lysosomal enzymes from the *trans*-Golgi network to the late endosome is a carbohydrate residue, **mannose 6-phosphate (M6P)**, which is formed in the *cis*-Golgi. The addition and initial processing of one or more preformed *N*-linked oligosaccharide precursors in the rough ER is the same for lysosomal enzymes as for membrane and secreted proteins, yielding core $Man_8(GlcNAc)_2$ chains (see Figure 13-18). In the *cis*-Golgi, the *N*-linked oligosaccharides present on most lysosomal enzymes undergo a two-step reaction sequence that generates M6P residues (Figure 14-22). The addition of M6P residues to the oligosaccharide chains of soluble lysosomal enzymes prevents these proteins from undergoing the further processing reactions characteristic of secreted and membrane proteins (see Figure 14-15).

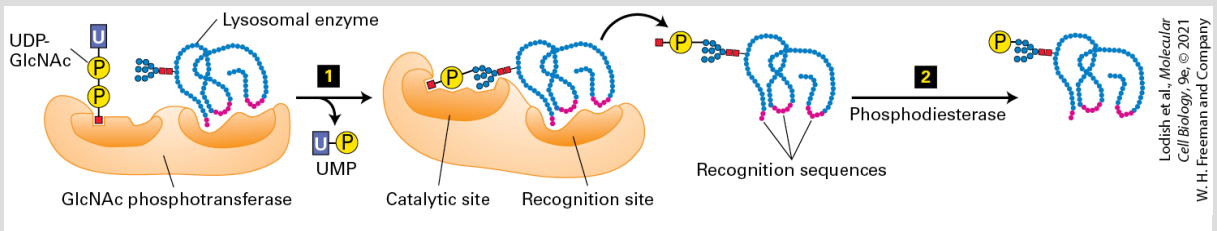


FIGURE 14-22 Formation of mannose 6-phosphate (M6P) residues that target soluble enzymes to lysosomes. The M6P residues that direct proteins to lysosomes are generated in the *cis*-Golgi by two Golgi-resident enzymes. Step **1**: An *N*-acetylglucosamine (GlcNAc) phosphotransferase transfers a phosphorylated GlcNAc group to carbon atom 6 of one or more mannose residues. Because only lysosomal enzymes contain sequences (red) that are recognized and bound by this enzyme, phosphorylated GlcNAc groups are added specifically to lysosomal enzymes. Step **2**: After release of the modified protein from the

phosphotransferase, a phosphodiesterase removes the GlcNAc group, leaving a phosphorylated mannose residue on the lysosomal enzyme. See A. B. Cantor et al., 1992, *J. Biol. Chem.* **267**:23349; and S. Kornfeld, 1987, *FASEB J.* **1**:462.

As shown in Figure 14-23, the segregation of M6P-bearing lysosomal enzymes from secreted and membrane proteins occurs in the *trans*-Golgi network. Here transmembrane *mannose 6-phosphate receptors* bind the M6P residues on lysosome-destined proteins very tightly and specifically. Clathrin/AP1-coated vesicles containing the M6P receptor and bound lysosomal enzymes then bud from the *trans*-Golgi network, lose their coats, and subsequently fuse with a late endosome by mechanisms described previously (Figure 14-18). Because M6P receptors can bind M6P at the slightly acidic pH (~ 6.5) of the *trans*-Golgi network, but not at a pH of less than 6, the bound lysosomal enzymes are released within late endosomes, which have an internal pH of 5.0–5.5. Furthermore, a phosphatase within late endosomes usually removes the phosphate from M6P residues on lysosomal enzymes, preventing any rebinding to the M6P receptor that might occur in spite of the low pH there. Vesicles budding from late endosomes, coated with a protein complex known as the **retromer**, recycle the M6P receptor back to the *trans*-Golgi network. Eventually, mature late endosomes fuse with lysosomes, delivering the lysosomal enzymes to their final destination.

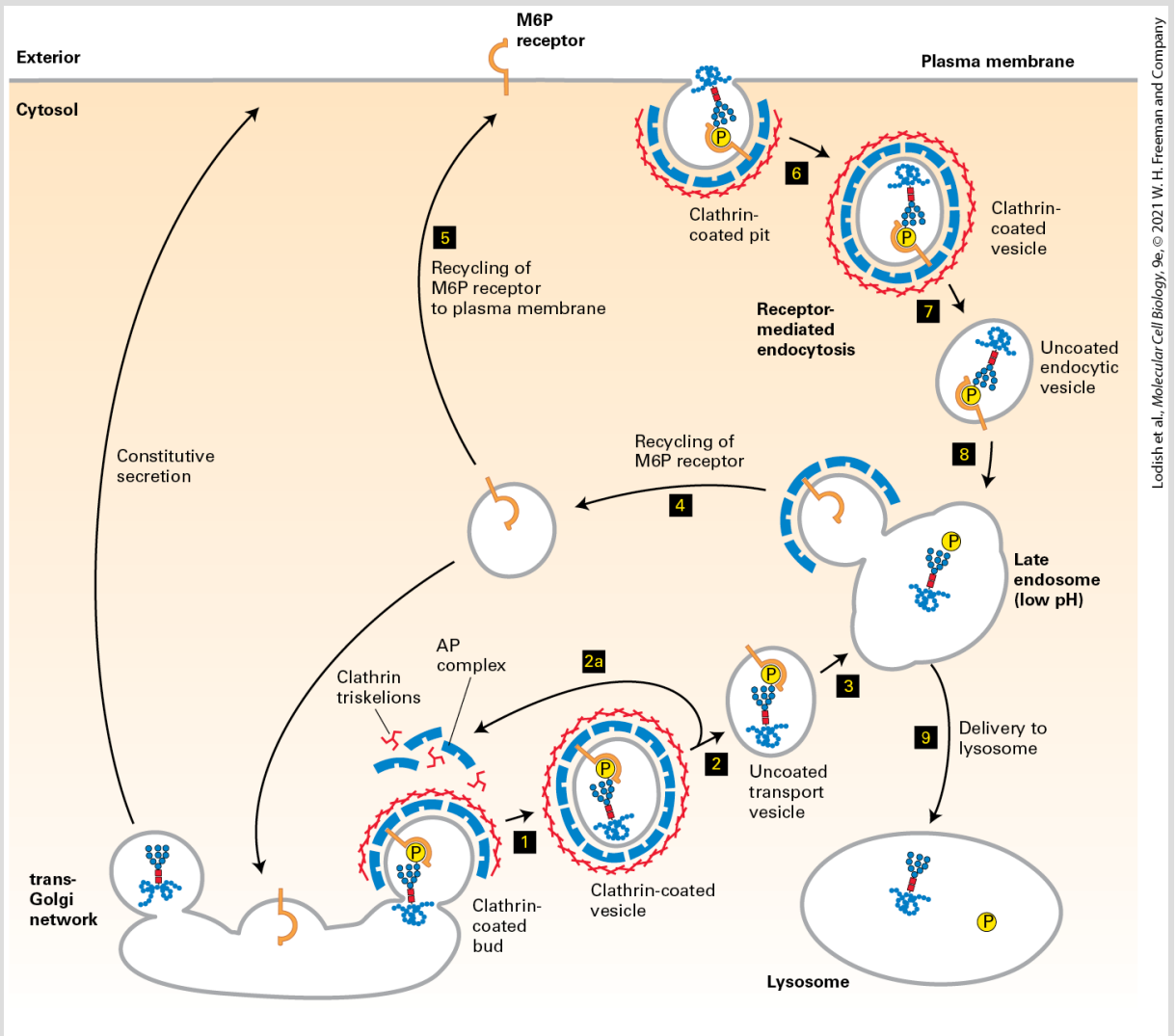


FIGURE 14-23 Trafficking of soluble lysosomal enzymes from the *trans*-Golgi network and cell surface to lysosomes. Newly synthesized lysosomal enzymes, produced in the ER, acquire mannose 6-phosphate (M6P) residues in the *cis*-Golgi (see Figure 14-22). For simplicity, only one phosphorylated oligosaccharide chain is depicted, although lysosomal enzymes typically have many such chains. In the *trans*-Golgi network, proteins that bear the M6P sorting signal interact with M6P receptors in the membrane and thereby are directed into clathrin/AP1-coated vesicles (step **1**). The coat surrounding released vesicles is rapidly depolymerized (step **2**), and the uncoated transport vesicles fuse with late endosomes (step **3**). Note that disassembled coat proteins are recycled for additional rounds of vesicle assembly (step **2a**). After the phosphorylated enzymes dissociate from the M6P receptors in late endosomes, free M6P receptors are recycled by incorporation into vesicles that bud from the late endosome (step **4**). Vesicles carrying free M6P receptors can return to the *trans*-Golgi or fuse with the plasma membrane delivering M6P receptors to

the cell surface (step 5). Phosphorylated lysosomal enzymes are also occasionally sorted from the *trans*-Golgi to the cell surface and secreted. These secreted enzymes can be retrieved by receptor-mediated endocytosis (steps 6–8). Ultimately, fusion of the late endosome with the lysosome delivers phosphorylated lysosomal enzymes to the interior of the lysosomes (step 9). See G. Griffiths et al., 1988, *Cell* **52**:329; S. Kornfeld, 1992, *Annu. Rev. Biochem.* **61**:307; and G. Griffiths and J. Gruenberg, 1991, *Trends Cell Biol.* **1**:5.

The sorting of soluble lysosomal enzymes in the *trans*-Golgi network (Figure 14-23, steps **1**–**4**) shares many features with the trafficking of proteins between the ER and *cis*-Golgi compartments mediated by COPII and COPI vesicles. First, M6P acts as a sorting signal by interacting with the luminal domain of a receptor protein in the donor membrane. Second, the membrane-embedded receptors with their bound ligands are incorporated into the appropriate vesicles — in this case, either GGA- or AP1-containing clathrin-coated vesicles — by interacting with the vesicle coat. Third, these transport vesicles fuse with only one specific organelle, here the late endosome, as the result of interactions between specific v-SNAREs and t-SNAREs. And finally, intracellular transport receptors dissociated from their bound ligand are recycled by retrograde vesicle trafficking.

Study of Lysosomal Storage Diseases Revealed Key Components of the Lysosomal Sorting Pathway

A group of genetic disorders termed *lysosomal storage diseases* are caused by the absence of one or more lysosomal enzymes. As a result, undigested glycolipids and extracellular components that would normally be degraded by lysosomal enzymes accumulate in lysosomes as large inclusions. Patients with lysosomal storage diseases can have a variety of developmental, physiological, and neurological abnormalities depending on the type and severity of the storage defect. *I-cell disease* is a particularly severe type of lysosomal storage disease in which multiple enzymes are missing from the lysosomes. Cells from affected individuals lack the *N*-acetylglucosamine phosphotransferase that is required for formation of M6P residues on lysosomal enzymes in the *cis*-Golgi (see Figure 14-22). Biochemical comparison of lysosomal enzymes from normal individuals with those from patients with I-cell disease led to the initial discovery of M6P as the lysosomal sorting signal. Lacking this signal, the lysosomal enzymes of affected individuals are secreted rather than being sorted to and sequestered in lysosomes.

When fibroblasts from patients with I-cell disease are grown in a medium containing lysosomal enzymes bearing M6P residues, the diseased cells acquire a nearly normal intracellular content of lysosomal enzymes. This finding indicates that the plasma membrane of these cells contains M6P receptors, which can internalize extracellular phosphorylated lysosomal enzymes by receptor-mediated endocytosis. This process, used by many cell-

surface receptors to bring bound proteins or particles into the cell, is discussed in detail in the next section. It is now known that even in normal cells, some M6P receptors are transported to the plasma membrane and some phosphorylated lysosomal enzymes are secreted (see Figure 14-23). The secreted enzymes can be retrieved by receptor-mediated endocytosis and directed to lysosomes. This pathway thus scavenges any lysosomal enzymes that escape the usual M6P sorting pathway.

Hepatocytes from patients with I-cell disease contain a normal complement of lysosomal enzymes and no inclusions, even in the absence of an exogeneous source of lysosomal enzymes bearing M6P residues. This finding implies that hepatocytes (the most abundant type of liver cell) employ a pathway for sorting lysosomal enzymes that does not rely on a targeting signal based on M6P. This pathway probably relies on some yet unidentified protein-based sorting signal present in lysosomal enzymes. ■

Protein Aggregation in the *trans*-Golgi May Function in Sorting Proteins to Regulated Secretory Vesicles

As noted in the chapter introduction, all eukaryotic cells continuously secrete certain proteins (constitutive secretion). Specialized secretory cells store other proteins in vesicles and

secrete them only when triggered by a specific stimulus. One example of such regulated secretion occurs in pancreatic β cells, which store newly made insulin in specialized secretory vesicles and secrete it in response to an elevation in blood glucose (see Figure 21-1b). These and other secretory cells simultaneously use two different types of secretory vesicles to move proteins from the *trans*-Golgi network to the cell surface: unregulated transport vesicles (also called *constitutive secretory vesicles*) and regulated transport vesicles.

A common mechanism appears to sort regulated proteins as diverse as ACTH (adrenocorticotropic hormone), insulin, and trypsinogen into regulated secretory vesicles. Evidence for such a common mechanism comes from experiments in which recombinant DNA techniques were used to induce the synthesis of insulin and trypsinogen in pituitary tumor cells already synthesizing ACTH. In these cells, which do not normally express insulin or trypsinogen, all three proteins segregate into the same regulated secretory vesicles and are secreted together when a hormone binds to a receptor on the pituitary cells and causes a rise in cytosolic Ca^{2+}.

Morphologic evidence suggests that the ability to form protein aggregates is a common feature of all three proteins that governs their packaging into regulated secretory vesicles. For instance, immature vesicles in this pathway — those that have just budded from the *trans*-Golgi network — contain diffuse aggregates of

secretory protein that are visible in the electron microscope. These aggregates are also found in vesicles that are in the process of budding, indicating that proteins destined for regulated secretory vesicles selectively aggregate together before their incorporation into the vesicles.

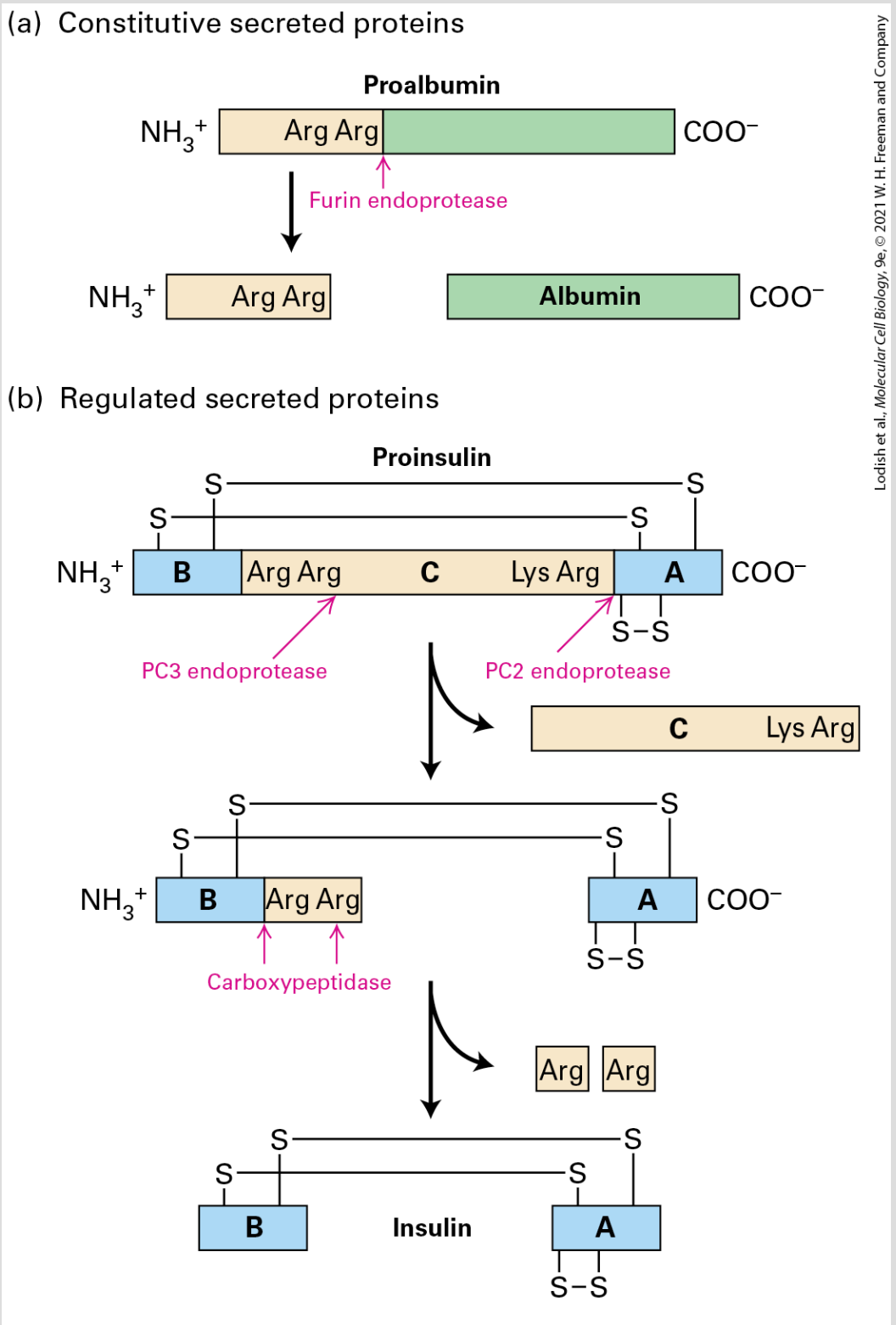
Some Proteins Undergo Proteolytic Processing After Leaving the *trans*-Golgi

For some secretory proteins (e.g., growth hormone) and certain viral membrane proteins (e.g., VSV G protein), removal of the N-terminal ER signal sequence from the nascent chain is the only known proteolytic cleavage required to convert the polypeptide to the mature, active protein (see Figure 13-10). However, some membrane proteins and many soluble secretory proteins are initially synthesized as relatively long-lived, inactive precursors, termed *proproteins*, that require further proteolytic processing to generate the mature, active proteins. Examples of proteins that undergo such processing are soluble lysosomal enzymes; many membrane proteins, such as influenza hemagglutinin (HA); and secreted proteins such as serum albumin, insulin, glucagon, and the yeast α mating factor. In general, the proteolytic conversion of a proprotein to the corresponding mature protein occurs after the

proprotein has been sorted in the *trans*-Golgi network to appropriate vesicles.

In the case of soluble lysosomal enzymes, the proproteins, called *proenzymes*, are sorted by the M6P receptor as catalytically inactive enzymes. Once in the late endosome or lysosome, a proenzyme undergoes a proteolytic cleavage that generates a smaller but enzymatically active polypeptide. Delaying the activation of lysosomal proenzymes until they reach the lysosome prevents them from digesting macromolecules in earlier compartments of the secretory pathway.

The proproteins of most constitutively secreted proteins (e.g., albumin) are cleaved only once at a site C-terminal to a dibasic recognition sequence such as Arg-Arg or Lys-Arg (Figure 14-24a). Proteolytic processing of proteins whose secretion is regulated generally entails additional cleavages. In the case of proinsulin, multiple cleavages of the single polypeptide chain yield the N-terminal B chain and the C-terminal A chain of mature insulin, which are linked by disulfide bonds, and the central C peptide, which is lost and subsequently degraded (Figure 14-24b).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-24 Proteolytic processing of proproteins in the constitutive and regulated secretory pathways. The processing of proalbumin and proinsulin is

typical of the constitutive and regulated pathways, respectively. The endoproteases that function in such processing cleave at the C-terminal end of a sequence of two consecutive amino acids. (a) The endoprotease furin acts on the precursors of constitutive secreted proteins. (b) Two endoproteases, PC2 and PC3, act on the precursors of regulated secreted proteins. The final processing of many such proteins is catalyzed by a carboxypeptidase that sequentially removes two basic amino acid residues at the C-terminus of a polypeptide. See D. Steiner et al., 1992, *J. Biol. Chem.* **267**:23435.

The processing cleavages are carried out by the members of a family of endoproteases located in the late secretory pathway, all of which cleave a protein chain on the C-terminal side of an Arg-Arg or Lys-Arg sequence. One, called *furin*, is found in all mammalian cells; it processes proteins such as albumin that are secreted constitutively. In contrast, the *PC2* and *PC3 endoproteases* are found only in cells that exhibit regulated secretion; these enzymes are localized to regulated secretory vesicles and proteolytically cleave the precursors of many hormones at specific sites.

Distinct Pathways Sort Membrane Proteins to the Apical or Basolateral Region of Polarized Cells

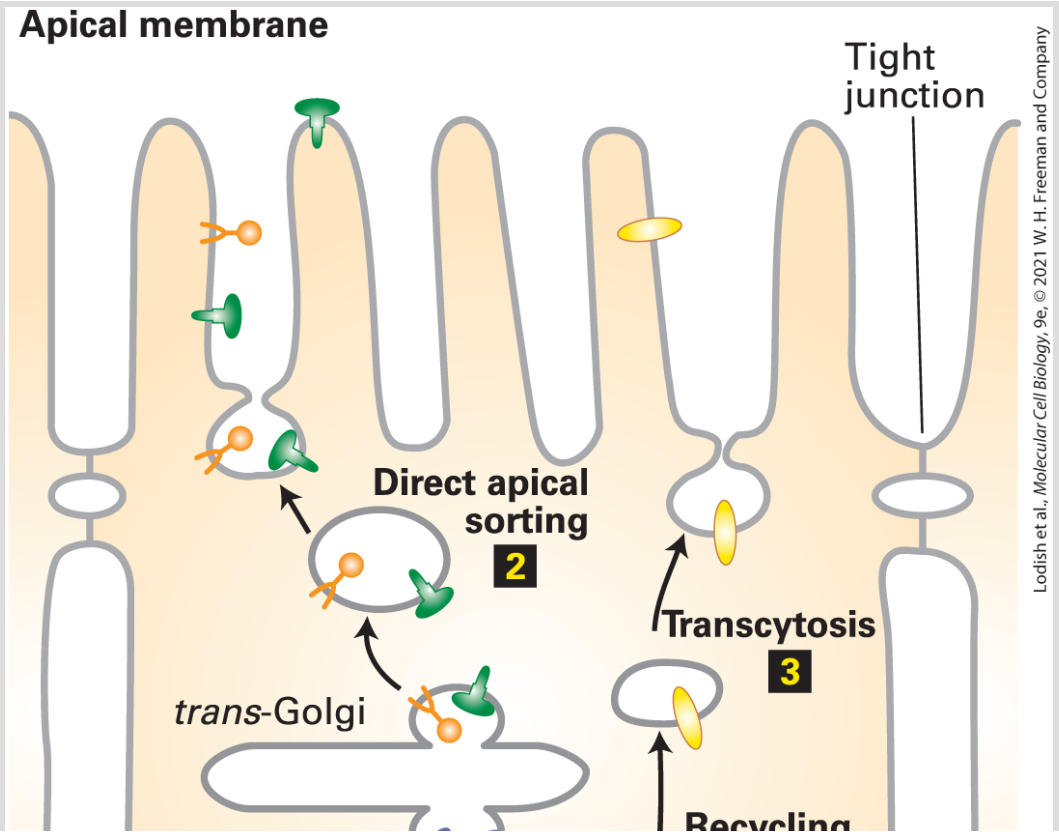
The plasma membrane of a polarized epithelial cell is divided into two domains: *apical* and *basolateral*. Tight junctions located between the two domains prevent the movement of plasma-membrane

proteins between them (see Figure 20-19). Several sorting mechanisms direct newly synthesized membrane proteins to either the apical or the basolateral domain of epithelial cells, and any one protein may be sorted by more than one mechanism. As a result of this sorting and the restriction on protein movement within the plasma membrane by tight junctions, distinct sets of proteins are found in the apical and basolateral domains. This preferential localization of certain transport proteins is critical to a variety of important physiological functions, such as absorption of nutrients from the intestinal lumen and acidification of the stomach lumen (see Figures 11-30 and 11-31).

In one mechanism, the sorting takes place in the *trans*-Golgi network. Microscopic and cell-fractionation studies indicate that proteins destined for either the apical or the basolateral membrane are initially transported together to the membranes of the *trans*-Golgi network. In some cases, proteins destined for the apical membrane are sorted into their own transport vesicles that bud from the *trans*-Golgi network and then move to the apical region, whereas proteins destined for the basolateral membrane are sorted into other vesicles that move to the basolateral region. The different vesicle types can be distinguished by their protein constituents, including distinct Rab and v-SNARE proteins, which apparently target them to the appropriate plasma-membrane domain. In this mechanism, segregation of proteins destined for the two domains

occurs as cargo proteins are incorporated into particular types of vesicles budding from the *trans*-Golgi network.

Such direct basolateral-apical sorting has been investigated in cultured Madin-Darby canine kidney (MDCK) cells, a line of cultured polarized epithelial cells (see Figure 4-4). In MDCK cells infected with the influenza virus, progeny viruses bud only from the apical membrane, whereas in cells infected with vesicular stomatitis virus, progeny viruses bud only from the basolateral membrane. This difference occurs because the HA glycoprotein of influenza virus is transported from the Golgi complex exclusively to the apical membrane and the VSV G protein is transported only to the basolateral membrane (Figure 14-25).



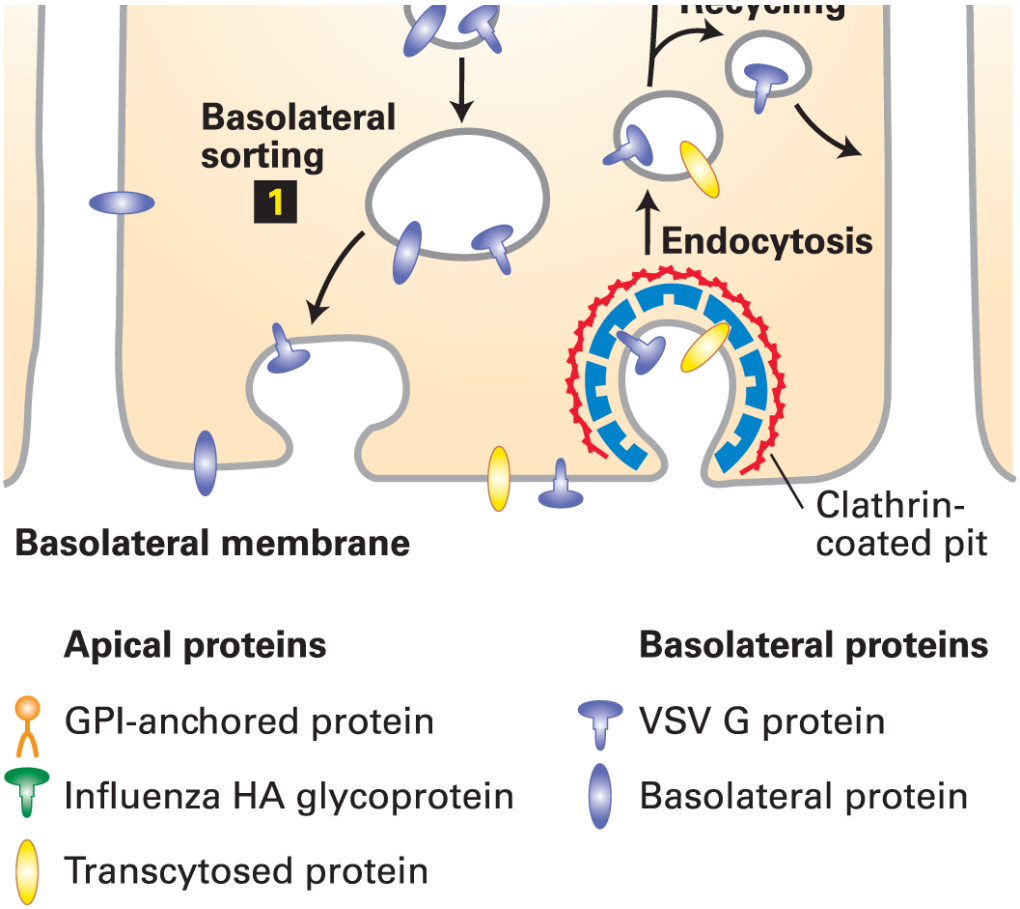


FIGURE 14-25 Sorting of proteins destined for the apical and basolateral plasma membranes of polarized cells. Cells that form an epithelial cell layer contain distinct apical and basolateral domains of the plasma membrane separated by tight junctions. Distinct sorting pathways in the late secretory pathway govern sorting to either the basolateral or apical membrane. **1**: In the *trans*-Golgi, one type of vesicle transports proteins such as VSV G protein and endogenous proteins that carry a basolateral targeting sequence directly to the basolateral membrane. **2**: A second type of vesicle budding from the *trans*-Golgi carries influenza HA protein and GPI-linked proteins directly to the apical membrane. **3**: Through a process known as *transcytosis*, proteins can move from the basolateral to the apical membrane. Transcytosis involves endocytosis in clathrin/AP coated vesicles from the basolateral membrane, recycling from an endocytic compartment of basolateral proteins back to the basolateral membrane, and delivery of transcytosed apical proteins to the apical membrane. See K. Simons and A. Wandinger-Ness, 1990, *Cell* **62**:207; and K. Mostov et al., 1992, *J. Cell Biol.* **116**:577.

Mutational studies on proteins, such as the VSV G protein, that are specifically targeted to the basolateral domain have defined Tyr-X-X-Φ and Asp-X-Leu-Leu sorting signals. As we have seen, proteins bearing these motifs are recruited into clathrin/AP-coated vesicles, strongly implicating clathrin-coated vesicles in the sorting of proteins to the basolateral membrane.

Among the cellular proteins that undergo similar apical-basolateral sorting in the Golgi are those with a *glycosylphosphatidylinositol (GPI) membrane anchor*. In MDCK cells and most other types of epithelial cells, GPI-anchored proteins are targeted to the apical membrane. In membranes, GPI-anchored proteins are clustered into lipid rafts, which are rich in sphingolipids (see Chapter 10). This finding suggests that lipid rafts are localized to the apical membrane along with proteins that preferentially partition them in many cells. However, the GPI anchor is not an apical sorting signal in all polarized cells; in thyroid cells, for example, GPI-anchored proteins are targeted to the basolateral membrane. Other than GPI anchors, no unique sequences have been identified that are both necessary and sufficient to target proteins to either the apical or basolateral domain. Instead, each membrane protein may contain multiple sorting signals, any one of which can target it to the appropriate plasma-membrane domain. The identities of these complex signals and of the vesicle coat proteins that recognize them are currently

being pursued for a number of different proteins that are sorted to specific plasma-membrane domains of polarized epithelial cells.

Another mechanism for sorting apical and basolateral proteins, also illustrated in Figure 14-25, operates in hepatocytes. The basolateral membranes of hepatocytes face the blood (like those of intestinal epithelial cells), and the apical membranes line the small intercellular channels into which bile is secreted. In hepatocytes, newly made apical and basolateral proteins are first transported in vesicles from the *trans*-Golgi network to the basolateral region and incorporated into the plasma membrane by exocytosis (i.e., fusion of the vesicle membrane with the plasma membrane). From there both basolateral and apical proteins are endocytosed in the same vesicles, but then their paths diverge. The endocytosed basolateral proteins are sorted into transport vesicles that recycle them to the basolateral membrane. In contrast, the apically destined endocytosed proteins are sorted into transport vesicles that move across the cell and fuse with the apical membrane, a process called **transcytosis**. This process is also used to move extracellular materials from one side of an epithelium to another. Even in epithelial cells, such as MDCK cells, in which apical-basolateral protein sorting occurs in the Golgi, transcytosis may provide an editing function by which an apical protein sorted incorrectly to the basolateral membrane is subjected to endocytosis and then correctly delivered to the apical membrane.

KEY CONCEPTS OF SECTION 14.4

Later Stages of the Secretory Pathway

- The *trans*-Golgi network is a major branch point in the secretory pathway where soluble secreted proteins, lysosomal proteins, and in some cells, membrane proteins destined for the basolateral or apical plasma membrane are segregated into different transport vesicles.
- Many vesicles that bud from the *trans*-Golgi network as well as endocytic vesicles bear a coat composed of AP (adapter protein) complexes and clathrin (see Figure 14-19).
- The pinching off of clathrin-coated vesicles requires dynamin, which forms a collar around the neck of the vesicle bud and hydrolyzes GTP (see Figure 14-20).
- Soluble enzymes destined for lysosomes are modified in the *cis*-Golgi by the addition of multiple mannose 6-phosphate (M6P) residues to their oligosaccharide chains.
- M6P receptors in the membrane of the *trans*-Golgi network bind proteins bearing M6P residues and direct them to late endosomes, where receptors and their ligand proteins dissociate. The receptors are then recycled to the Golgi or plasma membrane, and the lysosomal enzymes are delivered to lysosomes (see Figure 14-23).
- Regulated secreted proteins are concentrated and stored in secretory vesicles to await a neuronal or hormonal signal for exocytosis. Protein aggregation within the *trans*-Golgi network may play a role in sorting secreted proteins to the regulated secretory pathway.
- Many proproteins transported through the secretory pathway undergo post-Golgi proteolytic cleavages that yield the mature, active proteins. This proteolytic maturation can occur in vesicles carrying proteins from the *trans*-Golgi network to the cell surface, in the late endosome, or in the lysosome.
- In polarized epithelial cells, membrane proteins destined for the apical and basolateral domains of the plasma membrane are sorted in the *trans*-Golgi network into different transport vesicles (see Figure 14-25). Tyr-X-X-Φ and Asp-X-Leu-Leu sorting signals target proteins to the basolateral membrane, whereas a GPI anchor targets proteins to the apical membrane.
- In hepatocytes and some other polarized cells, all plasma-membrane proteins are directed first to the basolateral membrane. Apically destined proteins are then endocytosed and moved across the cell to the apical membrane (transcytosis).

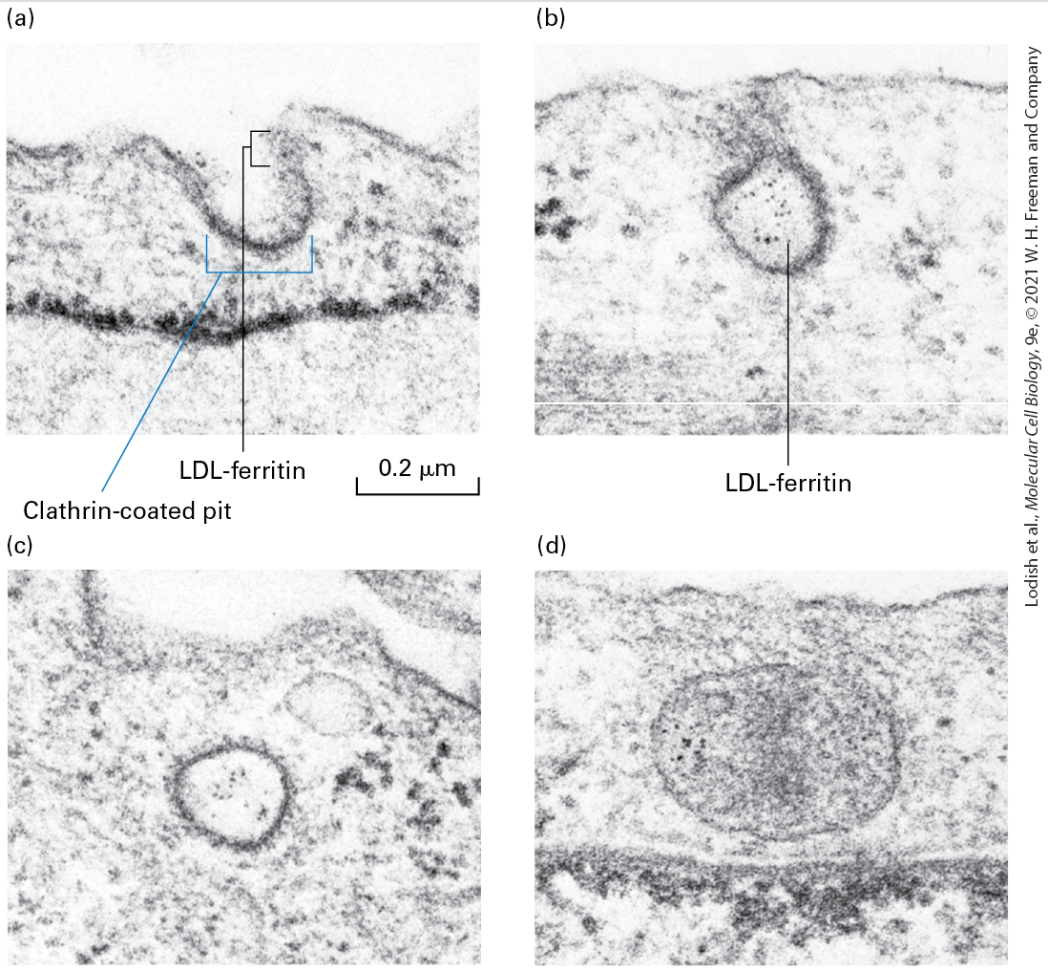
14.5 Receptor-Mediated Endocytosis

In previous sections, we have explored the main pathways of vesicle trafficking responsible for delivering soluble and membrane secretory proteins synthesized on the rough ER to the cell surface. The same basic principle of coated vesicle budding and fusion are responsible for the ability of animal cells to internalize macromolecules from their surroundings and to deliver these macromolecules to particular intracellular destinations. In this process, known as **receptor-mediated endocytosis**, a specific receptor on the cell surface binds tightly to an extracellular macromolecular ligand that it recognizes; then the complex of receptor and ligand are recruited to a plasma-membrane region that buds inward to form a coated transport vesicle. Plant cells and fungal cells take up small molecules by endocytosis, but generally do not take up macromolecules, which cannot readily pass through the cell wall because of their size.

Among the common macromolecules that vertebrate cells internalize by receptor-mediated endocytosis are cholesterol-containing low-density lipoprotein (LDL) particles, the iron-carrying protein transferrin, many protein hormones (e.g., insulin), and certain glycoproteins. Receptor-mediated endocytosis of such ligands generally occurs via clathrin/AP2-coated pits and vesicles in a process similar to the packaging of lysosomal enzymes by the

binding of M6P in the *trans*-Golgi network (see Figure 14-23). As noted earlier, some M6P receptors are found on the cell surface, and these receptors participate in the receptor-mediated endocytosis of lysosomal enzymes that are mistakenly secreted. In general, the transmembrane receptor proteins that function in the uptake of extracellular ligands are internalized from the cell surface during endocytosis and are then sorted and recycled back to the cell surface, much as M6P receptors are recycled to the plasma membrane and *trans*-Golgi. The rate at which a ligand is internalized is limited by the amount of its corresponding receptor on the cell surface.

Clathrin/AP2-coated pits make up about 2 percent of the surface of cells such as hepatocytes and fibroblasts. Many internalized ligands have been observed in clathrin/AP2-coated pits and vesicles, which are thought to function as intermediates in the endocytosis of most (though not all) ligands bound to cell-surface receptors (Figure 14-26). Some receptors are clustered over clathrin-coated pits even in the absence of ligand. Other receptors diffuse freely in the plane of the plasma membrane but undergo a conformational change when they bind to ligand, so that when the receptor-ligand complex diffuses into a clathrin-coated pit, it is retained there. Two or more types of receptor-bound ligands, such as LDL and transferrin, can be seen in the same coated pit or vesicle.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 14-26 The initial stages of receptor-mediated endocytosis of low-density lipoprotein (LDL) particles are revealed by electron microscopy. Cultured human fibroblasts were incubated in a medium containing LDL particles covalently linked to the electron-dense, iron-containing protein ferritin; each small iron particle in ferritin is visible as a small dot under the electron microscope. Cells were initially incubated at 4 °C; at this temperature LDL can bind to its receptor, but internalization does not occur. After excess LDL not bound to the cells was washed away, the cells were warmed to 37 °C and then prepared for microscopy at periodic intervals. (a) A coated pit, showing the clathrin coat on the inner (cytosolic) surface of the pit, soon after the temperature was raised. (b) A pit containing LDL, apparently closing on itself to form a coated vesicle. (c) A coated vesicle containing ferritin-tagged LDL particles. (d) Ferritin-tagged LDL particles in a smooth-surfaced

early endosome, 6 minutes after internalization began. See M. S. Brown and J. Goldstein, 1986, *Science* **232**:34.

[Republished with permission from *Nature*, from J. Goldstein et al., 1979, "Coated Pits, Coated Vesicles, and Receptor-Mediated Endocytosis," *Nature* **279**:679–685; permission conveyed through the Copyright Clearance Center, Inc.]

Cells Take Up Lipids from the Blood in the Form of Large, Well-Defined Lipoprotein Complexes

Lipids absorbed from the diet in the intestines or stored in adipose tissue can be distributed to cells throughout the body. To facilitate the mass transfer of lipids between cells, animals have evolved an efficient way to package hundreds or even thousands of lipid molecules into water-soluble, macromolecular carriers, called *lipoproteins*, which cells can take up from the circulation as an ensemble. A lipoprotein particle has a shell composed of proteins (*apolipoproteins*) overlying a cholesterol-containing phospholipid monolayer. The shell is amphipathic because its outer surface is hydrophilic, making the particle water soluble, and its inner surface is hydrophobic. Beneath the hydrophobic inner surface of the shell is a core of neutral lipids containing mostly cholesteryl esters, triglycerides, or both. Mammalian lipoproteins fall into different classes, defined by their differing buoyant densities. The class we will consider here is **low-density lipoprotein (LDL)**. A typical LDL

particle, depicted in Figure 14-27, is a sphere 20–25 nm in diameter. The amphipathic outer shell is composed of a phospholipid monolayer and a single molecule of a large protein known as *apoB-100*; the core of the particle is packed with cholesterol in the form of cholesteryl esters.

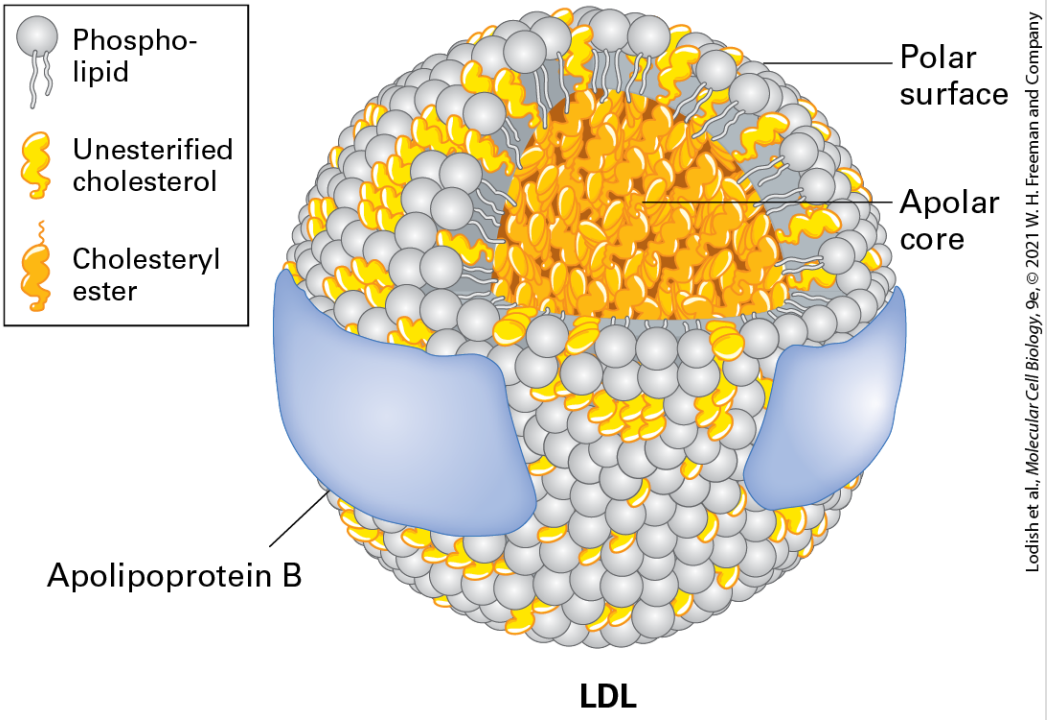
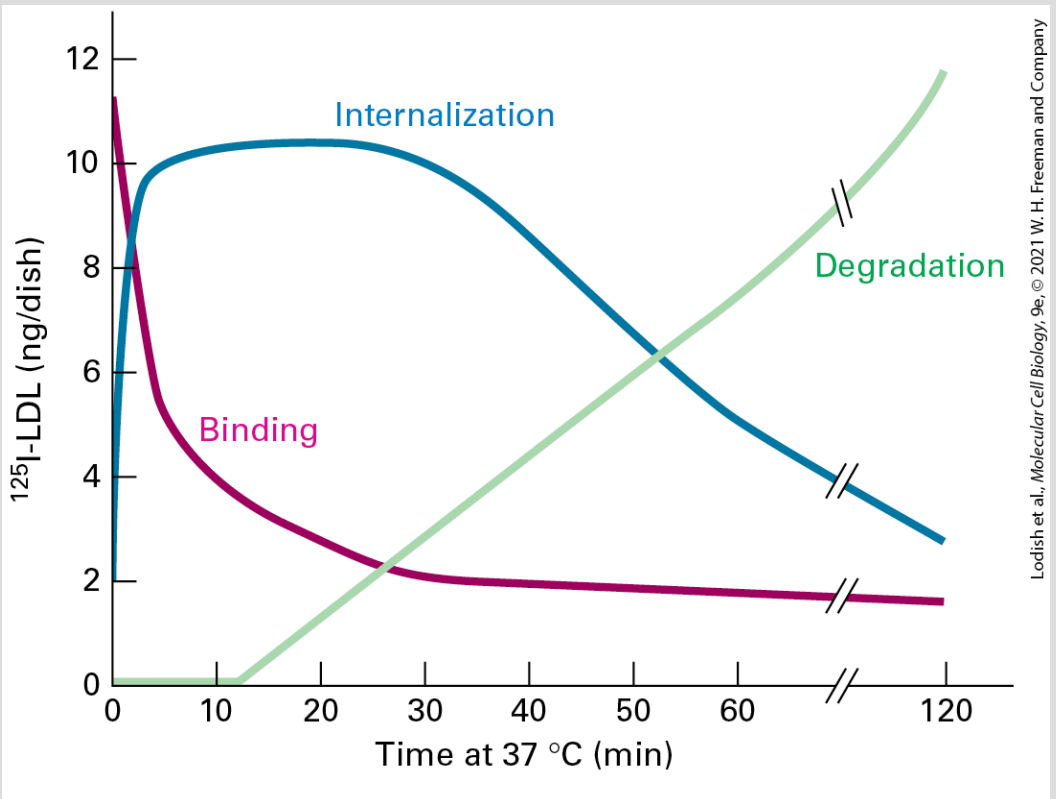


FIGURE 14-27 Model of low-density lipoprotein (LDL). All classes of lipoproteins have the same general structure: an amphipathic shell composed of apolipoprotein, a phospholipid monolayer (not bilayer), and cholesterol, and a hydrophobic core composed mostly of cholesteryl esters or triglycerides, or both, but with minor amounts of other neutral lipids (e.g., some vitamins). This model of LDL is based on electron microscopy and other low-resolution biophysical methods. LDL is unique in that it contains only a single molecule of one type of apolipoprotein (ApoB), which appears to wrap around the outside of the particle as a band of protein. The other lipoproteins contain multiple apolipoprotein molecules, often of different types. See M. Krieger, 1995, in E.

Haber, ed., *Molecular Cardiovascular Medicine, Scientific American Medicine*, pp. 31–47.

We have seen that a general feature of coated vesicles is the ability to select specific cargo molecules; this process of cargo selection was first studied in the context of endocytosis of labeled LDL particles added to intact cells. Two general experimental approaches have been used to study how labeled LDL particles enter cells. The first method makes use of LDL that has been labeled by the covalent attachment of radioactive ^{125}I to the side chains of tyrosine residues in apoB-100 on the surfaces of the LDL particles. After cultured cells are incubated for several hours with the labeled LDL, it is possible to determine how much LDL is bound to the surfaces of cells, how much is internalized, and how much of the apoB-100 component of the LDL is transported to the lysosome and there degraded by enzymatic hydrolysis to individual amino acids. The degradation of apoB-100 can be detected by the release of ^{125}I-tyrosine into the culture medium. Figure 14-28 shows the time course of these events in receptor-mediated cellular LDL processing, determined by pulse-chase experiments with a fixed concentration of ^{125}I-labeled LDL. These experiments clearly demonstrate the order of events: surface binding of LDL → internalization → degradation. The second approach involves tagging LDL particles with an electron-dense label that can be detected by electron microscopy. Such studies can reveal the details of how LDL particles first bind to the surface of cells at sites of nascent endocytic vesicles known as *clathrin-coated*

pits, then remain associated with the coated pits as they invaginate and bud off to form clathrin-coated vesicles, and are finally transported to endosomes.



EXPERIMENTAL FIGURE 14-28 Pulse-chase experiment demonstrates precursor-product relations in cellular uptake of LDL. Cultured normal human skin fibroblasts were incubated in a medium containing ^{125}I-LDL for 2 hours at 4 °C (the pulse). After excess ^{125}I-LDL not bound to the cells was washed away, the cells were incubated at 37 °C for the indicated amounts of time in the absence of external LDL (the chase). The amounts of surface-bound, internalized, and degraded (hydrolyzed) ^{125}I-LDL were measured. Binding, but not internalization or hydrolysis, of LDL apoB-100 occurs during the 4 °C pulse. The data show the very rapid disappearance of bound ^{125}I-LDL from the surface as it is internalized after the cells have been warmed to allow membrane movements. After a lag period of 15–20 minutes, lysosomal degradation of the internalized ^{125}I-LDL commences. See M. S. Brown and J. L. Goldstein, 1976, *Cell* **9**:663.

Receptors for Macromolecular Ligands Contain Sorting Signals That Target Them for Endocytosis

The key to understanding how LDL particles bind to the cell surface and are then taken up into endocytic vesicles was the discovery of the *LDL receptor*. The LDL receptor is an 839-residue type I membrane glycoprotein; it has a long N-terminal exoplasmic segment that contains the LDL-binding domain, a single transmembrane segment, and a short C-terminal cytosolic segment. Seven cysteine-rich repeats form the LDL-binding domain, which interacts with the apoB-100 molecule in an LDL particle. Figure 14-29 shows how the LDL receptor facilitates internalization of LDL particles by receptor-mediated endocytosis. After internalized LDL particles reach lysosomes, lysosomal proteases hydrolyze their surface apolipoproteins and lysosomal cholesteryl esterases hydrolyze their core cholesteryl esters. The unesterified cholesterol is then free to leave the lysosome and be used as necessary by the cell in the synthesis of membranes or various cholesterol derivatives.

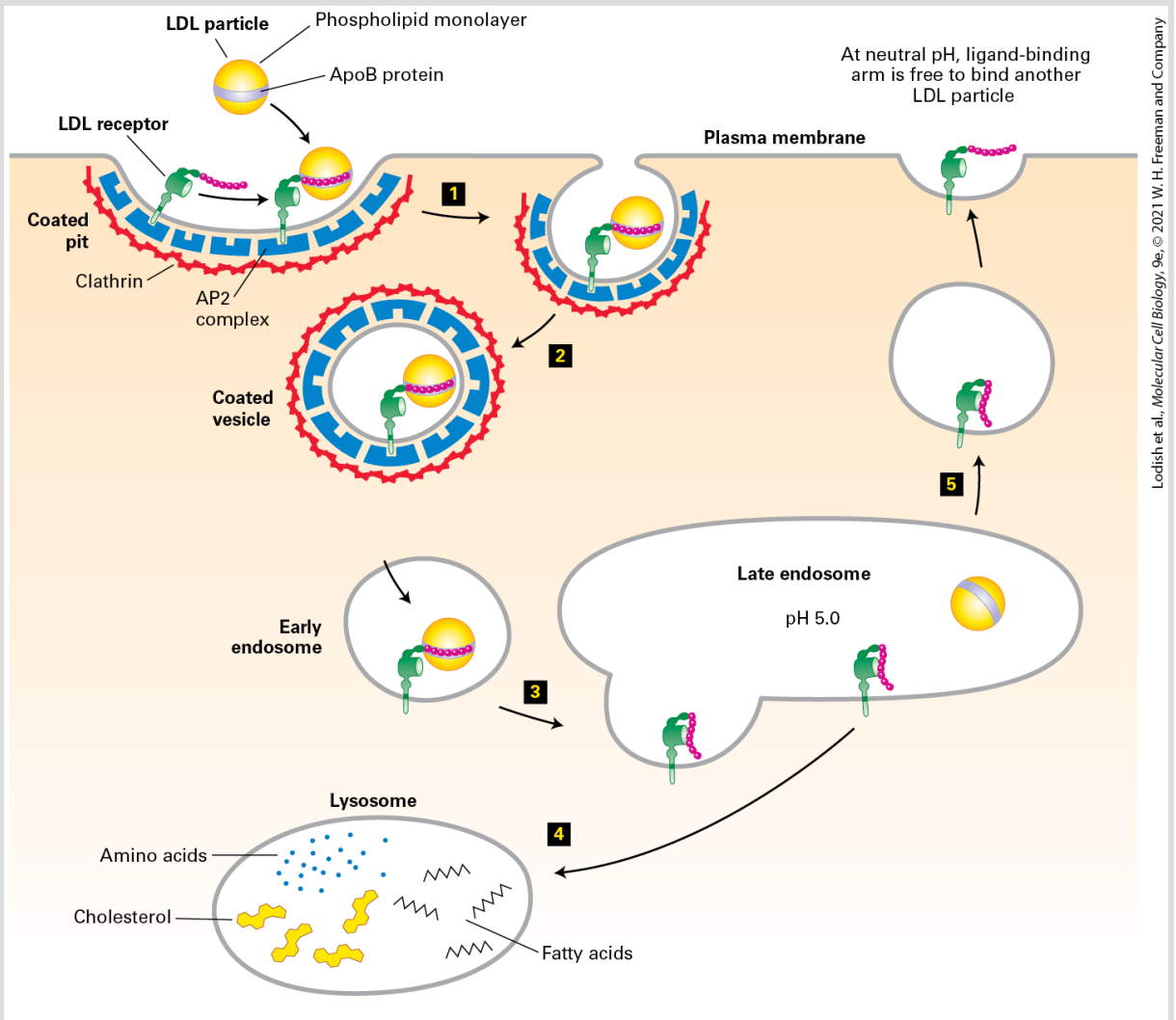


FIGURE 14-29 Endocytic pathway for internalizing low-density lipoprotein (LDL). Step 1: A cell-surface LDL receptor binds to an ApoB protein embedded in the phospholipid outer layer of an LDL particle. Interaction between the NPXY sorting signal in the cytosolic tail of the LDL receptor and the AP2 complex incorporates the receptor-ligand complex into a nascent endocytic vesicle. Step 2: Clathrin-coated pits containing receptor-LDL complexes are pinched off by the same dynamin-mediated mechanism used to form clathrin/AP1-coated vesicles on the *trans*-Golgi network (see Figure 14-19). Step 3: After the vesicle coat is shed, the uncoated endocytic vesicle (early endosome) fuses with a late endosome. The acidic pH in this compartment causes a conformational change in the LDL receptor that leads to release of the bound LDL particle. Step 4: The late endosome fuses with a lysosome, and the proteins and lipids of the free LDL particle are broken down into their constituent parts by enzymes in the lysosome. Step 5: The LDL receptor is recycled to the cell surface, where at the neutral pH of the exterior medium the receptor undergoes a

conformational change so that it can bind another LDL particle. See M. S. Brown and J. L. Goldstein, 1986, *Science* **232**:34; and G. Rudenko et al., 2002, *Science* **298**:2353.

Mutations in the human gene for the LDL receptor (*LDLR*) can cause a hereditary disease that is marked by elevated plasma LDL known as *familial hypercholesterolemia (FH)*. In patients who have one normal and one defective copy of the *LDLR* gene (heterozygotes), LDL in the blood is increased about twofold. Those with two defective *LDLR* genes (homozygotes) have LDL levels that are from fourfold to sixfold higher than normal. High levels of blood LDL can lead to excessive deposition of LDL, forming atherosclerotic plaque in coronary arteries. Without medical intervention, FH homozygotes usually die of heart attacks before reaching their late twenties.

A variety of mutations in the *LDLR* gene can cause FH. Some mutations prevent the synthesis of the LDL receptor protein; others prevent proper folding of the receptor protein in the ER, leading to its premature degradation (see Chapter 13); and still other mutations reduce the ability of the LDL receptor to bind LDL tightly. A particularly informative group of mutant receptors are expressed on the cell surface and bind LDL normally but cannot mediate the internalization of bound LDL. In individuals with this type of defect, plasma-membrane receptors for other ligands are internalized normally, but the mutant LDL receptor is not recruited into coated pits. Analysis of this mutant receptor and other mutant LDL

receptors generated experimentally and expressed in fibroblasts identified a four-residue motif in the cytosolic segment of the receptor that is crucial for its internalization: Asn-Pro-X-Tyr, where X can be any amino acid. This *NPXY sorting signal* binds to the AP2 complex, linking the clathrin/AP2 coat to the cytosolic segment of the LDL receptor in coated pits. A mutation in any of the conserved residues of the NPXY signal abolishes the ability of the LDL receptor to be incorporated into coated pits. ■

The Acidic pH of Late Endosomes Causes Most Receptor-Ligand Complexes to Dissociate

The overall rate of endocytic internalization of the plasma membrane is quite high: cultured fibroblasts regularly internalize 50 percent of their cell-surface proteins and phospholipids each hour. Most cell-surface receptors that undergo endocytosis will repeatedly deposit their ligands within the cell and then recycle to the plasma membrane to mediate internalization of ligand molecules once again. The LDL receptor, for instance, makes one round trip into and out of the cell interior every 10–20 minutes, for a total of about a hundred trips in its 20-hour life span.

Internalized receptor-ligand complexes commonly follow the pathway depicted for the M6P receptor in Figure 14-22 and the LDL

receptor in Figure 14-29. Endocytosed cell-surface receptors typically dissociate from their ligands within late endosomes, which appear as spherical vesicles with tubular branching membranes located a few micrometers from the cell surface.

The dissociation of receptor-ligand complexes in late endosomes occurs not only in the endocytic pathway but also in the delivery of soluble lysosomal enzymes via the secretory pathway (see Figure 14-23). As discussed in Chapter 11, the membranes of late endosomes and lysosomes contain V-class proton pumps that act in concert with Cl^- channels to acidify the vesicle lumen (see Figure 11-13). Most receptors, including the M6P receptor and the LDL receptor, bind their ligands tightly at neutral pH, but release their ligands if the pH is lowered to 6.0 or below. The late endosome is the first organelle encountered by receptor-ligand complexes whose luminal pH is sufficiently acidic to promote dissociation of most endocytosed receptors from their tightly bound ligands.

The mechanism by which the LDL receptor releases bound LDL particles is now understood in detail (Figure 14-30). At the endosomal pH of 5.0–5.5, histidine residues in a region known as the *β-propeller domain* of the receptor become protonated, forming a site that can bind with high affinity to the negatively charged cysteine-rich repeats in the LDL-binding domain. This intramolecular interaction sequesters the repeats in a conformation like that of a closed jackknife that cannot simultaneously bind to

apoB-100, thus causing release of the bound LDL particle. The LDL receptor in this relatively compact closed state is efficiently recycled back to the plasma membrane for another round of binding and uptake of LDL.

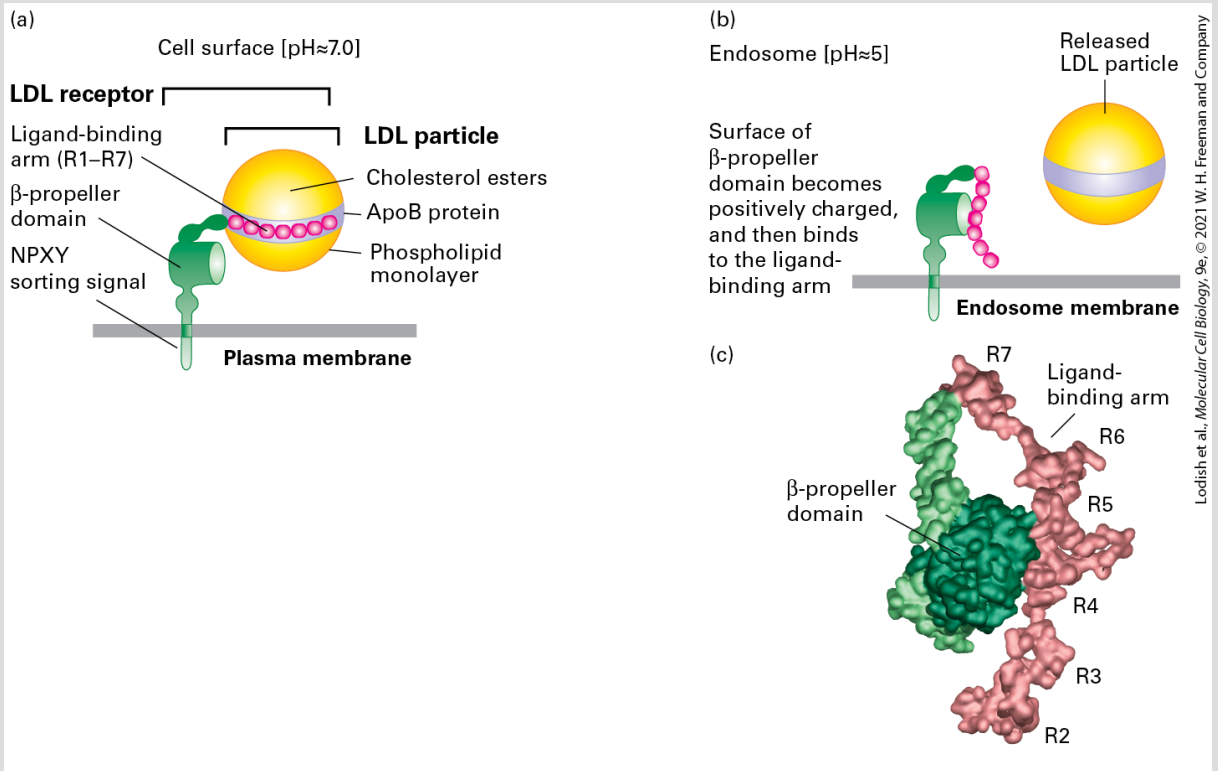


FIGURE 14-30 Model for pH-dependent binding of LDL particles by the LDL receptor. Schematic depiction of an LDL receptor at the neutral pH found at the cell surface and at the acidic pH found in the interior of the late endosome. (a) At the cell surface, apoB-100 on the surface of an LDL particle binds tightly to the receptor. Of the seven cysteine-rich repeats (R1–R7) in the ligand-binding arm, R4 and R5 appear to be most critical for LDL binding. (b) Within the endosome, histidine residues in the β-propeller domain of the LDL receptor become protonated. The positively charged propeller can bind with high affinity to the ligand-binding arm, which contains negatively charged residues, causing release of the LDL particle. (c) Experimental electron density and C_α backbone trace model of the extracellular region of the LDL receptor at pH 5.3 based on x-ray crystallographic analysis. In this conformation, extensive hydrophobic and ionic interactions occur between the β-propeller and the R4 and R5 repeats.

[Part (b) data from G. Rudenko et al., 2002, *Science* **298**:2353; PDB ID 1n7d.]

As described in Chapter 6, human mutations that cause a loss of function of the gene PCSK9 lead to a decrease in blood LDL. We now know that PCSK9 encodes a secreted protein whose normal function is to negatively regulate the LDL receptor. Biochemical and structural studies show that PCSK9 binds to the LDL-binding domain of the LDL receptor, thereby preventing the receptor from closing into a compact state. This more open conformation of the receptor cannot be recycled to the plasma membrane and is transported to the lysosome where it is degraded. Inhibitors of PCSK9 are a promising new class of LDL-lowering drugs.

Receptor-Mediated Endocytosis Can Down-Regulate Signaling Receptors

Signaling receptors are an important class of plasma membrane proteins that can undergo receptor-mediated endocytosis. These receptors bind to extracellular ligands and respond by generating an appropriate intracellular signal. As described in Chapter 16, the epidermal growth factor receptor (EGFR) activates a signal transduction pathway in response to the extracellular growth factor protein EGF. Treatment of cells with EGF both activates EGFR and stimulates receptor-mediated endocytosis of EGFR. Removal of EGFR from the plasma membrane by endocytosis leads to an

attenuation of the signaling response in an adaptive process, known as *down-regulation*, which extends the dynamic range of EGF concentrations over which the system can respond. The mechanism of EGFR down-regulation is one of the best understood examples of regulated endocytosis.

EGFR is a Type I cell surface glycoprotein with a large extracellular domain that binds the ligand EGF, a single transmembrane domain, and a cytoplasmic signaling domain. Normally, EGFR resides in the plasma membrane as a monomer, but when EGF is present, two copies of EGF bind to two copies of EGFR, forming a dimer that brings the cytoplasmic signaling domains of EFGR together and activates signal transduction (see Figure 16-8). Activated EGFR enters the endocytic pathway because the dimeric form of the receptor binds to clathrin/AP2-coated pits. Mutational studies of the cytoplasmic domain of EGFR have identified a di-leucine (Leu-Leu) sorting signal that is necessary for rapid endocytosis. Other proteins employ a cytoplasmic di-leucine sorting signal for entry into the endocytic pathway, but EGFR is unusual in that the cytoplasmic di-leucine signal only interacts with clathrin/AP2 complexes during ligand-induced endocytosis when the cytoplasmic domain is dimeric.

Endocytosis delivers EGFR to the late endosome, and from this location EGFR is available to recycle to the plasma membrane, in a pathway similar to that for the LDL receptor (see Figure 14-29).

Complete down-regulation depends on further transport of EGFR to the lysosome where it is degraded. In addition to enabling interaction with clathrin/AP2 complexes, ligand-induced activation of EGFR also triggers covalent modification of the cytoplasmic domain of EGFR by the small protein ubiquitin. As we will see in the next section, monoubiquitinylation marks EGFR for entry into multivesicular bodies that are ultimately degraded in lysosomes (see Figure 14-32).

KEY CONCEPTS OF SECTION 14.5

Receptor-Mediated Endocytosis

- Some extracellular ligands that bind to specific cell-surface receptors are internalized, along with their receptors, in clathrin-coated vesicles whose coats also contain AP2 complexes (see Figure 14-26).
- Sorting signals in the cytosolic domain of cell-surface receptors target them into clathrin/AP2-coated pits for internalization. Known signals include the Asn-Pro-X-Tyr, Tyr-X-X-Φ, and Leu-Leu sequences (see Table 14-2).
- The endocytic pathway delivers some ligands (e.g., LDL particles) to lysosomes, where they are degraded. Transport vesicles from the cell surface first fuse with late endosomes, which subsequently fuse with lysosomes.
- Most receptor-ligand complexes dissociate in the acidic milieu of the late endosome; the receptors are recycled to the plasma membrane, while the ligands are sorted to lysosomes (see Figure 14-29).
- Signaling receptors can be down-regulated by endocytosis. In these cases, the cytoplasmic domain of the receptors only interact with clathrin/AP2-coated pits when the receptor is in a ligand-activated state.

14.6 Directing Membrane Proteins and Cytosolic Materials to the Lysosome for Degradation

The major function of lysosomes is to degrade extracellular materials taken up by the cell and to degrade intracellular components under specific conditions. Materials to be degraded must be delivered to the lumen of the lysosome, where the various degradative enzymes reside. As we have just seen, endocytosed ligands (e.g., LDL particles) that dissociate from their receptors in the late endosome subsequently enter the lysosomal lumen when the membrane of the late endosome fuses with the membrane of the lysosome (see Figure 14-29).

It is apparent how the general vesicular trafficking mechanism discussed in this chapter can be used to deliver the luminal contents of an endosomal organelle to the lumen of the lysosome for degradation. In this section, we will discuss two further pathways. Membrane proteins delivered to the lysosome by the typical vesicular trafficking process we have discussed the previous section should ultimately be delivered to the membrane of the lysosome. How, then, are membrane proteins delivered to the interior of the lysosome for degradation? The first pathway discussed in this section, used to degrade endocytosed membrane proteins, utilizes an unusual type of vesicle that buds into the lumen of the endosome

to produce a multivesicular endosome. The second pathway that will be discussed is used by the cell for delivery of cytosolic materials to the lysosomal lumen for degradation. This second pathway, known as *autophagy*, involves the de novo formation of a double-membrane organelle known as an *autophagosome* that envelops cytosolic material, such as soluble cytosolic proteins, or sometimes organelles, such as peroxisomes or mitochondria. Both pathways lead to fusion of either the multivesicular endosome or autophagosome with the lysosome, depositing the contents of these organelles into the lysosomal lumen for degradation.

Multivesicular Endosomes Segregate Membrane Proteins Destined for the Lysosomal Membrane from Proteins Destined for Lysosomal Degradation

Resident lysosomal membrane proteins, such as V-class proton pumps and amino acid transporters, can carry out their functions and remain in the lysosomal membrane, where they are protected from degradation by the soluble hydrolytic enzymes in the lumen. Such proteins are delivered to the lysosomal membrane by transport vesicles that bud from either the *trans*-Golgi network or the endosome by the same basic mechanisms described in earlier sections. In contrast, endocytosed membrane proteins that are to be degraded are transferred in their entirety to the interior of the

lysosome by a specialized delivery mechanism. Lysosomal degradation of cell-surface receptors for extracellular signaling molecules is a common mechanism for controlling the sensitivity of cells to extracellular signals (see Chapter 15). Receptors that become damaged are also targeted for lysosomal degradation.

Early evidence that membranes can be delivered to the lumen of a membrane-bounded compartment came from electron micrographs showing membrane vesicles and fragments of membranes within endosomes and lysosomes. Parallel experiments in yeast revealed that endocytosed receptor proteins targeted to the vacuole (the yeast organelle equivalent to the lysosome) were primarily associated with membrane fragments and small vesicles within the interior of the vacuole rather than with the vacuole surface membrane.

These observations suggest that endocytosed membrane proteins can be incorporated into specialized vesicles that form at the endosomal membrane (Figure 14-31). Although these vesicles are similar in size and appearance to transport vesicles, they differ topologically. Transport vesicles bud *outward* from the surface of a donor organelle into the cytosol, whereas vesicles within the endosome bud *inward* from the surface into the lumen (away from the cytosol). Mature endosomes containing numerous vesicles in their interior are usually called **multivesicular endosomes** (or bodies). The surface membrane of a multivesicular endosome then fuses with the membrane of a lysosome, thereby delivering its

internal vesicles and the membrane proteins they contain into the lysosome interior for degradation. Thus the sorting of proteins in the endosomal membrane determines which ones will remain on the lysosome surface (e.g., pumps and transporters) and which ones will be incorporated into internal vesicles and ultimately degraded in lysosomes.

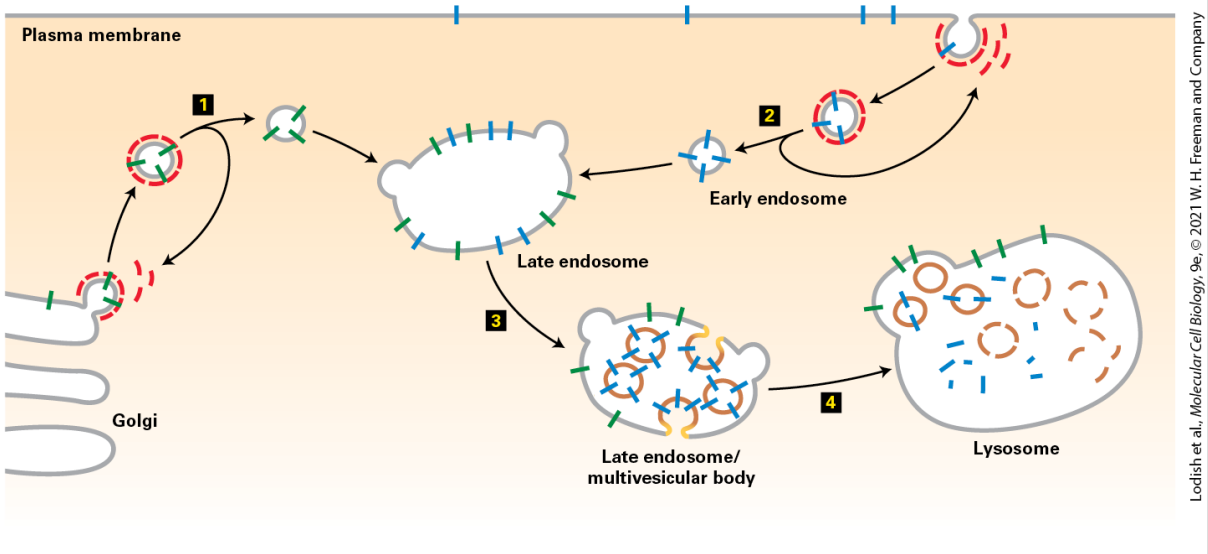
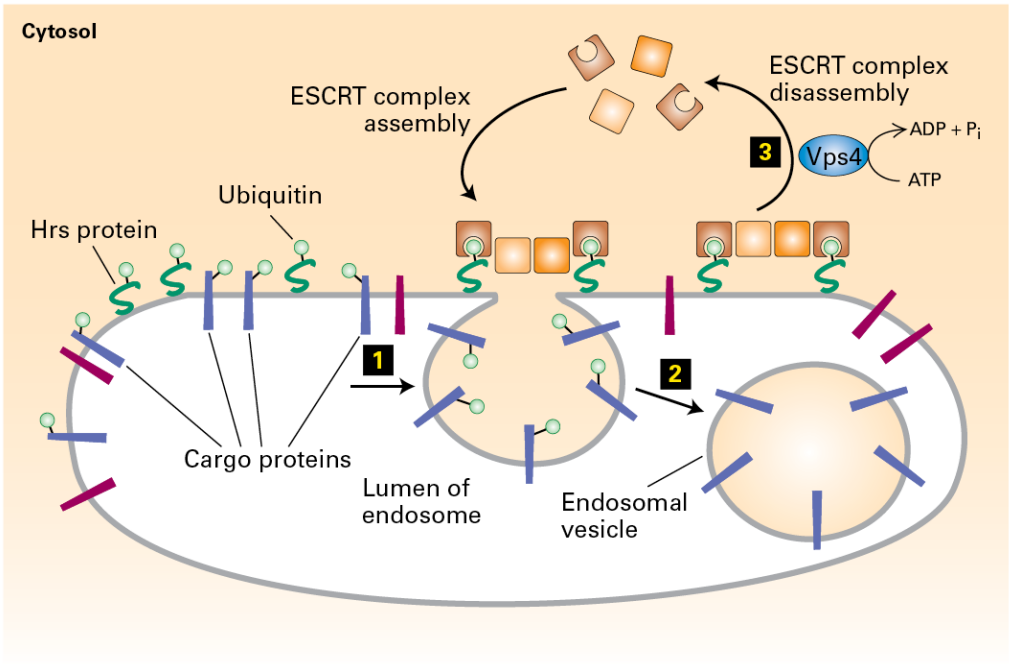


FIGURE 14-31 Delivery of plasma-membrane proteins to the lysosome interior for degradation. Early endosomes carrying endocytosed plasma-membrane proteins (blue) and vesicles carrying lysosomal membrane proteins (green) from the *trans*-Golgi network fuse with the late endosome, transferring their membrane proteins to the endosomal membrane (steps **1** and **2**). Proteins to be degraded, such as those from the early endosome, are incorporated into vesicles that bud *into* the interior of the late endosome, eventually forming a multivesicular endosome containing many such internal vesicles (step **3**). Fusion of a multivesicular endosome directly with a lysosome releases the internal vesicles into the lumen of the lysosome, where they can be degraded (step **4**). Because proton pumps and other lysosomal membrane proteins normally are not incorporated into internal endosomal vesicles, they are delivered to the lysosomal membrane and are protected from degradation. See F. Reggiori and D. J. Klionsky, 2002, *Eukaryot. Cell* **1**:11; and D. J. Katzmann et al., 2002, *Nat. Rev. Mol. Cell Biol.* **3**:893.

Many of the proteins required for inward budding of the endosomal membrane were first identified by mutations in yeast that blocked the delivery of membrane proteins to the interior of the vacuole. The current model of endosomal budding to form multivesicular endosomes is based primarily on studies in yeast, but mammalian cells appear to have essentially the same mechanism (Figure 14-32). Most cargo proteins that enter a multivesicular endosome are tagged with ubiquitin. Cargo proteins destined to enter a multivesicular endosome usually receive their ubiquitin tags at the plasma membrane, the *trans*-Golgi network, or the endosomal membrane. We have already seen how ubiquitin tagging can serve as a signal for degradation of cytosolic or misfolded ER proteins by proteasomes (see Chapters 3 and 13). When used as a signal for proteasomal degradation, the ubiquitin tag usually consists of a chain of covalently linked ubiquitin molecules (polyubiquitin), whereas ubiquitin used to tag proteins for entry into the multivesicular endosome usually takes the form of a single (monoubiquitin) molecule. In the membrane of the endosome, a ubiquitin-tagged peripheral membrane protein, known as *Hrs*, facilitates recruitment of a set of at least three different protein complexes to the membrane. These **ESCRT** (*e*ndosomal *s*orting *c*omplexes *r*equired for *t*ransport) **proteins** include the ubiquitin-binding protein *Tsg101*. The membrane-associated ESCRT proteins act to drive vesicle budding directed into the interior of the endosome as well as the loading of specific monoubiquitinylated membrane cargo proteins into the vesicle buds. Finally, the ESCRT proteins pinch off the

vesicle by forming a filamentous spiral inside the neck of a vesicle bud, releasing it and the specific membrane cargo proteins it carries into the interior of the endosome. An ATPase, known as *Vps4*, uses the energy from ATP hydrolysis to disassemble the ESCRT proteins, releasing them into the cytosol for another round of budding. In the fusion event that pinches off a completed endosomal vesicle, the ESCRT proteins and Vps4 may function like SNAREs and NSF, respectively, in the typical membrane-fusion process discussed previously (see Figure 14-11).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-32 Model of the mechanism for formation of multivesicular endosomes. In endosomal budding, ubiquitinylated Hrs on the endosomal membrane directs the loading of specific membrane cargo proteins (blue) into vesicle buds and then recruits cytosolic ESCRT protein complexes to the membrane (step 1). Note that both Hrs and the recruited cargo proteins are tagged with ubiquitin. After the set of bound ESCRT complexes mediates the completion and pinching off of the inwardly budding vesicles (step 2), these

complexes are disassembled by the ATPase Vps4 and returned to the cytosol (step **3**). See text for discussion. See O. Pornillos et al., 2002, *Trends Cell Biol.* **12**:569.

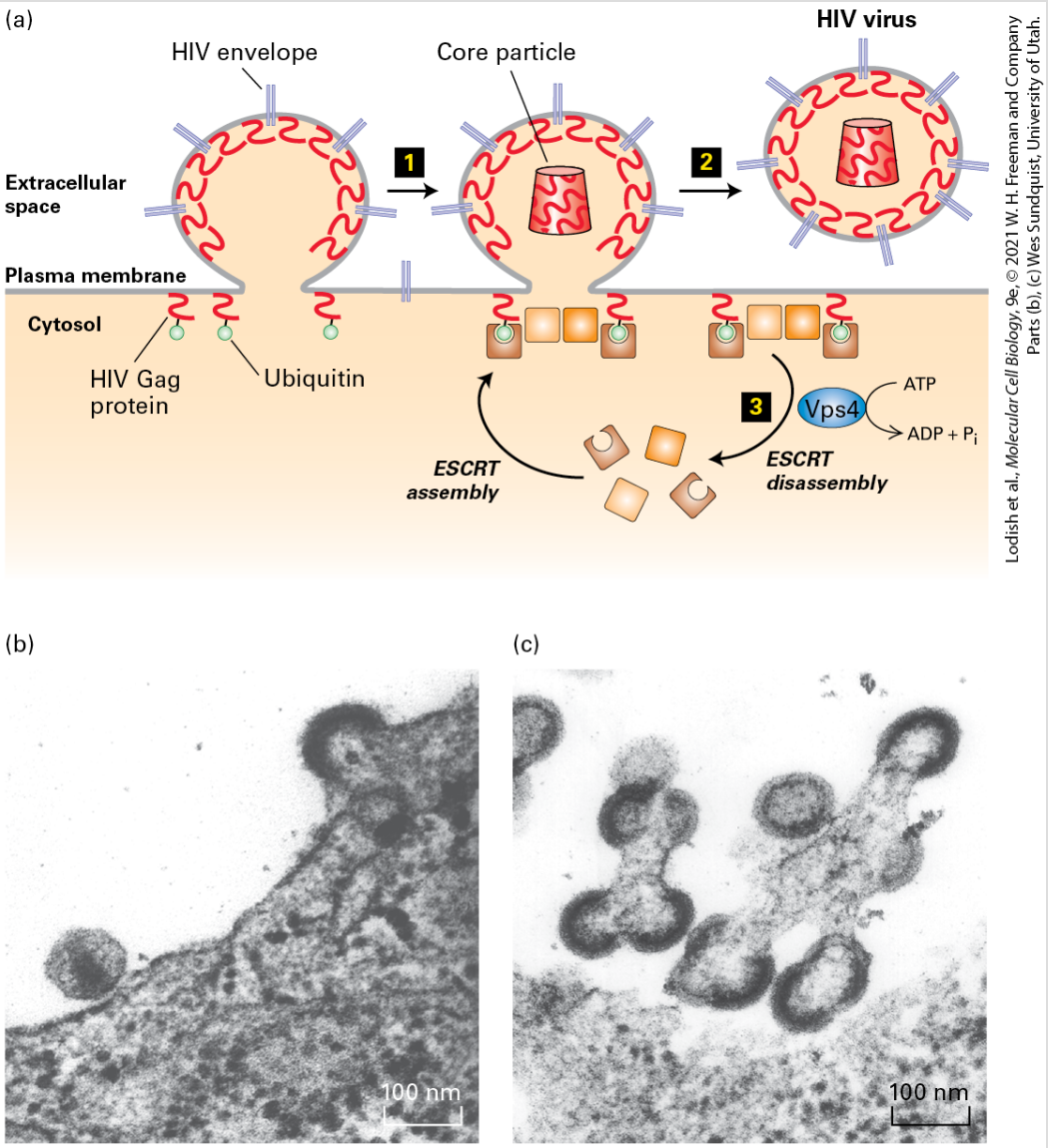
Retroviruses Bud from the Plasma Membrane by a Process Similar to Formation of Multivesicular Endosomes

The vesicles that bud into the interior of endosomes have a topology similar to that of enveloped virus particles that bud from the plasma membrane of virus-infected cells. Moreover, recent experiments have demonstrated that a common set of proteins is required for both types of membrane-budding events. In fact, the two processes so closely parallel each other in mechanistic detail as to suggest that enveloped viruses have evolved mechanisms to recruit the cellular proteins used in inward endosomal budding for their own purposes.

The human immunodeficiency virus (HIV) is an enveloped retrovirus that buds from the plasma membrane of infected cells in a process driven by the viral Gag protein, the major structural component of completed virus particles. Gag protein binds to the plasma membrane of an infected cell, and some 4000 Gag molecules polymerize into a spherical shell, producing a structure that looks like a vesicle bud protruding outward from the plasma membrane.

Mutational studies with HIV have revealed that the N-terminal segment of Gag protein is required for association with the plasma membrane, whereas the C-terminal segment is required for pinching off of complete HIV particles. For instance, if the portion of the viral genome encoding the C-terminus of Gag is removed, HIV buds will form in infected cells, but pinching off does not occur, and thus no free virus particles are released.

The first indication that HIV budding employs the same molecular machinery as vesicle budding into endosomes came from the observation that Tsg101, an ESCRT protein, binds to the C-terminus of Gag protein. Subsequent findings have clearly established the mechanistic parallels between the two processes. For example, Gag is ubiquitinylated as part of the process of virus budding, and in cells with mutations in Tsg101 or Vps4, HIV virus buds accumulate but cannot pinch off from the membrane (Figure 14-33). Moreover, when a segment from the cellular Hrs protein is added to a truncated Gag protein by construction of the appropriate hybrid gene, proper budding and release of virus particles is restored. Taken together, these results indicate that Gag protein mimics the function of Hrs, redirecting ESCRT proteins to the plasma membrane, where they can function in the budding of virus particles.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
Parts (b), (c) Wes Sundquist, University of Utah.

FIGURE 14-33 Mechanism for budding of HIV from the plasma membrane. Proteins required for the formation of multivesicular endosomes are exploited by HIV for virus budding from the plasma membrane. (a) Budding of HIV particles from HIV-infected cells occurs by a mechanism similar to that shown in Figure 14-32, using the virally encoded Gag protein and cellular ESCRT and Vps4 proteins (steps 1–3). Ubiquitinylated Gag located near a budding particle functions like Hrs. (b) In wild-type cells infected with HIV, virus particles bud from the plasma membrane and are rapidly released into the extracellular space. (c) In cells that lack the functional ESCRT protein Tsg101, the viral Gag protein forms dense viruslike structures, but budding of these structures from

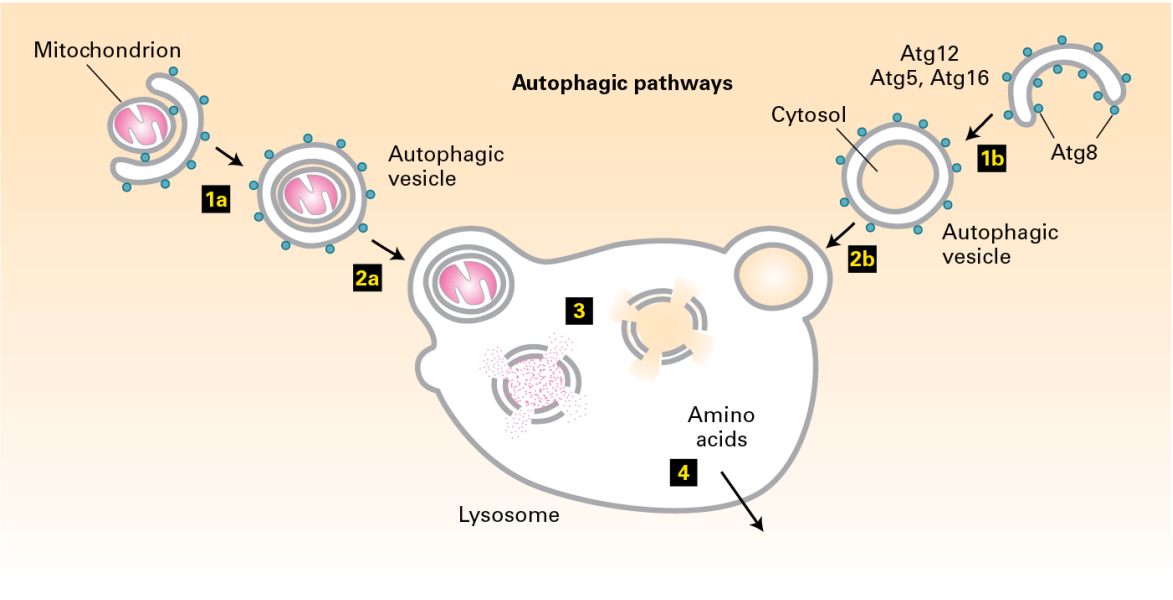
the plasma membrane cannot be completed, and chains of incomplete viral buds still attached to the plasma membrane accumulate.

Other enveloped retroviruses, such as murine leukemia virus and Rous sarcoma virus, have also been shown to require ESCRT complexes for their budding, although each virus appears to have evolved a somewhat different mechanism to recruit ESCRT complexes to the site of virus budding.

The Autophagic Pathway Delivers Cytosolic Proteins or Entire Organelles to Lysosomes

When cells are placed under stressful conditions, such as starvation, they have the capacity to recycle macromolecules for use as nutrients in a process of lysosomal degradation known as **autophagy** ("eating oneself"). The autophagic pathway begins with the formation of a flattened double-membrane cup-shaped structure that envelops a region of the cytosol or an entire organelle (e.g., a mitochondrion), forming an *autophagosome*, or *autophagic vesicle* (Figure 14-34). The outer membrane of an autophagosome can fuse with a lysosome, delivering a large vesicle, bounded by a single membrane bilayer, to the interior of the lysosome. Lipases and proteases within the lysosome degrade the autophagosome and its contents into their molecular components, just as they do when the

contents of multivesicular endosomes are delivered to the lysosome. Amino acid permeases in the lysosomal membrane then allow for the transport of free amino acids back into the cytosol for use in synthesis of new proteins.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 14-34 The autophagic pathway. The autophagic pathway allows cytosolic proteins and organelles to be delivered to the lysosome interior for degradation. In the autophagic pathway, a cup-shaped structure forms around a portion of the cytosol (*right*) or an organelle such as a mitochondrion (*left*). Continued addition of membrane eventually leads to the formation of an autophagosome that envelops its contents in two complete membranes (steps **1a** and **1b**). Fusion of the outer membrane with the membrane of a lysosome releases a single-membrane vesicle and its contents into the lysosome interior (steps **2a** and **2b**). After degradation of the protein and lipid components by hydrolases in the lysosome interior (step **3**), the released amino acids are transported across the lysosomal membrane into the cytosol (step **4**). Proteins known to participate in the autophagic pathway include Atg8, which forms a coat structure around the autophagosome.

By studying mutants with defects in the autophagic pathway, scientists have identified processes other than recycling of cellular

components during starvation that also depend on autophagy. Experiments carried out principally in *Drosophila* and mice have shown that autophagy participates in a type of quality-control mechanism that removes organelles that have ceased to function properly. In particular, the autophagic pathway can target dysfunctional mitochondria that have lost their integrity and no longer have an electrochemical gradient across their inner membrane. In certain cell types, pathogenic bacteria and viruses that are multiplying in the cytosol of host cells can be targeted to the autophagic pathway for destruction in the lysosome as part of a host defense mechanism against infection.

In each of these processes, and in all eukaryotic organisms, the autophagic pathway takes place in three basic steps. Although the mechanisms underlying each of these steps are relatively poorly understood, they are thought to be related to the basic mechanisms for vesicular trafficking discussed in this chapter.

Autophagosome Nucleation

The autophagosome is thought to originate from a fragment of a membrane-bounded organelle. The origin of this membrane has been difficult to trace because no known unique integral membrane proteins, which might serve to identify the source of this membrane, are known to be required for the formation of the autophagosome. Studies in yeast have shown that some mutants

defective in Golgi trafficking are also defective in autophagy, suggesting that the autophagosome is initially derived from a fragment of the Golgi. Autophagy that is induced by starvation appears to be a nonspecific process in which a random portion of the cytoplasm, including organelles, becomes enveloped by an autophagosome. In these cases, the site of nucleation is probably random. In cases in which organelles are enveloped by the autophagosome in response to nutrient starvation, a signal based on ubiquitin modification on the surface of the organelle targets nucleation of the autophagosome.

Autophagosome Growth and Completion

New membrane must be delivered to the autophagosome membrane in order for this cup-shaped organelle to grow. This growth probably occurs by the fusion of transport vesicles with the membrane of the autophagosome. About 30 proteins that participate in the formation of autophagosomes have been identified in genetic screens for yeast mutants that are defective in autophagy. One of these proteins is Atg8, shown in Figure 14-34, which is covalently linked to the lipid phosphatidylethanolamine and thus becomes attached to the cytoplasmic face of the autophagosome. Association of Atg8 with a membrane vesicle appears to be the key step in enabling a vesicle to fuse with the growing autophagosome.

Fusion of Atg8-containing vesicles with the autophagosome involves the formation of a cytosolic assembly of Atg12, Atg5, and Atg16.

Atg12 is similar in structure to ubiquitin, and a set of proteins related to ubiquitin-conjugating enzymes are responsible for covalently joining Atg12 to Atg5 by a process similar to that used for covalently joining ubiquitin to a target protein (see Figure 3-32). The covalently linked Atg12-Atg5 dimer then co-assembles with Atg16 to form a polymeric complex localized to the site of a growing autophagosome. By an unknown mechanism, this cytosolic complex is thought to bring about the fusion of Atg8-containing vesicles into a cup-shaped autophagosome.

Autophagosome Targeting and Fusion

The outer membrane of the completed autophagosome is thought to contain a set of proteins that target it for fusion with the membrane of a lysosome. Two vesicle-tethering proteins have been found to be required for autophagosome fusion with a lysosome, but the corresponding SNARE proteins have not been identified. Fusion of the autophagosome with the lysosome occurs after Atg8 has been released from the autophagosome membrane by proteolytic cleavage; this proteolysis step occurs only after the autophagosome has completely formed a sealed double-membrane system. Thus Atg8 protein appears to mask fusion proteins and to prevent premature fusion of the autophagosome with the lysosome.

KEY CONCEPTS OF SECTION 14.6

Directing Membrane Proteins and Cytosolic Materials to the Lysosome for Degradation

- Endocytosed membrane proteins destined for degradation in the lysosome are incorporated into vesicles that bud into the interior of the endosome. Multivesicular endosomes, which contain many of these internal vesicles, can fuse with the lysosome to deliver the vesicles to the interior of the lysosome (see Figure 14-31).
- Some of the cellular components (e.g., ESCRT) that mediate inward budding of endosomal membranes are used in the budding and pinching off of enveloped viruses, such as HIV, from the plasma membrane of virus-infected cells (see Figures 14-32 and 14-33).
- In the process of autophagy, a portion of the cytoplasm or an entire organelle (e.g., a mitochondrion) can be enveloped in a flattened membrane and eventually incorporated into a double-membrane autophagosome. Fusion of the outer vesicle membrane with the lysosome delivers the enveloped contents to the interior of the lysosome for degradation (see Figure 14-34).

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Analyze the Data
- Classic Experiment 14-1: Following a Protein out of the Cell
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

adapter protein (AP) complexes
anterograde
ARF protein
autophagy
cisterna
cisternal maturation
clathrin
constitutive secretion
COPI
COPII
dynamin

endocytic pathway
ESCRT proteins
exocytosis
late endosome
low-density lipoprotein (LDL)
mannose 6-phosphate (M6P)
multivesicular endosomes
Rab proteins
receptor-mediated endocytosis
regulated secretion
retrograde
retromer
secretion (sec) mutants
secretory pathway
sorting signals
transcytosis
trans-Golgi network
transport vesicles
t-SNAREs
v-SNAREs

Review the Concepts

1. What two modern methods allow the process of intercompartmental transport to be followed? What basic features do these experimental approaches to study protein transport have in common?

2. Vesicle budding is associated with coat proteins. What is the role of coat proteins in vesicle budding? How are coat proteins recruited to membranes? What kinds of molecules are likely to be included or excluded from newly formed vesicles? What is the best-known example of a protein likely to be involved in the pinching off of vesicles?
3. Treatment of cells with the drug *brefeldin A (BFA)* has the effect of causing Golgi membranes in the process of budding COPI vesicles to shed their vesicle coats. The result of BFA treatment is to cause the vast majority of Golgi proteins to be relocated in the ER. What inferences can be made from this observation regarding roles of coat proteins, other than promoting vesicle formation? Predict what type of mutation in ARF might have the same effect as treating cells with BFA.
4. Microinjection of an antibody known as *EAGE*, which reacts with the "hinge" region of the β subunit of COPI, causes accumulation of Golgi enzymes in transport vesicles and inhibits anterograde transport of newly synthesized vesicles from the ER to the plasma membrane. What effect does the antibody have on COPI activity? Explain the results.
5. Specificity in fusion between vesicles involves two discrete and sequential processes. Describe the first of the two processes and its regulation by GTPase switch proteins. What effect on the size of early endosomes might result from overexpression of a mutant form of Rab5 that is stuck in the GTP-bound state?

6. *Sec18* is a yeast gene that encodes NSF. Mutations of this gene produce class C mutants. What is the mechanistic role of NSF in membrane trafficking? And, as indicated by its class C phenotype, why does an NSF mutation produce accumulation of vesicles at what appears to be only one stage of the secretory pathway?
7. What feature of procollagen synthesis provided early evidence for the Golgi cisternal maturation model?
8. Sorting signals that cause retrograde transport of a protein in the secretory pathway are sometimes known as *retrieval sequences*. List the two known examples of retrieval sequences for soluble and membrane proteins of the ER. How does the presence of a retrieval sequence on a soluble ER protein result in its retrieval from the *cis*-Golgi complex? Describe how the concept of a retrieval sequence is essential to the cisternal-maturation model.
9. Clathrin adapter protein (AP) complexes bind directly to the cytosolic face of membrane proteins and also interact with clathrin. What are the four known adapter protein complexes? What observation regarding AP3 suggests that clathrin is an accessory protein to a core coat composed of adapter proteins?
10. I-cell disease is a classic example of an inherited human defect in protein targeting that affects an entire class of proteins: the soluble enzymes of the lysosome. What is the molecular defect in I-cell disease? Why does it affect the targeting of an entire class of proteins? What other types of mutations might produce the same phenotype?

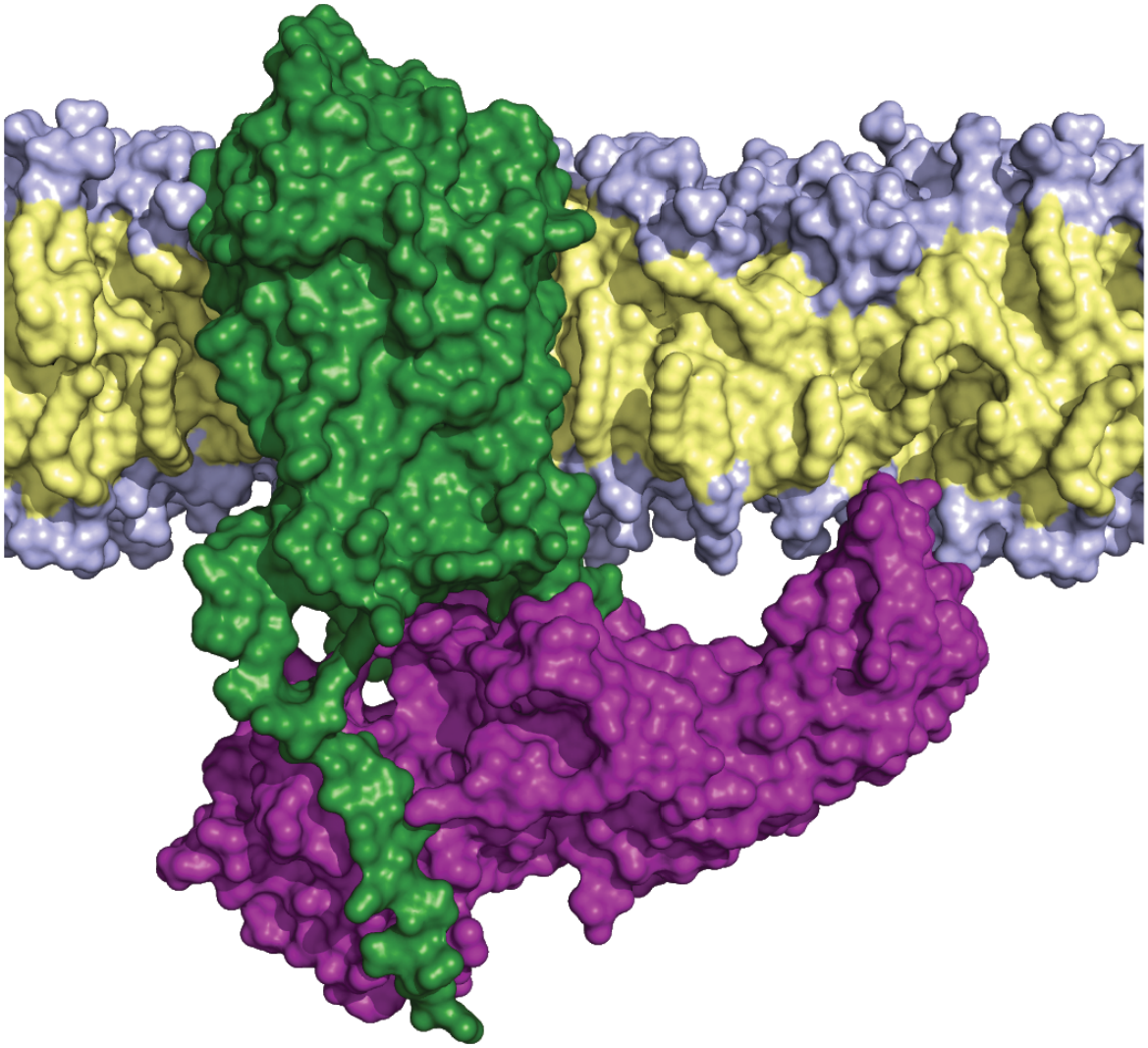
11. The *trans*-Golgi network is the site of multiple sorting processes as proteins and lipids exit the Golgi complex. Compare and contrast the sorting of proteins to lysosomes with the packaging of proteins into regulated secretory vesicles such as those containing insulin. Compare and contrast the sorting of proteins to the basolateral versus apical cell surfaces in MDCK cells and in hepatocytes.
12. What does the budding of influenza virus and vesicular stomatitis virus (VSV) from polarized MDCK cells reveal about the sorting of newly synthesized plasma membrane proteins to the apical or basolateral domains? Now consider the following result: a peptide with a sequence identical to that of the VSV G protein cytoplasmic domain inhibits targeting of the G protein to the basolateral surface and has no effect on HA targeting to the apical membrane, but a peptide in which the single tyrosine residue is mutated to an alanine has no effect on G protein basolateral targeting. What does this tell you about the sorting process?
13. Describe the role of pH in regulating the interaction between mannose 6-phosphate and the M6P receptor. Why does a rise in endosomal pH lead to the secretion of newly synthesized lysosomal enzymes into the extracellular medium?
14. What mechanistic features are shared by (a) the formation of multivesicular endosomes by budding into the interior of an endosome and (b) the outward budding of HIV virus at the cell surface? You wish to design a peptide

inhibitor/competitor of HIV budding and decide to mimic a portion of the HIV Gag protein in a synthetic peptide. Which portion of the HIV Gag protein would be a logical choice? What normal cellular process might this inhibitor block?

15. The endocytic and autophagic pathways serve two fundamentally different roles, but both deliver their vesicles to the lysosome. What are the fundamental differences between the two pathways? Describe the three basic steps in the formation and fusion of autophagosomes.
16. Describe the location and pH sensitivity of the receptor-ligand interaction in the LDL receptor-mediated endocytosis pathway.
17. What do mutations in the cytoplasmic domain of the LDL receptor that cause familial hypercholesterolemia reveal about the receptor-mediated endocytosis pathway?

CHAPTER 15

Receptors, Hormones, and Cell Signaling



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

Structure of a cell surface G protein–coupled receptor (green) bound to β-arrestin (purple). G protein–coupled receptors that are in the active state for a long period of time become

phosphorylated, triggering binding of an arrestin and inhibition of further signaling by the receptor.

[Data from Y. Kang et al., 2015, *Nature* **523**:561–567, PDB ID 4zwj, and custom PDB.]

OUTLINE

15.1 Signal Transduction Pathways: From Extracellular Signal to Cellular Response
15.2 Studying Cell-Surface Receptors and Signal Transduction Proteins
15.3 Structure and Mechanism of G Protein–Coupled Receptors
15.4 Regulating Metabolism of Many Cells: G Protein–Coupled Receptors That Activate or Inhibit Adenylyl Cyclase
15.5 Regulating Protein Secretion and Muscle Contraction: Ca^{2+} Ions as Second Messengers in Multiple Signal Transduction Pathways
15.6 Vision: How the Eye Senses Light

No cell lives in isolation; life requires that all cells sense stimuli in their environment and respond appropriately with adaptations that affect their growth, health, function, or development. Many cells sense physical stimuli such as mechanical pressure (touch) and heat (see Chapter 21). Phototactic single-celled algae sense light (see Chapter 1), and in this chapter we see how cells in the human retina sense light and respond by sending signals to the brain. Metazoan cells sense a wide variety of environmental chemicals, including nutrients such as sugars and amino acids as well as oxygen, toxic compounds, and diverse molecules that convey taste (tastants) and odor (odorants). Many cells sense a reduction in oxygen levels in the blood and respond by activating the HIF1α transcription factor which, in turn, activates expression of genes that enable the cells and the organism to adapt to low oxygen (see Chapter 21).

Many types of cells release particular chemicals that can influence the behavior of another cell; these are often called *extracellular* (outside of the cell) **signaling molecules**, or simply *signals*. The release and reception of such signals is a fundamental process, known as **cellular communication**, that shapes the development and function of every living organism. Cells that communicate through extracellular signals include single-cell prokaryotic and eukaryotic microorganisms — such as bacteria, yeasts, algae, slime molds, and protozoans. In Chapter 1 we saw how secreted molecules called *pheromones* coordinate the aggregation of free-living yeast cells for sexual mating.

This chapter and the next focuses on hormones and other extracellular signaling molecules that function *within* multicellular animals. These signaling molecules function to control a variety of processes, including the metabolism of sugars, fats, and amino acids; the induction or repression of specific genes that in turn affect the growth, differentiation, and function of cells and tissues; the synthesis and secretion of many proteins; and the composition of intracellular and extracellular fluids. For example, when the concentration of certain hormones rises, that informs cells that the organism is threatened (adrenaline and the fight-or-flight response) or that excess nutrients are present and should be stored for future use (insulin triggering the uptake of glucose into muscle for storage as glycogen).

The primary biological function of many of the molecules used as signals — hormones, pheromones, and some neurotransmitters — is the transfer of information. Some molecules used as signals have multiple functions. For example, some nutrients both act as signals and serve as fundamental building blocks of biomolecules (amino acids, sugars) or as sources of energy (oxygen, glucose, fatty acids).

The cell that receives a signal — termed the **target cell** — must be able to detect the signaling molecule. Typically, the signaling molecule binds noncovalently to a specific protein in the cell, typically called the **receptor**, that contains a binding site for the signaling molecule. The signaling molecule that binds a receptor is often called the **ligand** of the receptor. Receptors bind a single type of molecule or a group of closely related molecules.

After detecting a signaling molecule, the cell needs to generate a response. Upon binding their ligand, receptors undergo a conformational change that enables them to bind to other proteins in the cytosol or attached to a membrane, often the plasma membrane (Figure 15-1, steps **1** and **2**). In binding to these proteins, the receptor may activate (or inhibit) them, sometimes by enzymic modification, such as phosphorylation. Often there are multiple biochemical or biophysical steps that form a pathway from the initial detection of the signal by the receptor to the final (downstream) responses of the cell. The series of steps leading from initial detection to the final response is termed a **signal transduction pathway**. Signal transduction pathways ultimately

lead to activation (or in some cases inhibition) of one or more **effector proteins** — often enzymes or transcription factors or cytoskeletal proteins — that lead directly to changes in cellular activities (Figure 15-1, steps 3 – 5). Some signal transduction pathways contain just one or two intermediates; others can involve over a dozen.

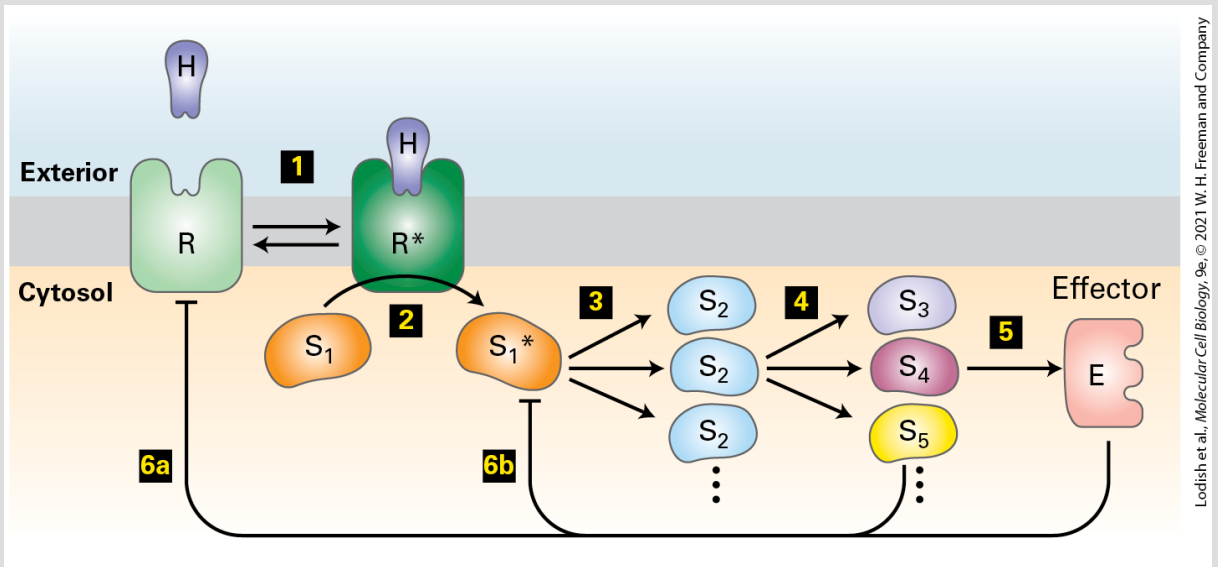


FIGURE 15-1 Signal transduction pathways. In this generalized pathway, a receptor protein R binds a ligand H (step 1), triggering a conformational change in the receptor (R^*) that enables it to bind to and activate a signal transduction protein S_1 (step 2). In many signaling pathways this is a GTP-binding protein or a protein kinase or phosphatase. S_1 in turn binds to and activates (or inhibits) other signal transduction proteins (step 3), S_2, that in turn activate other signaling proteins, here S_3, S_4, and S_5 (step 4). Some signaling proteins (here S_2) can activate multiple types of different downstream proteins. Certain signal transduction proteins bind and activate an effector protein, E (step 5), which can be an enzyme, transcription factor, transport protein, ion channel, or other type of protein. Signal amplification is a characteristic of many signal transduction pathways, in which activation of a single receptor molecule can lead to activation of multiple signal transduction and effector proteins (step 3). Virtually all signal transduction pathways have feedback controls whereby a protein in the signal transduction pathway (here S_5) or an effector protein modifies either the receptor (step 6a) or an early protein in the pathway

(step 6b), thus inhibiting or blocking an early step in that pathway. In many pathways these feedback controls trigger receptor degradation, lowering the numbers of receptors and thus the sensitivity of the cell to the ligand.

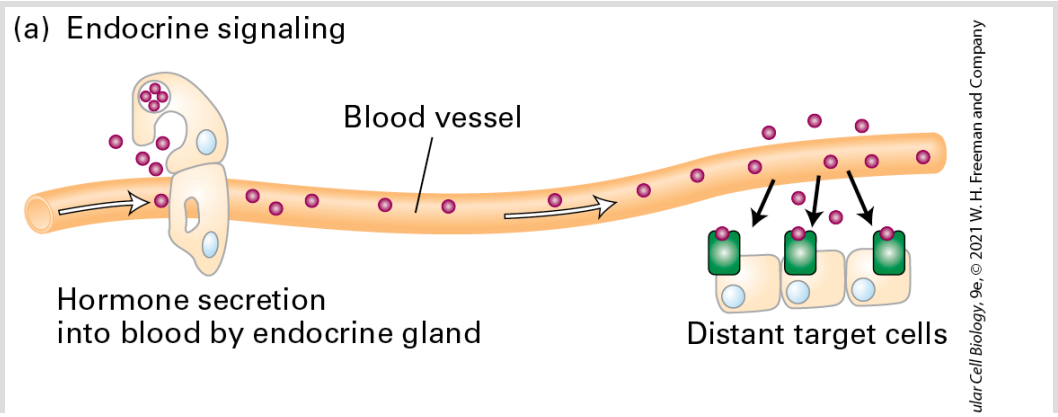
In any cell in multicellular organisms, multiple receptors are simultaneously activated by different extracellular signals, and each of these receptors in turn can activate a different signal transduction pathway. Complex information-processing networks help coordinate the cell's responses to these multiple signals and lead to an appropriate change in the cell's pattern of gene expression and metabolism.

In this chapter, we first discuss some basic properties of receptors and the signal transduction pathways they activate, then introduce key techniques for studying receptors and proteins in signaling pathways. We next focus on the largest and evolutionarily most ancient class of receptors — the **G protein–coupled receptors**. We will see several of the many important signal transduction pathways they activate and the functions these pathways play in biology.

15.1 Signal Transduction Pathways: From Extracellular Signal to Cellular Response

Signaling Molecules Can Act Locally or at a Distance

As noted above, cells respond to many different types of signals — some originating from outside the organism, some internally generated. Within a multicellular organism, signaling molecules that are generated internally can be described by how they reach their target. Some signaling molecules are transported long distances from their site of synthesis — for example, by movement through the blood; others have more local effects. In animals, signaling by extracellular molecules can be classified into three types based on the distance over which the signal acts (Figure 15-2a–c).



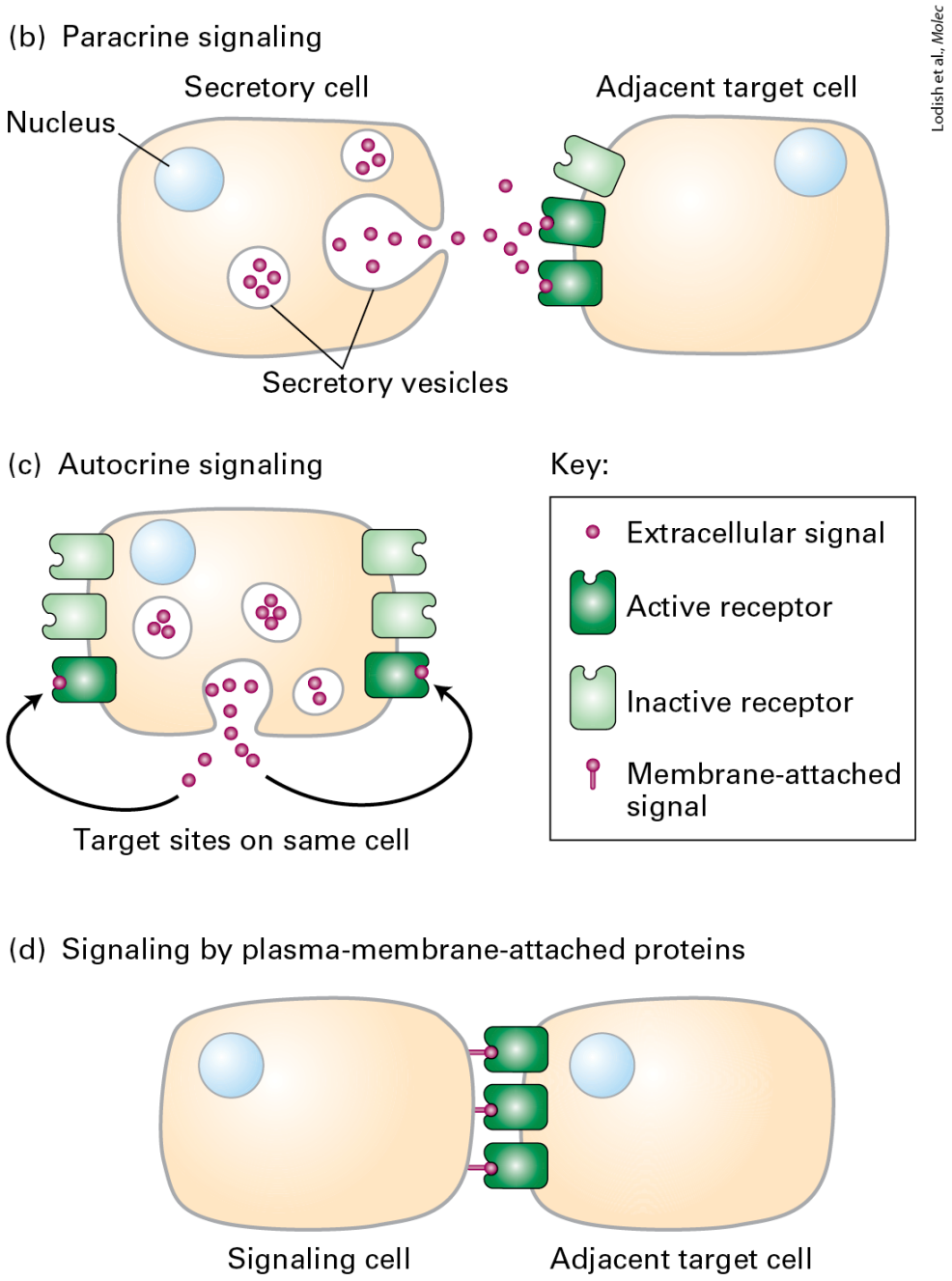


FIGURE 15-2 Types of extracellular signaling. (a–c) Cell-to-cell signaling by extracellular chemicals occurs over distances from a few micrometers in paracrine and autocrine signaling to several meters in endocrine signaling. (d) Proteins attached to the plasma membrane of one cell can interact directly with cell-surface receptors on adjacent cells.

In *endocrine* signaling, the signaling molecules act on target cells distant from their site of synthesis. The signaling molecules are synthesized and secreted by signaling cells (e.g., those found in endocrine glands) and transported through the circulatory system to the target cells. The term **hormone** generally refers to signaling molecules that mediate endocrine signaling. Insulin secreted by the pancreas and epinephrine secreted by the adrenal glands are examples of endocrine signals that travel through the blood, and thus they are termed *hormones*. Defects in endocrine signaling can result in human diseases — for example, defects in insulin signaling can cause diabetes.

In **paracrine** signaling, the signaling molecules released by a cell affect only those target cells in close proximity. An example is a neuron releasing a neurotransmitter that acts on an adjacent neuron. In addition to neurotransmitters, many of the protein **growth factors** that regulate development in multicellular organisms act at short range. Several of the developmentally important signaling proteins that we discuss in Chapter 16 diffuse away from the signaling cell, forming a concentration gradient, and induce different responses in nearby cells depending on their distances from the signaling cells and thus on the local concentration of the signaling protein. In some cases, paracrine signaling molecules, exemplified by TGF-β, are secreted by a cell and then trapped nearby in the web of extracellular macromolecules

termed the *extracellular matrix* until freed to bind to cell-surface receptors on a nearby cell (see Chapters 16 and 20).

In *autocrine* signaling, cells respond to substances that they themselves release. This type of signaling is particularly characteristic of tumor cells. Many tumor cells release growth factors that stimulate inappropriate, unregulated self-proliferation.

Some signaling molecules are integral membrane proteins located on the cell surface (Figure 15-2d). The targets of these relatively immobile membrane signals are receptors on the surface of adjacent cells, whose proliferation or differentiation are controlled by the signal. In other cases, proteolytic cleavage of a membrane-bound signaling protein releases the extracellular domain, which functions as a soluble signaling molecule that can act locally or at a distance (see Chapter 16).

Some signaling molecules can act at both short and long ranges. For example, epinephrine (also known as *adrenaline*) functions as a neurotransmitter (paracrine signaling) and as a hormone (endocrine signaling). As a hormone, it helps initiate the fight-or-flight response to a sudden danger in the environment discussed in Section 15.4.

Signal Transduction Pathways Can Produce Rapid, Short-Term or Slow, Long-Term Changes in Cells, or Both

The effector proteins in signal transduction pathways are often enzymes or transcription factors that induce two major types of cellular responses: rapid short-term changes in protein activities (seconds-to-minutes) or slower long-term changes (hours-to-days). Rapid changes are typically consequences of modifications to specific preexisting enzymes and other proteins that alter their activity or function. Changes to these proteins are often initiated by covalent modifications such as phosphorylation or ubiquitinylation or by binding of ions or molecules such as Ca^{2+} or cAMP. Such modifications can induce changes in cellular metabolism of sugars, amino acids, and lipids; can induce secretion of hormones; and can induce in nerve cells the electric signals called *action potentials* (Chapter 23). Slower, long-term changes are typically consequences of changes in gene expression, either activating or inhibiting synthesis of specific proteins. In this case, the effector proteins are usually transcription factors. Long-term changes in the amounts and activities of specific proteins within cells are critical for cell proliferation, cell differentiation, and organismal development.

Receptors Are Allosteric Proteins That Activate Signal Transduction Pathways

The interacting surfaces of a receptor and ligand are complementary (see Figure 2-12), held together by multiple weak, noncovalent forces, including ionic, van der Waals, and hydrophobic interactions. In most cases, the binding of the ligand to the receptor causes the receptor to change from an inactive conformation (R) to an active conformation (R*).

$$\mathrm{R} + \mathrm{L} \rightarrow \mathrm{RL} \rightarrow \mathrm{R^*L}$$

Only the active R* conformation of the receptor can trigger the subsequent steps in the signal transduction pathway (see Figure 15-1) because only it can bind to, and subsequently activate or inhibit, a protein that initiates the signaling pathway. The site on the receptor that interacts with the next molecule in the signal transduction pathway typically is not the site where the ligand binds. Thus receptors are considered to be *allosteric* proteins (see Chapter 3). As with all allosteric proteins, there is an equilibrium between the R and R* forms of the receptor.

$$\begin{aligned}
&\mathrm{R} = \text{inactive} \quad \mathrm{R^*} = \text{active} \\
&\text{No ligand } \mathrm{R} \rightleftarrows \mathrm{R^*} \\
&\text{With ligand } \mathrm{R} \rightleftarrows \mathrm{R^*}\ (shifts\ right)
\end{aligned}$$

In the absence of the ligand, the equilibrium is far to the left; most of the receptor is in the inactive R conformation. In the presence of ligand, the equilibrium is changed, such that more receptors are in the active R* conformation. In other words, the ligand binds much more tightly to the R* than the R conformation, thus pushing the equilibrium to the right and activating the signal transduction pathway.

Receptors Can Be in the Cytosol, Nucleus, or on the Cell Surface Membrane

How a signaling molecule gains access to its receptor's binding site depends on where the receptor is located, whether inside the cell or partially exposed to the extracellular fluid at the cell surface.

Extracellular signal molecules that bind to receptors inside the cell must be able to cross the plasma membrane. Some are hydrophobic and can spontaneously diffuse across the plasma membrane's bilayer without the need for transport proteins (see Chapter 11). Others, such as amino acids, are hydrophilic and, as discussed in Chapter 11, can enter the cell only with the aid of transporter proteins in the plasma membrane. Once inside the cell, these signaling molecules bind to receptors that usually are water-soluble

proteins located in the cytosol, the nucleus, or some other intracellular organelle (Figure 15-3).

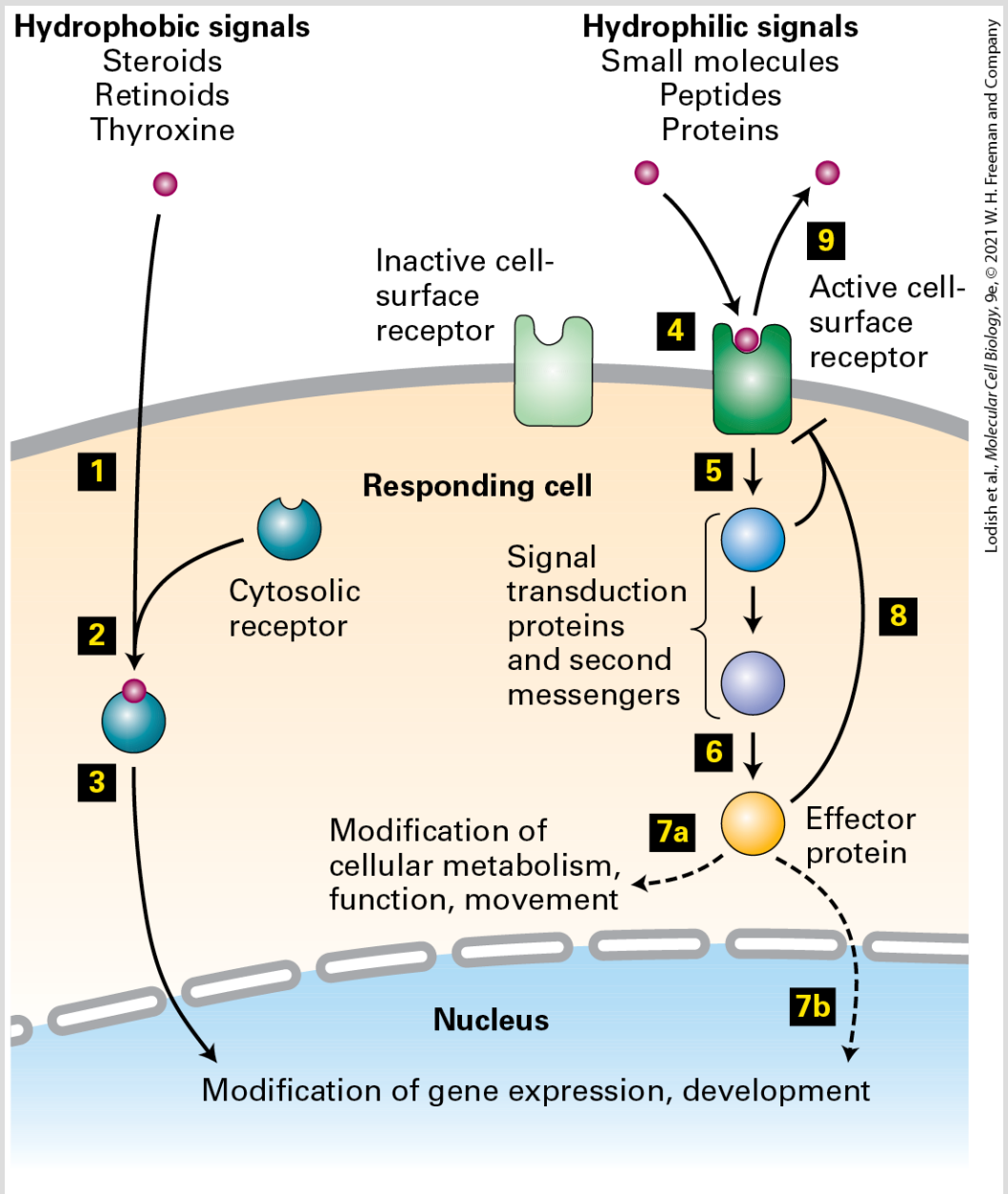


FIGURE 15-3 Overview of cell signaling. Hydrophobic signaling molecules, such as steroids and related molecules, diffuse through the plasma membrane (step 1) and bind to receptors in the cytosol (step 2). The receptor-signal complex moves into the nucleus (step 3), where it can bind to transcription-control regions in DNA and activate or repress gene expression. The majority of

signaling molecules, including small molecules (adrenaline, acetylcholine), peptides (yeast mating factors, glucagon), and proteins (insulin, growth hormone), are hydrophilic and cannot diffuse across the cell membrane. These molecules bind to specific cell-surface receptor proteins, triggering a conformational change in the receptor, thus activating it (step **4**). The activated receptor then activates one or more downstream signal transduction proteins or small-molecule second messengers (step **5**), which eventually leads to activation of one or more effector proteins (step **6**). The end result of a signaling cascade can be modification of specific cytosolic effector proteins, predominantly enzymes, leading to short-term changes in cellular function, metabolism, or movement (step **7a**). Alternatively, an effector protein can move into the nucleus, triggering a long-term change in gene expression (step **7b**). Termination or down-modulation of the cellular response is often caused by negative feedback from intracellular signaling molecules (step **8**) and by removal of the extracellular signal (step **9**).

One large class of hydrophobic signaling molecules spontaneously diffuses through the plasma membrane and binds to members of the nuclear receptor family of transcription factors. These signaling molecules are primarily steroids, retinoids, vitamin D, and thyroid hormone (thyroxine). In most cases, as we saw in Chapter 8, the receptor is located in the cytosol, and the receptor-hormone complex moves into the nucleus, binds to specific regulatory sequences in DNA, and activates or represses expression of specific target genes (see Figure 15-3). In other cases, the receptor is already positioned in the nucleus, bound to its corresponding enhancers or promoters on the DNA. The conformational change in the receptor induced by ligand binding permits transcriptional co-activators to bind and induce gene transcription.

In this chapter and the next, we focus on extracellular signaling molecules — small molecules, peptides, and proteins — that initiate signal transduction pathways without crossing the plasma membrane. Indeed, they are frequently too large and too hydrophilic to diffuse through the plasma membrane. These bind to **cell-surface receptors** that are integral membrane proteins embedded in the plasma membrane (see Figure 15-3). Cell-surface receptors generally consist of three discrete topological domains, or segments: an extracellular domain facing the extracellular fluid, a plasma-membrane-spanning (transmembrane) domain, and an intracellular domain facing the cytosol.

The signaling molecule binds to a structurally complementary site of the receptor that is accessible from the extracellular fluid. The binding site may be located on the extracellular domain or within the membrane-spanning domain of the receptor. Binding of the ligand induces a conformational change in the receptor that is transmitted through the membrane-spanning domain to the cytosolic domain. This allosteric change results in the receptor's binding to, and subsequently activating or inhibiting, other proteins in the cytosol or attached to the plasma membrane. Thus the information that an extracellular signal is present is *transduced* across the plasma membrane and delivered to the cell interior.

In eukaryotes, there are about a dozen classes of cell-surface receptors, which activate several types of intracellular signal

transduction pathways. Our knowledge of signal transduction has advanced greatly in recent years, in part because these receptors and pathways are highly conserved and function in essentially the same way in organisms as diverse as worms, flies, mice, and humans. Studies using one organism often provide insights into signal transduction pathways in other organisms. For example, information gleaned from genetic studies in flies or worms can be combined with the information from biochemical analyses of mammalian cells to enable researchers to trace many entire signaling pathways from binding of ligand to final cellular response.

Most Receptors Bind Only a Single Type of Ligand or a Group of Closely Related Ligands

Like an enzyme binding to a substrate, each type of receptor binds only a single type of molecule or a group of very closely related ones. For example, the growth factor receptor binds growth factor, but not insulin or other protein hormones (Figure 15-4). Testosterone receptors bind the steroid testosterone but not other steroid hormones such as estrogen. The *binding specificity* of a receptor refers to its ability to bind or not bind closely related substances.

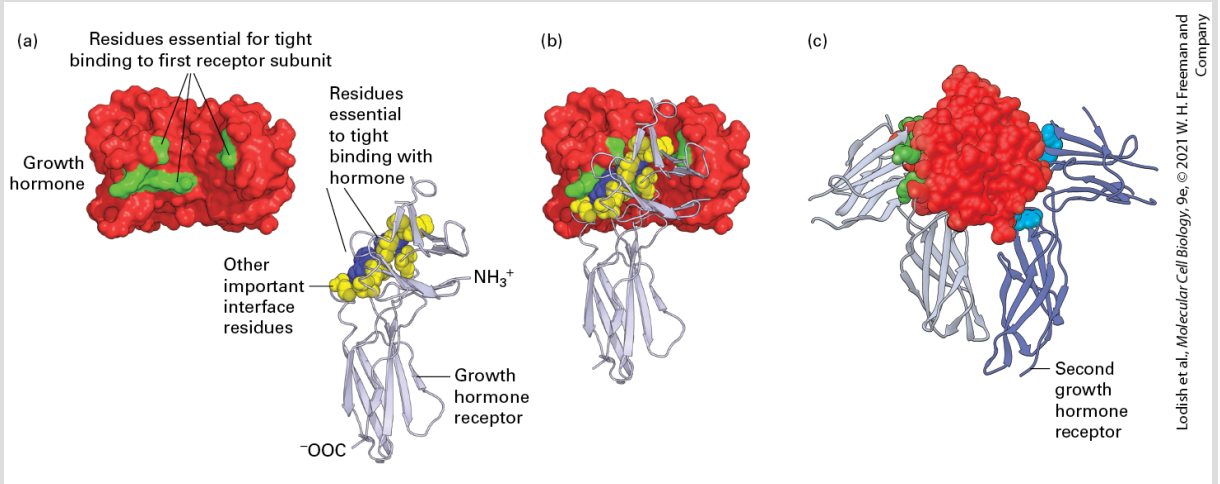


FIGURE 15-4 A single growth hormone protein binds simultaneously to two growth hormone receptors through multiple weak, noncovalent forces. (a) As determined from the three-dimensional structure of the 1 growth hormone: 2 growth hormone receptor complex, 28 amino acids in the hormone are at the binding interface with the first receptor protein. To determine which amino acids are important in ligand-receptor binding, researchers mutated each of these amino acids one at a time to alanine and measured the effect on receptor binding. From this study, it was found that only eight amino acids on growth hormone (green) contribute 85 percent of the energy that is responsible for tight receptor binding; these amino acids are distant from each other in the primary sequence but adjacent in the folded protein. Similar studies showed that two tryptophan residues (blue) in the receptor contribute most of the energy responsible for tight binding of growth hormone, although other amino acids at the interface with the hormone (yellow) are also important. (b) Binding of growth hormone to one receptor molecule is followed by (c) binding of a second receptor (purple) to the opposing side of the hormone; this involves the same set of yellow and blue amino acids on the receptor but different residues (light blue) on the hormone.

[Data from B. Cunningham and J. Wells, 1993, *J. Mol. Biol*. **234**:554; and T. Clackson and J. Wells, 1995, *Science* **267**:383.]

Organisms have evolved to be able to use a single ligand to stimulate different cell types to respond in distinct ways. Different cell types often have different receptors that bind the same ligand, and

activation of each receptor type induces a different intracellular signal transduction pathway. For instance, the surfaces of skeletal muscle cells, heart muscle cells, and the pancreatic acinar cells that produce hydrolytic digestive enzymes each have different types of receptors for acetylcholine. In a skeletal muscle cell, release of acetylcholine from a motor neuron innervating the cell (paracrine signaling) triggers muscle contraction by activating an ion channel. In heart muscle, the release of acetylcholine by certain neurons activates a G protein–coupled receptor that through a signal transduction pathway opens a K^+ channel, reducing the rate of contraction and lowering the heart rate. Acetylcholine binding to a receptor on pancreatic acinar cells induces a rise in the cytosolic concentration of Ca^{2+} that triggers secretion of digestive enzymes to facilitate digestion of a meal. Thus the activation by acetylcholine of different types of acetylcholine receptors produces different cellular responses, depending on the receptor and cell type.

Sometimes, the same receptor found on different cell types can, in response to binding the ligand, activate different signal transduction pathways leading to activation of different effectors. For example, the same receptor for the hormone epinephrine (the β-adrenergic receptor) is found on liver, muscle, and fat (adipose) cells. As we see in Section 15.4, epinephrine binding to its receptor stimulates depolymerization of glycogen to glucose in the first two cell types but hydrolysis and secretion of stored fat in adipose cells. In these ways, the same ligand can induce different cells to respond in a

variety of ways, often in a manner that coordinates the overall response of the organism. This property is known as the *effector specificity* of the receptor-ligand complex.

Most Receptors Bind Their Ligands with High Affinity

Binding of a single ligand to a receptor usually can be viewed as a simple reversible reaction, where the receptor is represented as R, the ligand as L, and the activated receptor-ligand complex as R*L:

$$\mathrm{R} + \mathrm{L} \quad \underset{k_{\mathrm{on}}}{\overset{k_{\mathrm{off}}}{\leftrightarrows}} \quad \mathrm{R}^{*}\mathrm{L} \tag{15-1}$$

where k_{on} is the kinetic rate constant for formation of a receptor-ligand complex from free ligand and receptor and k_{off} is the rate constant for dissociation of a ligand from its receptor. We define [R] and [L] as the concentrations of free receptor (i.e., receptor without bound ligand) and ligand, respectively, and [R*L] as the concentration of the receptor-ligand complex. At equilibrium, the rate of formation of the receptor-ligand complex, $[\mathrm{R}][\mathrm{L}]\ k_{\mathrm{on}}$, is equal to the rate of its dissociation, $[\mathrm{R}^{*}\mathrm{L}]\ k_{\mathrm{off}}$: thus at equilibrium,

$$[\mathrm{R}]\ [\mathrm{L}]\ k_{\mathrm{on}} = [\mathrm{R}^{*}\mathrm{L}]\ k_{\mathrm{off}}$$

The affinity of a signal molecule's binding to its receptor is typically quantified by the **dissociation constant** for binding (K_d). The dissociation constant K_d is the ratio of the two rate constants, $K_\mathrm{d} = k_\mathrm{off}/k_\mathrm{on}$. Equivalently, at equilibrium we can write this equation as:

$$K_\mathrm{d} = \frac{[\mathrm{R}][\mathrm{L}]}{[\mathrm{R}^*\mathrm{L}]} \tag{15-2}$$

The lower the value of the dissociation constant K_d, the tighter the binding (see Chapter 2 and below) and the more *stable* the receptor-ligand complex.

A way of seeing this key point is that when the concentration of ligand equals K_d (when $[\mathrm{L}] = K_\mathrm{d}$), then the concentration of free receptor [R] must equal the concentration of the receptor-ligand complex $[\mathrm{R}^*\mathrm{L}]$. Since the total amount of receptor equals the sum of the free and ligand-bound receptors:

$$[\mathrm{R_T}] = [\mathrm{R}^*\mathrm{L}] + [\mathrm{R}] \tag{15-3}$$

when the system is at equilibrium, half of the receptors have a ligand bound. The smaller the K_d, the lower the ligand concentration required to bind 50 percent of the cell-surface

receptors. A common method for experimentally determining the K_d of ligand/receptor binding is described in Section 15.2. The K_d for a binding reaction is similar to the Michaelis constant, K_m, which reflects the affinity of an enzyme for its substrate (see Chapter 3).

Signaling systems have evolved such that when an extracellular signal is not present or is present at only very low concentrations, very little receptor is bound to the ligand ([R*L] is low), and the cells are in what is called the *basal* or *unstimulated state*. A rise in the concentration of an extracellular signaling molecule, as when hormones are released into the blood, usually induces a proportional response in the target cell as more and more cellular receptors bind to their ligand and switch to their active conformations.

We can see this principle in practice by comparing the fraction of receptors bound by the hormone insulin at different levels of insulin present in the blood. The K_d for binding of insulin to its receptor on liver cells is 1.4×10^{-10} M (0.14 nM). The normal concentration of insulin in the blood is about 5×10^{-12} M. By substituting this value of L and the K_d into Equation 15-2, we can calculate the fraction of total insulin receptors with bound insulin, $[R^*L]/[R_T]$ at equilibrium as 0.0344; that is, about 3 percent of the total insulin receptors will have insulin bound to them in this unstimulated state. Ingesting glucose during a meal induces insulin secretion from the

pancreas (see Chapter 21). If the insulin concentration rises fivefold to 2.5×10^{-11} M, the number of receptor-hormone complexes will rise proportionately, almost fivefold, and about 15 percent of the total receptors will have insulin bound to them. If the extent of the induced cellular response is proportional to the number of active insulin-receptor complexes [RL], as is often the case, then the cellular response (changes in the activities of proteins and genes in the signal transduction pathway) will also increase by about fivefold.

We now turn our attention to events in the signal transduction pathway after the receptor is activated, focusing on the small molecules and proteins that most commonly participate in signal transduction pathways.

Second Messengers Are Used in Most Signal Transduction Pathways

Many signal transduction pathways include a step in which specific small molecules such as cyclic AMP and cyclic GMP are synthesized or in which the cytosolic concentration of an ion such as Ca^{2+} is increased. Such small molecules and ions in signaling pathways are often called **second messengers,** the first messenger being the extracellular signaling molecule itself (Figure 15-5). These second messenger molecules and ions then bind to and activate (or sometimes inhibit) specific intracellular proteins. Both the change

in concentration of a second messenger and its action on another protein are key steps in the pathway that ultimately results in a change in the physiological state of the cell in response to the initial signal/stimulus (first messenger). As an example, increases in the concentration of Ca^{2+} in the cytosol can in different cells trigger protein secretion, cell movement, or muscle contraction.

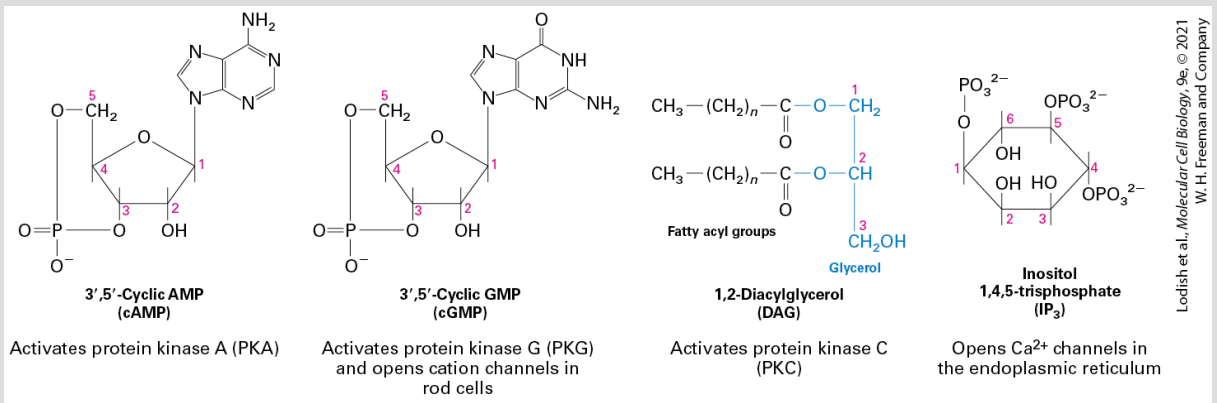


FIGURE 15-5 Four common compounds that function as intracellular second messengers. The major direct effect or effects of each compound are indicated below its structural formula. Calcium ions (Ca^{2+}) and several membrane-bound phosphatidylinositol derivatives also act as second messengers.

Protein Kinases and Phosphatases Participate in Signal Transduction Pathways by Covalently Modifying and Thus Activating or Inhibiting a Wide

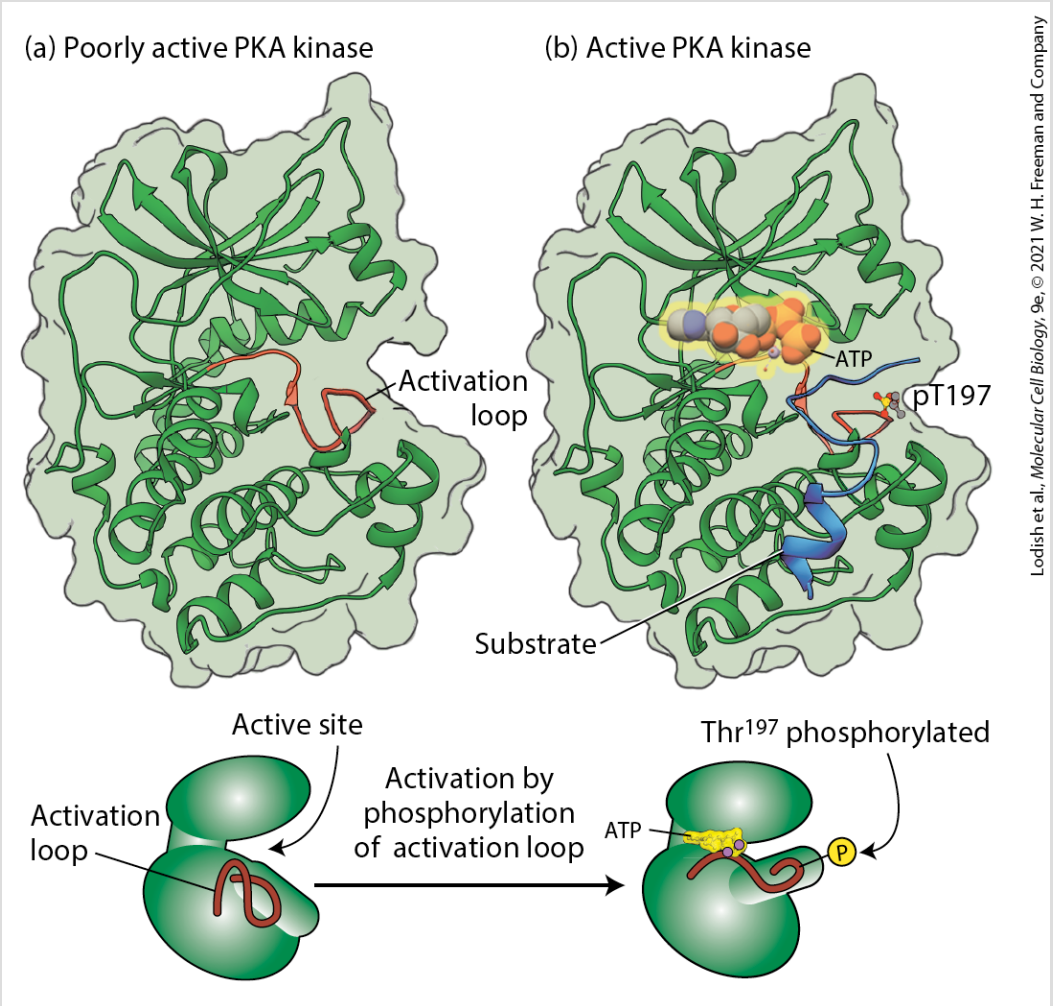
Variety of Proteins That Control Cellular States

Commonly in a signal transduction pathway, proteins are activated or deactivated at multiple steps. One of the most common and reversible mechanisms for changing the activity of a protein is phosphorylation. **Protein kinases (PKs)** add phosphate groups to specific residues of target proteins, while **protein phosphatases** remove phosphate groups from these residues (see Figure 3-36). Depending on the target protein, phosphorylation can change its enzymatic activity or its ability to bind to other proteins and thus plays a critically important role in determining the functional state of a cell. A typical mammalian cell expresses a hundred or more different types of kinases, each of which can phosphorylate — and thus affect the activity of — multiple target proteins.

The catalytic domains of all known protein kinases have a similar three-dimensional structure (see Figure 3-37) and highly conserved amino acids cluster around the catalytic site. Some form the *nucleotide-binding pocket* that binds ATP, positioning it such that its gamma phosphate can be transferred to the OH moiety of its substrate (see Figure 3-38). As detailed in Chapter 3, all kinases recognize their specific protein substrates by binding not only to the side chain of the serine, threonine, or tyrosine to be phosphorylated but also to specific amino acid residues that surround the target residue. Thus one can analyze the amino acid sequences

surrounding tyrosine, serine, and threonine residues in a protein and make a good prediction as to which kinases might phosphorylate those residues.

The catalytic activity of all kinases is highly regulated; commonly it is modulated by the binding of other proteins to the kinase and by changes in the concentrations of various small intracellular signaling molecules and metabolites. Many kinases are themselves activated by phosphorylation, resulting in cascades of kinases, one phosphorylating the next in the cascade. An example is protein kinase A (PKA), a kinase that phosphorylates proteins that regulate many aspects of cell metabolism (see Section 15-4). In its inactive, unphosphorylated form, critical residues in the activation loop are in the wrong position to bind ATP and its protein substrates (see Figures 3-38 and 15-6a). When a critical threonine residue in the activation loop is phosphorylated by PKA or by another kinase, the conformation of the loop changes, creating the binding site for ATP and its substrates, thus activating catalytic activity (Figure 15-6b). In this chapter and the next we see that many other kinases are activated in a similar manner.



EXPERIMENTAL FIGURE 15-6 Molecular structures of protein kinase A (PKA) in its catalytically inactive, unphosphorylated form (a) and after activation by phosphorylation of a threonine residue in the activation loop (b). Phosphorylation of PKA at threonine-197 (pT197) leads to a marked conformational change in the activation loop, a change that promotes binding of both ATP and its substrate proteins. A similar phosphorylation-dependent mechanism activates many other kinases (see Figure 3-38). Phosphorylation of threonine-197 is also required for binding of PKA to its inhibitory subunit R (see Figure 15-21a) and thus for its regulation by cAMP (see Section 15.4). PKA can autophosphorylate itself on T197 and other cellular kinases can also autophosphorylate this residue.

[Data from J.M. Steichen et.al., *J. Biol. Chem*. **287**:14672.]

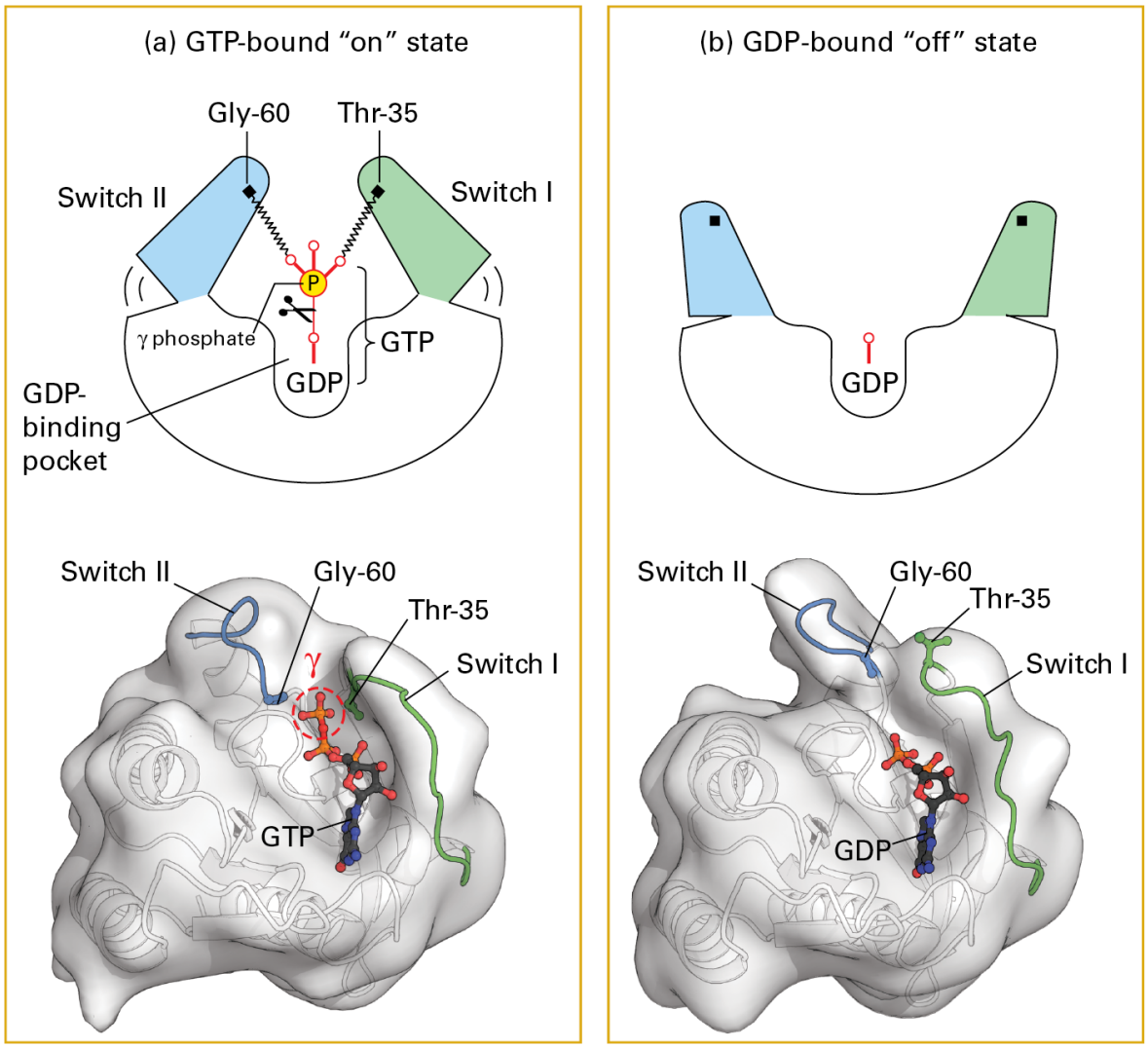
Many proteins are substrates for multiple kinases, each of which usually phosphorylates different amino acids in the protein. Each phosphorylation event has the potential to modify the activity of a particular target protein in a different way, some activating its function, others inhibiting it. An example we will encounter later is glycogen phosphorylase kinase, a key regulatory enzyme in glucose metabolism (see Figure 15-22). When activated by a signal transduction pathway, this enzyme phosphorylates and thereby activates the enzyme that breaks down glycogen into glucose. In many cases, addition of a phosphate group to an amino acid creates a binding surface that allows a second protein to bind; in the following chapter and Chapter 19 we will encounter many examples of kinase-driven assembly of multiprotein complexes.

Importantly, the activity of all protein kinases is opposed by the activity of protein phosphatases, some of which themselves are regulated by extracellular signals. Thus the activity of a protein in a cell can be a complex function of the activities of multiple kinases and phosphatases that act on it. Several examples of this phenomenon that occur in regulation of the cell cycle are described in Chapter 19.

GTP-Binding Proteins Are Frequently Used in Signal Transduction Pathways as On/Off Switches

Many signal transduction pathways contain members of the **GTPase superfamily of proteins**, which are found in all prokaryotic and eukaryotic cells. All of these proteins bind GTP and all can be turned off and on like a molecular switch. They exist in two forms (see Figure 3-35): (1) an active ("on") form with bound GTP (guanosine triphosphate) and (2) an inactive ("off") form with bound GDP (guanosine diphosphate). The GTP-bound "on" form binds to and modulates the activity of specific target proteins, whereas the "off" form cannot affect the activity of target proteins.

Two large classes of GTPase switch proteins are used in signal transduction pathways. **Heterotrimeric G proteins** directly bind to and are activated by certain cell-surface receptors. As we see in Section 15.3, G protein–coupled receptors function as guanine nucleotide exchange factors (GEFs): that is, they activate the heterotrimeric G protein to which they are coupled by triggering its release of GDP. The G protein then spontaneously binds GTP, favored by the high intracellular concentration of GTP relative to that of GDP. Binding of GTP induces a conformational change in the G protein to the active, "on" form (see Figure 3-35). The principal conformational changes involve two highly conserved segments of the GTP-binding protein, termed *switch I* and *switch II*, that allow the protein to bind to and activate downstream signaling proteins (Figure 15-7).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 15-7 Switching mechanism of monomeric G proteins. The ability of a G protein to interact with other proteins and thus transduce a signal differs between the GTP-bound "on" state and GDP-bound "off" state. (a) In the active "on" state, two domains, termed *switch I* (green) and *switch II* (blue), are bound to the terminal γ phosphate of GTP through interactions with the backbone amide groups of conserved threonine and glycine residues. When bound to GTP in this way, the two switch domains are in a conformation such that they can bind to and thus activate specific downstream effector proteins. (b) Removal of the γ phosphate by GTPase-catalyzed hydrolysis causes switch I and switch II to relax into a different conformation, the inactive "off" state; in this state, they are unable to bind to effector proteins. The ribbon models shown here represent both conformations of Ras, a monomeric G protein. A similar spring-loaded mechanism switches the alpha subunit in

heterotrimeric G proteins between the active and inactive conformations by movement of three, rather than two, switch segments.

[Data from I. Vetter and A. Wittinghofer, 2001, *Science* **294**:1299.]

After a brief time, the active G protein is shut off, converting back to the inactive form. A GTPase, which is often part of the GTP-binding protein itself, slowly hydrolyzes the bound GTP to GDP and P_i, thereby altering the conformation of the switch I and switch II segments so that they are unable to bind to the target protein.

The rate of GTP hydrolysis determines the length of time the GTP-binding protein remains in the active conformation and is able to signal its downstream target proteins: the slower the rate of GTP hydrolysis, the longer the protein remains in the active, GTP-bound state. The rate of GTP hydrolysis is often modulated by other proteins. For instance, both *GTPase-activating proteins (GAPs*) and *regulators of G protein signaling (RGSs)* accelerate GTP hydrolysis (see Figure 3-35). Many regulators of G protein activity are themselves controlled by extracellular signals.

Monomeric (often called *low-molecular-weight*) **G proteins**, including Ras and various Ras-like proteins such as Ran and Sar, do not directly bind to receptors but act as intermediate proteins in signal transduction pathways (Chapter 16). Monomeric G proteins play crucial roles in many pathways that regulate cell division and cell

motility. For this reason, mutations in genes encoding these G proteins frequently lead to cancer.

Signal Amplification and Feedback Repression Characterize Most Signal Transduction Pathways

Two essential features of most signal transduction pathways are *signal amplification* and *feedback repression*. In amplification (see Figure 15-1, step **3**), a single signaling molecule binding to one receptor protein can substantially alter the activity of a biochemical pathway within a cell. For instance, we see later in the chapter that a single photon absorbed by a photoreceptor in a rod cell in the eye (the physical stimulus) produces a measurable change in the membrane electrical potential. The process that enhances the effect of a signaling molecule is called *amplification*.

In amplification, the activation of a single receptor molecule can lead to activation of more than one molecule of a downstream intermediate protein in the signal transduction pathway. Each intermediate protein can also activate multiple proteins downstream of it, resulting in a cascade of activation. The extent of amplification depends on the number of steps in the signal transduction cascade and on the relative concentrations of the various components. As discussed later in the chapter, we are able to

detect a flash of as few as five photons because the signal transduction pathway amplifies the signal from a few photons enough to change the activities of Na^+ and Ca^{2+} ion channels responsible for the change in membrane potential. The ability to respond to a single photon makes night vision possible.

Cells must be able to suppress or even turn off a signal transduction pathway once the cell has adequately responded to the signal or the signal has been removed. Virtually all signal transduction pathways have feedback controls that will shut down the pathway: once activated, an effector protein modifies either the receptor or one of the proteins that operates in the early stages of the signal transduction pathway, thus blocking the downstream cellular responses (see Figure 15-1, steps **6a** and **6b**). In many pathways, these feedback controls trigger loss of receptor activity or even receptor degradation, lowering the number of functional receptors. The cell thus becomes much less sensitive to the ligand. Such feedback responses are sometimes referred to as *adaptation*.

We next describe how cell-surface receptors and signal transduction proteins are identified and characterized biochemically, focusing on components of signal transduction pathways that have been conserved through evolution. We then turn to an in-depth discussion of a very large and evolutionarily conserved class of receptors found in organisms from fungi to humans: the G protein–coupled receptors. Other large classes of cell-surface receptors and

the intracellular signaling pathways they activate are explored in Chapter 16.

KEY CONCEPTS OF SECTION 15.1

Signal Transduction Pathways: From Extracellular Signal to Cellular Response

- All cells communicate through extracellular signals. In unicellular organisms, extracellular signaling molecules regulate interactions between individuals, while in multicellular organisms, they mainly regulate physiology and development.
- Extracellular signals include membrane-anchored and secreted proteins or peptides (e.g., vasopressin and insulin), small hydrophobic molecules (e.g., steroid hormones and thyroxine), small hydrophilic molecules (e.g., epinephrine), gases (e.g., O_2, nitric oxide), and physical stimuli (e.g., light, touch).
- Hydrophobic signaling molecules interact with cytosolic or nuclear receptors and mainly affect gene expression.
- Binding of a hydrophilic extracellular signaling molecule to a cell-surface receptor triggers a conformational change in the receptor, which in turn leads to activation of intracellular signal transduction pathways that ultimately modulate cellular metabolism, function, or gene expression.
- Signals from one cell act on distant cells in endocrine signaling, on nearby cells in paracrine signaling, or on the signaling cell itself in autocrine signaling.
- Most receptors bind only a single type of ligand or a group of closely related ligands and with high affinity.
- Second messengers such as cAMP and Ca^{2+} are found in many signal transduction pathways.
- Protein kinases, protein phosphatases, and GTP-binding switch proteins participate in most signal transduction pathways.
- Signal transduction pathways allow amplification of an extracellular signal, enabling activation of a relatively small number of receptors to trigger major changes in cell metabolism, movements, or gene expression.
- Virtually all signal transduction pathways have feedback controls whereby an effector protein modifies and inhibits an early protein in the pathway, blocking an early step in that pathway.

15.2 Studying Cell-Surface Receptors and Signal Transduction Proteins

How a cell responds to external signals depends on the cell's complement of receptors that bind the signals as well as on the signal transduction pathways activated by these receptors. In this section, we examine experimental techniques used to characterize receptors and other signal transduction proteins. Many of these methods are also used to characterize the receptors that mediate endocytosis (see Chapter 14) or cell adhesion (see Chapter 20). We conclude the section with a discussion of techniques commonly used to measure the activity of signal transduction pathway components, such as kinases and GTP-binding switch proteins, and to study the mechanisms that regulate these signaling proteins.

Binding Assays Are Used to Detect Receptors and Determine Their Affinity and Specificity for Ligands

An experiment called a *binding assay* is used to quantify the number of receptors on a cell surface and determine how tightly they bind to a ligand (K_d). Receptors (R) are usually detected and quantified by their ability to bind ligands (L) that have been added to the fluid

surrounding intact cells, cell fragments, or purified receptors. In many experiments, the ligands are radioactively or fluorescently labeled to permit easy measurement of their binding. The ligand and receptor are allowed to incubate together until they reach equilibrium binding. Then unbound ligand is removed and the amount of ligand bound to receptors (RL) is measured — for example, by putting the sample in a machine that can detect the amount of radioactivity or fluorescence. One often measures the amount of ligand bound to the receptor at different ligand concentrations. A plot of the amount of ligand bound to the receptor [RL] versus the amount of free ligand [L], called a *ligand-binding* or *saturation curve*, can be seen in the red line in Figure 15-8. Such a plot generally follows Equation 15-4, which is simply an algebraic transformation of Equations 15-2 and 15-3.

$$[\mathrm{RL}] = \frac{R_T}{\frac{K_\mathrm{d}}{[\mathrm{L}]} + 1} \tag{15-4}$$

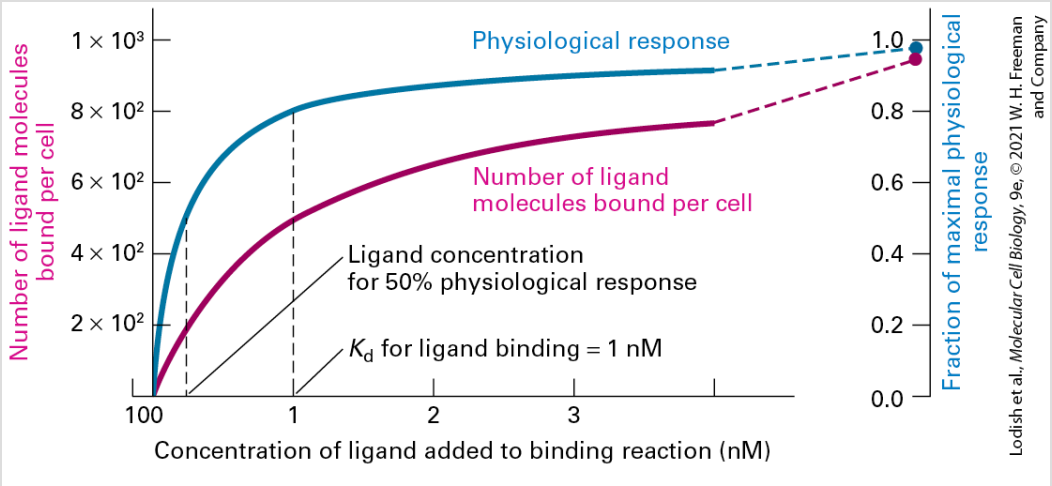


FIGURE 15-8 Binding assays determine the K_d and the number of receptors per cell, but the maximal physiological response to an external signal occurs when only a fraction of the receptors are occupied by ligand. In a typical experiment to determine the affinity of a receptor for a ligand, radioactively or otherwise labeled ligand is incubated with control cells that do not express the receptor of interest and with the same cells that have been altered by recombinant DNA techniques to express the receptor on their surface. Incubation is generally for an hour at 4 °C; the low temperature is used to prevent endocytosis of the cell-surface receptors. The cells are then separated from unbound ligand, usually by centrifugation and washing with buffer, and the amount of radioactivity bound to the cells is measured. Background binding by control cells is subtracted from the binding to the receptor-expressing cells, and the amount of bound ligand per cell is calculated and plotted (red curve) as a function of the ligand concentration. Note that even at relatively high ligand concentrations, the number of receptor-bound ligand molecules approaches, but does not equal, the number of cell-surface receptors. Nonetheless, by fitting of the data to Equation 15-4, one can determine that these cells express 1000 receptors for this ligand and that the K_d for binding of the ligand is 1 nM. In parallel experiments, the physiological response of the cell to increasing concentrations of ligand is also measured (blue curve). Typically, the plots of the extent of ligand binding to the receptor and of physiological response at different ligand concentrations differ. In the example shown here, 50 percent of the maximal physiological response is induced at a ligand concentration at which only 18 percent of the receptors are occupied. Likewise, 80 percent of the maximal response is induced when the

ligand concentration equals the K_d value, at which 50 percent of the receptors are occupied.

The concentration of unbound ligand ([L]) is shown on the horizontal axis, and the amount of receptor-ligand complex, usually expressed as the amount of receptor-bound ligand (RL), is shown on the vertical axis. In most cases, the absolute amount of ligand initially added into the extracellular fluid, often referred to as the *medium*, is vastly in excess of the total amount of cell-surface receptors. Thus even after some of the ligand has bound to the receptors, there is essentially no reduction in the concentration of unbound ligand in the medium. As a consequence, one can assume that the concentration of free (unbound) ligand [L] at the end of the incubation is equal to the concentration of ligand initially added to the fluid. In practice, as the concentration of ligand in the medium increases, the number of receptor-ligand complexes approaches but never actually reaches the total number of cell-surface receptors R_T (see Figure 15-8). The maximal amount of binding (B_{max}) that actually occurs is proportional to the total number of receptors present on the cells (R_T).

The maximal amount of binding (B_{max}) is one of two parameters that characterize a simple receptor-ligand binding assay, where each receptor protein has a single binding site and thus can bind only one ligand (see Equations 15-2 and 15-3 and Chapter 2). The other is the dissociation constant (K_d), which is an indication of how tightly the

ligand binds to the receptor. Note that the concentration of ligand that results in half of the receptors being bound at equilibrium ($[\mathrm{R}] = [\mathrm{R{*}L}] = 0.5 \times [\mathrm{R_T}]$) is equal to K_d (Equation 15-2). The tighter the binding of a ligand to its receptor, the lower the concentration of ligand needed to achieve binding to 50 percent of the receptors.

Computer curve-fitting programs are typically used to calculate the $\mathrm{R_T}$ and K_d values from data such as those in Figure 15-8. Using this approach, one can calculate from this binding curve that there are 1000 receptors per cell surface and that the K_d for binding this ligand is 1×10^{-9} M, or 1 nM.

Near-Maximal Cellular Response to a Signaling Molecule Usually Does Not Require Activation of All Receptors

The extent of the cellular response to a particular signaling molecule depends on the number of receptor-ligand complexes. In general, much less than 100 percent of its receptors need to be bound to the ligand to induce the maximal cellular response. One can observe this phenomenon by comparing the extent of the cellular response (blue curve in Figure 15-8) and the amount of ligand-bound receptors at different concentrations of the ligand. For example, an erythroid progenitor cell in the bone marrow has 1000 surface receptors for erythropoietin, the protein hormone that

induces these cells to proliferate and differentiate into red blood cells; the K_d for erythropoietin binding is 1 nM. But only 180 of these receptors (18 percent of the total) need to bind erythropoietin to induce 50 percent of the maximal cellular response (i.e., the division of the progenitor cell in this case). Thus the shape of the curve showing the dependence of a cellular response on ligand concentration differs from that of the curve for ligand binding (red curve in Figure 15-8). The ligand concentration needed for a cellular response that is 50 percent of the maximum is considerably lower than the K_d value.

Sensitivity of a Cell to External Signals Is Determined by the Number of Cell-Surface Receptors and Their Affinity for Ligand

Because the cellular response to a particular signaling molecule depends on the *number* of receptor-ligand complexes, the fewer receptors for a ligand present on the surface of a cell, the less *sensitive* the cell is to that ligand. As a consequence, a higher ligand concentration is necessary to induce the physiological response than would be the case if more receptors were present. Of course, if the cell has no receptors, there can be no response. In contrast, if the level of a receptor for a particular ligand is increased, the cell

will become more sensitive to the ligand. That is, fewer ligand molecules need be present to induce a response.

Epidermal growth factor (EGF), as its name implies, stimulates the proliferation of many types of epithelial cells (see Chapters 16 and 20), including those that line the ducts of the mammary gland. In about 25 percent of breast cancers, the tumor cells have elevated levels of one particular type of EGF receptor called *HER2*. The overproduction of HER2 makes the cells hypersensitive to ambient levels of EGF and related hormones. Normal levels of EGF are too low to stimulate cell proliferation, but in the presence of excess HER2, even normal levels of EGF stimulate growth of tumor cells. We will see in Chapters 16 and 25 that researchers have developed monoclonal antibodies that bind HER2 and thereby block signaling by EGF; these antibodies have proved useful in treatment of breast cancer patients whose tumors overexpress HER2.

The HER2–breast cancer connection vividly demonstrates the importance of regulating the number of receptors for a given signaling molecule. Such regulation can occur at the levels of transcription, translation, and post-translational processing. Alternatively, endocytosis can remove receptors from the cell surface, often leading to their degradation in lysosomes, reducing their numbers so much that the cellular response is effectively eliminated. As we discuss in later sections of this chapter, other mechanisms can reduce a receptor's affinity for ligand. This

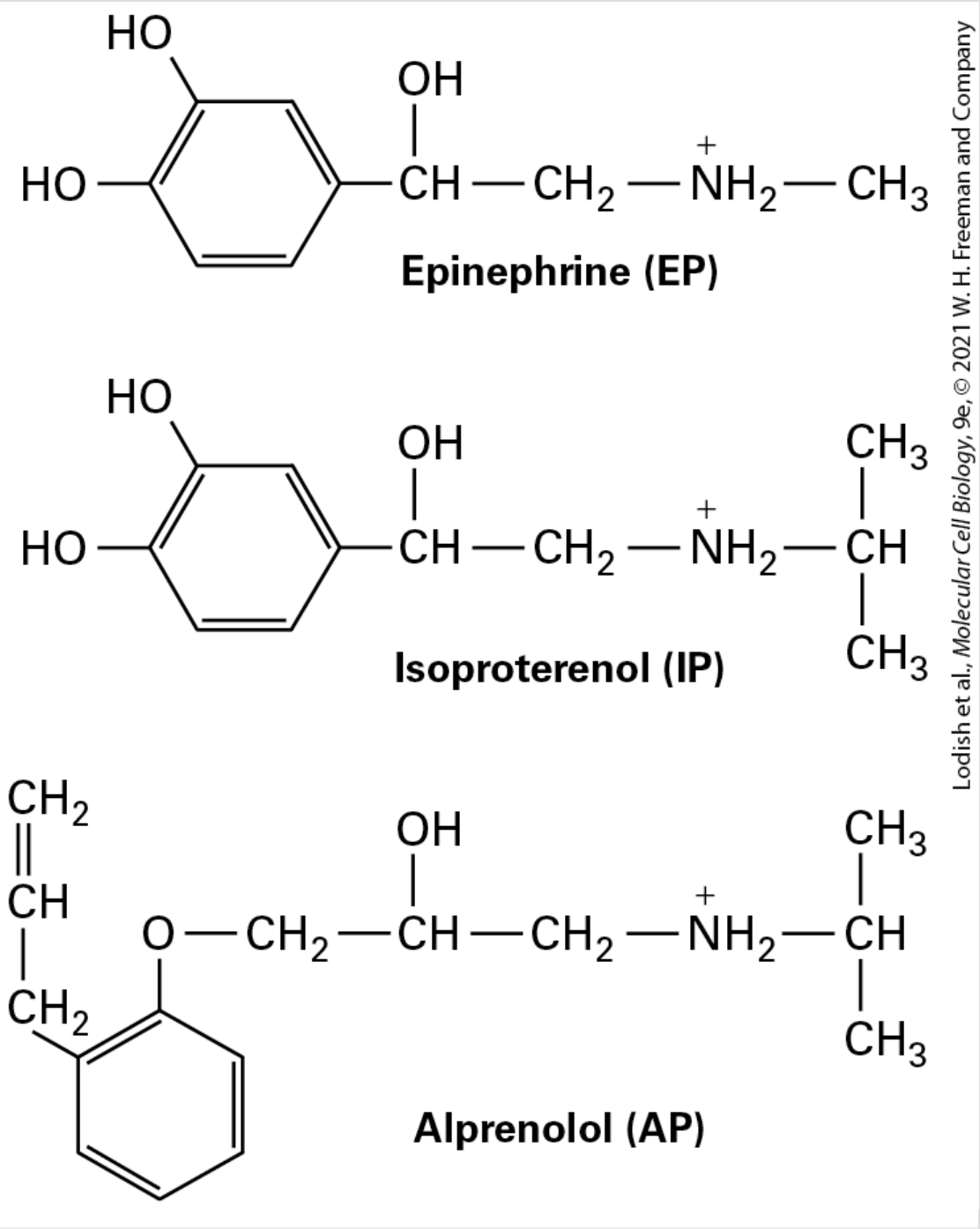
reduction of a cell's sensitivity to a particular ligand, called *desensitization*, can dramatically reduce the ability of cells to respond appropriately to external signals.

Chemical Analogs of Signaling Molecules Are Used to Study Receptors and Are Widely Used as Drugs

Synthetic analogs of natural signaling molecules, including hormones, are widely used both in research on cell-surface receptors and clinically as drugs. These analogs fall into two major classes. **Agonists** mimic the function of a natural hormone by binding to its receptor and inducing the allosteric $\mathrm{LR} \rightarrow \mathrm{LR}^*$ conformational change that activates the downstream signal transduction pathway and thus the normal cellular response to the hormone. Many synthetic agonists bind much more tightly to the receptor than does the natural hormone.

In contrast, **antagonists** bind to the receptor, usually at the normal ligand-binding site, but do not induce a conformational change. As a consequence, antagonists induce no cellular response. By occupying ligand-binding sites on a receptor, an antagonist can block binding of a natural hormone (or agonist) and thus reduce the hormone's usual physiological activity. In other words, antagonists inhibit receptor signaling stimulated by the natural ligand.

Consider, for instance, the drug *isoproterenol*, one of the first epinephrine agonists used to treat asthma. Isoproterenol is made by the chemical addition of two methyl groups to epinephrine (Figure 15-9). Isoproterenol is an agonist of the epinephrine receptor (called the β_2 *adrenergic receptor*, see Section 15.4) on bronchial smooth muscle cells. The drug binds about tenfold more tightly than does epinephrine; that is, it has a tenfold lower K_d. When activated, β_2 adrenergic receptors cause bronchial smooth muscle to relax, opening the air passages in the lungs. Because of this effect, isoproterenol and other agonists that have greater affinity and greater specificity for the β_2 adrenergic receptor are used in treating bronchial asthma, chronic bronchitis, and emphysema.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 15-9 Structures of the natural hormone epinephrine, the synthetic agonist isoproterenol, and the synthetic antagonist alprenolol. As discussed in the text, isoproterenol and alprenolol, both of which bind to receptors for epinephrine, are used as drugs to treat different conditions.

A different type of epinephrine-responsive receptor is present on cardiac muscle cells (called the *β_1-adrenergic receptor*). Its activation increases the heart's contraction rate. Antagonists of this receptor, such as alprenolol (see Figure 15-9) and related compounds, are referred to as *beta-blockers*; such antagonists are used to slow heart contractions in the treatment of cardiac arrhythmias and angina. ■

Receptors Can Be Purified by Affinity Chromatography Techniques

In order to study how receptors function, it is first necessary to purify them. For example, once an investigator has a sample of a purified protein, its molecular structure with and without a bound ligand can be determined in order to elucidate the conformational changes that occur upon ligand binding. Separating integral membrane cell-surface receptors from other cellular proteins is very challenging. A typical mammalian cell can have between 1000 to 50,000 copies of a single type of cell-surface receptor. This may seem like a large number, but consider that this same cell contains $\sim 10^{10}$ total protein molecules, with $\sim 10^6$ of them in the plasma membrane alone; these receptors constitute only 0.1 to 5 percent of plasma-membrane proteins. This low abundance complicates the isolation and purification of cell-surface receptors. Purification of these integral membrane proteins is also difficult because the membrane must first be solubilized with a non-ionic detergent (see

Figure 10-23) while maintaining the three-dimensional structure of the receptor and its ability to bind ligand. Only then can the receptors be separated from other cellular molecules.

Recombinant DNA techniques can often be used to generate cells that express large amounts of receptor proteins. Even then, special techniques are necessary to isolate the receptors from other membrane proteins. A type of *affinity chromatography* is often used in purifying cell-surface receptors that retain their ligand-binding ability when they are solubilized by detergents (see Figure 3-44c). Either an antibody that recognizes the receptor or a ligand for the receptor is chemically linked to the beads used to form a column. A crude, detergent-solubilized preparation of membrane proteins is then passed through the column. Only the receptor protein, together with other proteins tightly bound to it, will stick to the column; other proteins are washed away. Once the other proteins are removed, the receptors can be released ("eluted") from the column either by passage of an excess of the soluble ligand through the column (ligand affinity chromatography) or by changing chemical conditions such as pH to release the receptor from the antibody (antibody affinity chromatography). In some cases, a receptor can be purified as much as 100,000-fold in a single affinity-chromatographic step. After purification, membrane receptors can be studied either in their detergent-soluble form or after reconstitution into membrane bilayers in vitro (see Chapter 11).

Immunoprecipitation Assays and Affinity Techniques Can Be Used to Study the Activity of Protein Kinases

Following ligand binding, receptors activate one or more signal transduction proteins that, in turn, can affect the activity of multiple target effector proteins (see Figures 15-1 and 15-3). Understanding a signaling cascade requires the researcher to be able to quantitatively measure the activity of these signal transduction proteins. As kinases are found in many signaling cascades, in this section we describe several assays used for measuring their activity.

As noted above, any particular kinase can be in an inactive form or in a partially or completely active form, depending in part on which signal transduction pathways have been activated. The amount of activity exhibited by any given kinase will depend on the amount of kinase protein present and the fraction of that kinase protein that is in the active form. Thus investigators are often interested in determining the fraction of a particular kinase protein in a cell that is enzymatically active as well as the total enzymatic activity of that kinase.

Immunoprecipitation of Kinases

Immunoprecipitation assays, a type of antibody affinity chromatography (see Figure 3-44c), are frequently used to measure the activity of a particular kinase in a cell extract. As there are hundreds of kinases in eukaryotic cells, it is necessary first to purify the kinase of interest in order to measure its enzymatic activity without interference from other kinases. In one version of the immunoprecipitation method, an antibody specific for the desired kinase is first incubated with small beads coated with protein A; this causes the antibody to bind to the beads via its Fc segment (see Figure 4-33). The beads are then mixed with an extract of the whole cell or of an organelle, such as the nucleus, then recovered by centrifugation and washed extensively with a salt solution to remove weakly bound proteins that are unlikely to be binding specifically to the antibody. Thus only cell proteins that specifically bind to the antibody — the kinase itself and proteins tightly bound to the kinase — are present on the beads.

To measure its activity, the kinase isolated on the beads is exposed to its substrates. The beads are incubated in a buffered solution with a substrate protein and γ-[^{32}P] ATP, in which only the γ phosphate of the ATP is radiolabeled. The amount of [^{32}P] transferred covalently to the substrate protein is a measure of the amount of active kinase on the beads. That amount can be determined by separating the substrate protein and measuring the substrate-associated radioactivity. For example, the substrate can be isolated by polyacrylamide gel electrophoresis followed by autoradiography to

determine the amount of covalently attached [^{32}P] (see Figure 3-42), or it can be isolated by immunoprecipitation with an antibody specific for the substrate followed by measurement of the radioactivity in the immunoprecipitate. By comparing extracts from cells before and after ligand addition, for example, whether or not a particular kinase is activated in the signal transduction pathway(s) triggered by that ligand can be readily determined.

Western Blotting with a Monoclonal Antibody Specific for a Phosphorylated Amino Acid in a Protein

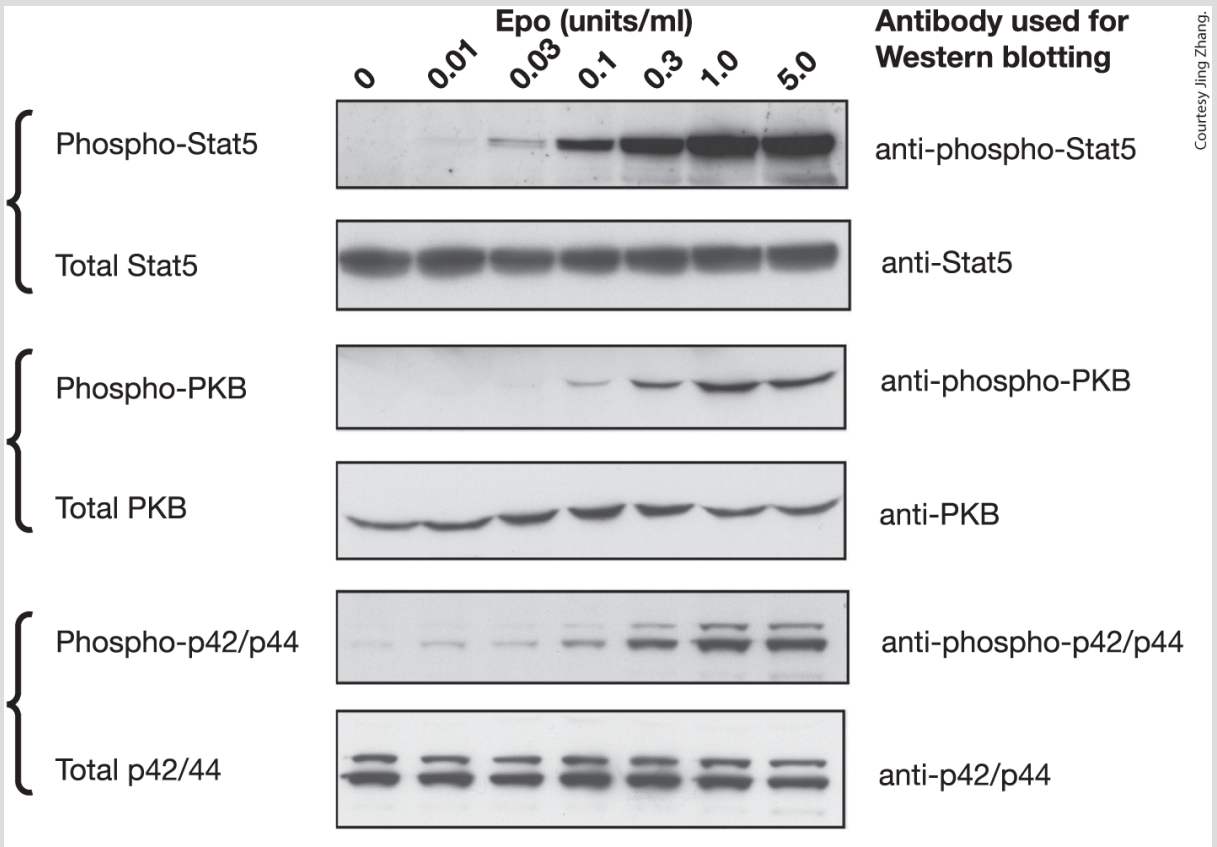
We noted above that many protein molecules can be phosphorylated by more than one kinase, each kinase phosphorylating the molecule on different serine, threonine, or tyrosine residues. To understand how such protein substrates are phosphorylated and how such multiple phosphorylations influence the protein's activity and alter cell function, it is sometimes necessary to measure the extent of phosphorylation of only one of several phosphorylation sites in a specific protein, usually doing so before and after ligand addition.

Antibodies play a crucial role in detecting phosphorylation events at specific sites on the protein. It is often possible to generate an antibody that can bind to the protein of interest only when the site

of interest on the protein — the hydroxyl group on the side chain of serine, threonine, or tyrosine — has been phosphorylated. To do so, a mouse is immunized (see Chapter 24) with a small peptide, approximately 15-amino-acids long, that has the sequence of the phosphorylation site of interest and that is phosphorylated. The phosphopeptide is called an *antigen* and it is usually chemically synthesized and then coupled to an adjuvant (see Chapter 24) to increase its immunogenicity, prior to immunizing the mouse. From the immunized mouse a set of monoclonal antibodies can be generated (see Figure 4-6). Then a particular monoclonal antibody is selected that reacts *only* with the phosphorylated but not the nonphosphorylated peptide; such an antibody generally will bind to the protein only when it is phosphorylated on the particular amino acid of interest. This specificity is possible because the antibody binds simultaneously to the phosphorylated amino acid and to side chains of adjacent amino acids.

As an example of the use of such antibodies, Figure 15-10 shows that three signal transduction proteins in erythroid progenitor cells become phosphorylated on specific amino acid residues within 10 minutes of stimulation by varying concentrations of the hormone erythropoietin (Epo); phosphorylation, which is the first step in triggering the differentiation of these cells into red blood cells, increases with Epo concentration. As is described in the next chapter, each of these signal transduction proteins activates

different downstream effectors, including different transcription factors.



EXPERIMENTAL FIGURE 15-10 Activation by the hormone erythropoietin (Epo) of three signal transduction proteins via their phosphorylation. Mouse erythrocyte progenitor cells were treated for 10 minutes with different concentrations of the hormone erythropoietin (Epo). Extracts of the cells were analyzed by Western blotting with three different antibodies specific for the phosphorylated forms of three signal transduction proteins and three other antibodies that recognize a nonphosphorylated segment of amino acids in each of the same proteins. These data show that with increasing concentrations of Epo, the three proteins become phosphorylated. Treatment with 1 unit Epo/ml is sufficient to maximally phosphorylate and thus activate all three pathways. $\mathrm{Stat5} = \mathrm{transcription\ factor}$ phosphorylated on tyrosine 694; $\mathrm{Stat5} = \mathrm{protein\ kinase\ B}$ phosphorylated on serine 473; $\mathrm{p42/p44} = \mathrm{p42/p44}$ MAP kinase phosphorylated on threonine 202 and tyrosine 204.

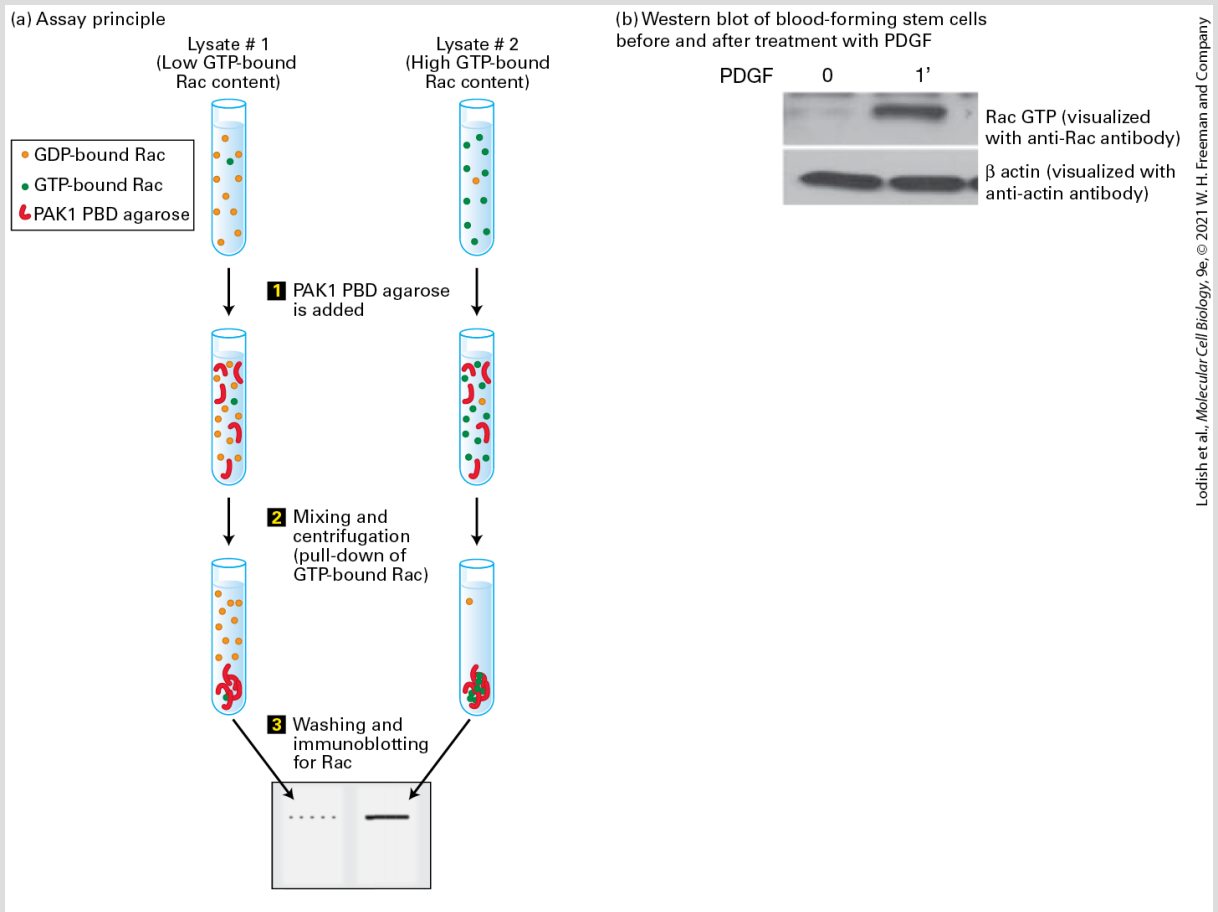
[Data from J. Zhang et al., 2003, *Blood* **102**:3938.]

GTP-Binding Signal Transduction Proteins Can Be Isolated and Their Activities Measured Using Pull-Down Assays

We noted that two large classes of GTPase switch proteins are used in signal transduction pathways: heterotrimeric G proteins that directly bind to and are activated by certain cell-surface receptors, and monomeric (often called *low-molecular-weight*) G proteins.

The principal assay for measuring activation of this class of proteins takes advantage of the fact that each switch protein has one or more targets to which it binds, but only when it is bound to GTP, not GDP; the target protein usually has a specific binding domain that binds to the switch segments of that GTP-bound protein. Pull-down assays used to quantify the activation of a specific GTP-binding protein are similar to immunoprecipitation assays, except that, rather than an antibody, the domain of the target protein that binds the GTP-bound GTP-binding protein is immobilized on small beads (Figure 15-11a). The beads are mixed with a cell extract; the beads bind any of the GTP-bound GTP-binding protein present in the extract; and then the beads are separated from the unbound components in the cell extract by centrifugation. The amount of the active, GTP-bound GTP-binding protein on the beads can be quantified using an antibody to the GTP-binding protein and Western blotting. Figure 15-11b shows

that in blood-forming stem cells, the fraction of the small GTPase Rac that has a bound GTP increases markedly after stimulation by the hormone platelet-derived growth factor (PDGF), indicating that Rac is a signal transduction protein activated by the PDGF receptor.



EXPERIMENTAL FIGURE 15-11 A pull-down assay shows that the small GTP-binding protein Rac is activated by platelet-derived growth factor (PDGF). Like other small GTPases, Rac regulates molecular events by cycling between an inactive GDP-bound form and an active GTP-bound form. In its active (GTP-bound) state, Rac binds specifically to the Rac binding (PBD) domain of p21-activated protein kinase (PAK1) to control downstream signaling cascades. (a) Assay principle: the Rac-binding PBD domain is generated by recombinant DNA techniques and attached to agarose beads, then mixed with cell extracts (step 1). The beads are specifically recovered by centrifugation (step 2), and the amount of GTP-bound Rac is quantified by Western blotting using an anti-Rac antibody (step 3).

(b) Western blot showing activation of Rac after treatment of blood-forming stem cells for 1 minute with the hormone platelet-derived growth factor (PDGF). A Western blot for actin serves as a control to show that the same amount of total protein is loaded on each lane of the gel.

[Part (b) Republished with permission from The American Society of Hematology, from G. Ghiaur et. al., 2006, "Inhibition of RhoA GTPase Activity Enhances Hematopoietic Stem and Progenitor Cell Proliferation and Engraftment," *Blood* **108**:2087–2094; permission conveyed through Copyright Clearance Center, Inc.]

Concentrations of Free Ca^{2+} in the Mitochondrial Matrix, ER, and Cytosol Can Be Measured with Targeted Fluorescent Proteins

The concentration of the second messenger Ca^{2+} changes in response to many different signal transduction systems, and such changes can influence the activities of many cellular proteins. Chapter 4 describes how the fluorescent small-molecule dye *fura-2* can be used to measure the concentration of free Ca^{2+} (Ca^{2+} not bound to proteins) in the cytosol of live cells. Fura-2 fluorescence at a certain wavelength increases when Ca^{2+} is bound. The technique depends on the fact that a chemically modified version of fura-2, in which the carboxylic acid residues are in an ester linkage, spontaneously diffuses into the cytosol from extracellular fluids. Enzymes in the cell cytosol cleave off these ester groups, forming

free fura-2 that cannot diffuse out of the cytosol and whose fluorescence reports the concentration of free Ca^{2+} in the cytosol.

Several proteins also emit light or their emitted light changes color when they bind Ca^{2+}. An advantage of these proteins is that they can be used experimentally to determine the free Ca^{2+} concentration not just in the cytosol but also within subcellular compartments. Cells can be transfected with cDNA expression vectors that encode these proteins, which have been altered to include a signal sequence that targets the protein to a particular subcellular compartment (Chapter 13).

One such protein is aequorin, a calcium-activated bioluminescent protein isolated from the hydrozoan *Aequorea victoria*. As expressed in cells by recombinant DNA technologies, aequorin consists of a protein subunit fused with a signal sequence that targets it to a specific organelle, such as the ER lumen (see Figure 13-6) or the mitochondrial intermembrane space or matrix (see Figure 13-26). Aequorin contains three EF hands (see Chapter 3) that function as binding sites for Ca^{2+}. When the small-molecule prosthetic group coelenterazine is added to the culture medium, it diffuses into the cell and binds to aequorin, and the cell emits light at a specific wavelength proportional to the concentration of free Ca^{2+} in that subcellular space.

KEY CONCEPTS OF SECTION 15.2

Studying Cell-Surface Receptors and Signal Transduction Proteins

- The concentration of ligand at which half the ligand's receptors are occupied, the dissociation constant (K_{d}), can be determined experimentally and is a measure of the affinity of the receptor for the ligand.
- The near-maximal response of a cell to a particular ligand generally occurs at ligand concentrations at which less than 100 percent of its receptors are bound to the ligand.
- Affinity chromatography techniques can be used to purify receptors even when they are present in low abundance.
- Immunoprecipitation using antibodies specific for protein kinases can be used to measure the activity of these kinases.
- Western blotting assays using antibodies specific for phosphorylated peptides can measure phosphorylation of a specific amino acid on any desired protein within a cell.
- Pull-down assays using the protein-binding domain of a target protein can be used to quantify activation of a GTP-binding protein within a cell.
- The concentration of free Ca^{2+} in a subcellular space can be measured using fura-2 or genetically engineered aequorin proteins.

15.3 Structure and Mechanism of G Protein–Coupled Receptors

In this section, we introduce the most numerous class of receptors: the G protein–coupled receptors (GPCRs). The human genome encodes about 800 functional GPCRs, which constitute about 4 percent of the identified human proteins. These detect and respond to many different types of signals, including odorants, tastants, neurotransmitters, hormones involved in glycogen and fat metabolism, growth factors, and even physical stimuli (photons of light). Many GPCRs are found primarily in cells of the central nervous system and are used in neuronal signaling, as we will learn in Chapter 23. GPCRs are of immense medical importance, as approximately 35 percent of all drugs used in humans are agonists or antagonists of GPCRs, and at last count 134 GPCRs are targets of drugs approved in the United States or the European Union. Table 15-1 describes just a few of these drugs; note the wide variety of receptors targeted by these drugs and the diverse diseases that they treat.

TABLE 15-1 • Human G Protein–Coupled Receptors of Pharmaceutical Importance

Receptor	Natural Ligand	Location	Physiological Function	Drug	Medical Use
Histamine H2 receptor	Histamine	Acid-secreting cells of the stomach	Stimulates acid secretion	Cimetidine (Tagamet) Ranitidine (Zantac) (antagonists)	Prevent acid stomach; treat peptic ulcers
Histamine H1 receptor	Histamine	Smooth muscles, vascular	Increases vascular permeability	Fexofenadine (Allegra) Loratadine	Reduce symptoms of allergy

		endothelial cells	and causes symptoms of allergy	(Claritin) (antagonists)	
Serotonin $5HT_{2A}$	Serotonin	Central nervous system	Synaptic transmission between neurons	Clozapine, risperidone (antagonists)	Treat schizophrenia
Serotonin $5HT_{1a}$	Serotonin	Central nervous system	Synaptic transmission between neurons	Buspirone (BuSpar) (agonist)	Treat depression, general anxiety disorder
Angiotensin AT_1	Angiotensin II	Vascular smooth muscle cells	Constrict blood vessels and increase blood pressure	Losartan (Cozarr) (antagonist)	Reduce hypertension
β_2-adrenergic receptor	Epinephrine	Smooth muscle cells lining the airway	Facilitate respiration	Salmeterol (Severent) (agonist)	Treatment of asthma, chronic obstructive pulmonary disease
$CysLT_1$	Leukotrienes	Lungs, bronchial tubes, mast cells	Contracts smooth muscles	Montelukast (Singulair) (antagonist)	Treatment of asthma, seasonal allergies

SOURCE: A. Wise et al., 2002, *Drug Discovery Today* **7**:235–246.

All GPCR signal transduction pathways share the following common elements: (1) a membrane-imbedded receptor that contains seven membrane-spanning α helices; (2) a heterotrimeric G protein, which functions as a receptor-activated switch by cycling between an active GTP-bound and an inactive GDP-bound state; (3) a membrane-bound effector

protein; and (4) proteins that participate in amplification and desensitization of the signaling pathway. Protein kinases and second messengers such as cAMP and Ca^{2+} often participate in GPCR-mediated signal transduction pathways. These pathways usually have short-term effects in the cell by quickly modifying the activities of existing proteins, including enzymes and ion channels. Thus these pathways allow cells to respond rapidly to a variety of signals, whether they be physical stimuli such as light or hormonal stimuli such as epinephrine.

In this section, we discuss the basic structures and mechanisms of GPCRs and their associated heterotrimeric G proteins. The following sections describe several well-understood and important examples of signal transduction pathways activated by these receptors.

All G Protein–Coupled Receptors Share the Same Basic Structure

All G protein–coupled receptors have the same orientation in the membrane and contain seven transmembrane α-helical regions (H1–H7), four extracellular segments, and four cytosolic segments (Figure 15-12). Invariably the N-terminus is on the exoplasmic face and the C-terminus is on the cytosolic face of the plasma membrane. G protein–coupled receptors are stably anchored in the hydrophobic core of the plasma membrane by the many hydrophobic amino acids on the outer surfaces of the seven membrane-spanning α-helices.

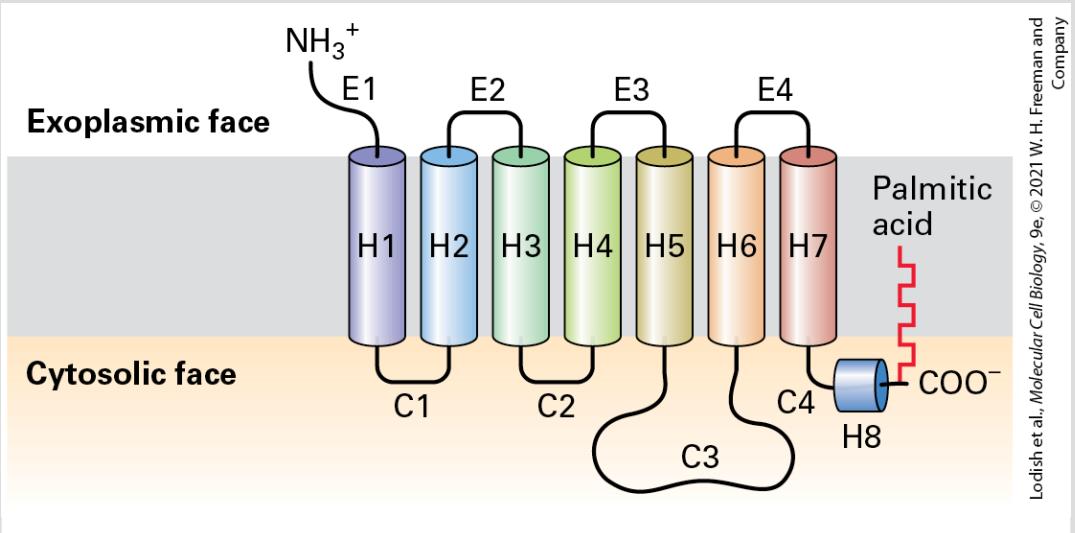


FIGURE 15-12 General structure of G protein–coupled receptors. All receptors of this type have the same orientation in the membrane and contain seven transmembrane α-helical regions (H1–H7), four extracellular segments (E1–E4), and four cytosolic segments (C1–C4). In many GPCRs a short α helix H8 lies parallel to the membrane surface, and a covalently bound palmitic acid group anchors the C-terminus to the membrane surface.

The human GPCRs are divided into several families. Members of each family are especially similar to one another in amino acid sequence and structure. The families differ from each other in the details of how ligands bind the receptors.

Figure 15-13 shows the variation in ligand-binding sites between three of the five different families of GPCRs. In family A, the largest family in mammals, segments of several membrane-embedded α helices and extracellular loops form the ligand-binding site. That site is exposed to the exoplasmic surface, but can extend down toward the middle of the membrane bilayer. Examples of family A receptors are the **β-adrenergic receptors**, which bind hormones such as epinephrine and norepinephrine (Figure 15-14a). These are some of the best studied GPCRs and are frequently used as archetypical examples of GPCRs (Figure 15-14b). As an example, in binding of epinephrine (adrenaline) to the β-adrenergic receptor, side chains of 15 amino acids

located in four transmembrane α helices and the extracellular loop E2 make noncovalent bonds with the bound ligand (Figure 15-14c); the two hydroxyl groups on epinephrine become hydrogen bonded to the hydroxyl groups on the side chains of two serine (S) residues (designated S203 and S207 in transmembrane helix #5) as well as to the side chain of an asparagine residue (N293) in transmembrane helix #6.

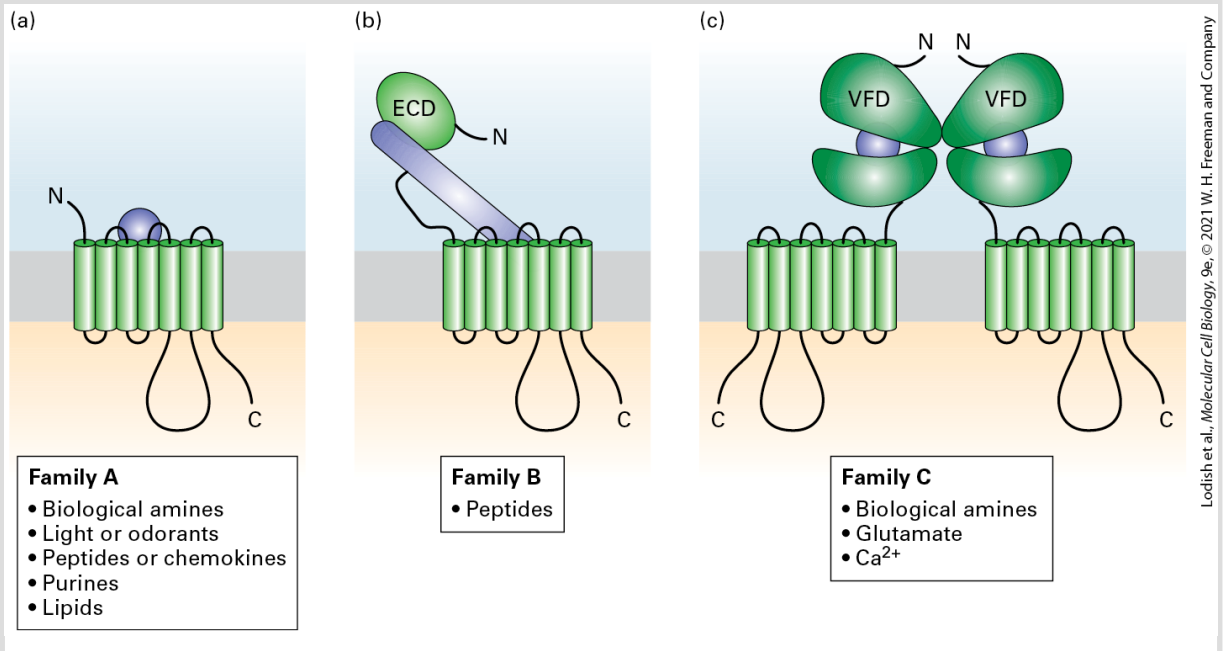
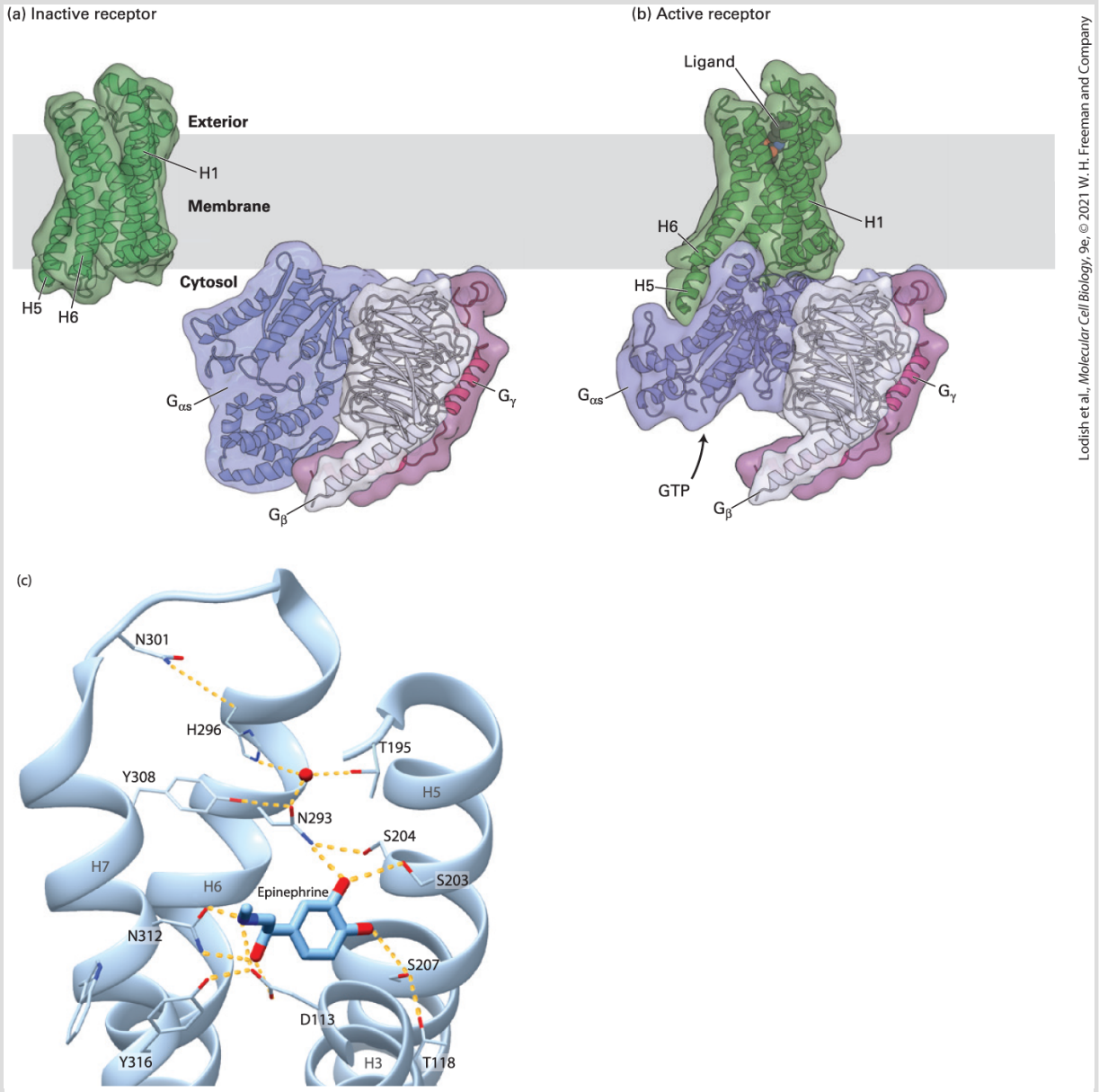


FIGURE 15-13 Three principal classes of G protein–coupled receptors that bind their ligand in different ways. See text for details.

[Data from J. Tesmer, 2016, *Nat. Rev. Mol. Cell Biol.* **17**:439.]



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 15-14 Structure of the β_2-adrenergic receptor in the inactive state and in the active state with a bound ligand and with its associated heterotrimeric G protein, G_s. (a) The three-dimensional structure of the β_2-adrenergic receptor bound to an antagonist (not shown) and thus in the inactive state. Placed next to it are the three-dimensional structures of the subunits of the heterotrimeric G protein G_s : $G_{\alpha s}$ (dark purple), G_β (light purple), and G_γ (pink), showing the inability of the resting receptor to bind to and activate $G_{\alpha s}$. (b) The overall structure of the active receptor complex shows the adrenergic receptor bound to an agonist (blue and red spheres) and engaged in extensive interactions with a segment of $G_{\alpha s}$. (c) Binding of epinephrine to the side chains of multiple amino acids in the interior of the β_2-adrenergic receptor.

[Data for parts (a) and (b) from S. Rasmussen et al., 2011, *Nature* **476**:549, PDB ID 3sn6; and V. Cherezov et al., 2007, *Science* **318**:1258, PDB ID 2rh1. Part (c) from A. M. Ring et al., 2013, *Nature* **502**:575–579 https://doi.org/10.1038/nature12572.]

The glucagon receptor is a member of family B of the GPCRs (Figure 15-13b). Its ligand *glucagon* is a 29-amino-acid peptide hormone secreted by the α cells of the pancreatic islets; as we will see later in this chapter, it acts on the liver to trigger glycogen breakdown and secretion of glucose into the bloodstream. While the glucagon receptor has the seven transmembrane α helices characteristic of GPCRs, it also has a large exoplasmic domain that is connected to the first transmembrane α helix. This exoplasmic domain binds tightly to the C-terminus of glucagon, positioning glucagon's N-terminus to bind to the pocket formed, as in β-adrenergic and other GPCRs, by residues from several transmembrane helices (Figure 15-15).

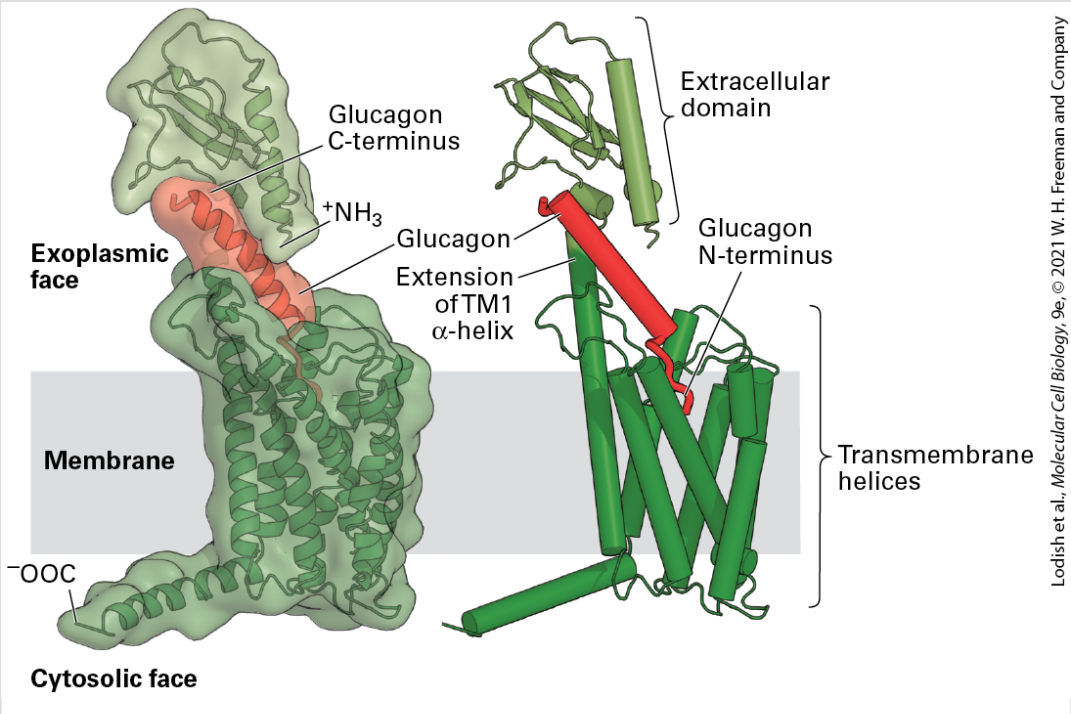


FIGURE 15-15 Structure of the glucagon receptor with a bound glycogen molecule. The seven transmembrane α helices of the glucagon receptor and the exoplasmic extension of transmembrane helix 1 are colored dark green and the N-terminal exoplasmic domain light

green. The C-terminus of the 29-amino-acid peptide glucagon (red) is bound to the exoplasmic domain of the receptor, and the glucagon N-terminus is thought to insert into a binding pocket that is in the center of the seven transmembrane α helices.

[Data from P. Siu et al., 2014, *Nature* **499**:444, custom PDB.]

The glutamate receptor is an excitatory neurotransmitter receptor in family C of GPCRs (Figure 15-13c). The long extracellular segment forms a sort of venus flytrap domain that clamps around and binds the glutamate ligand. Many members of this class of GPCRs dimerize through their venus flytrap domains.

Like most receptors, GPCRs are allosteric proteins (Chapter 3). Binding of the ligand at the extracellular surface causes a conformational change in the receptor ($R \rightarrow R^*$) that enables the cytosol-facing segment of the receptor to bind to and activate the heteromeric G protein, the next intracellular component of the GPCR signal transduction pathway — hence the name *G-protein coupled receptor*.

Ligand-Activated G Protein–Coupled Receptors Catalyze Exchange of GTP for GDP on the α Subunit of a Heterotrimeric G Protein

The first signal transduction step after the activation of GPCRs by their ligands is the activation of heterotrimeric G proteins (Figure 15-16). These G proteins contain three subunits, designated α, β, and γ. Guanine nucleotides (GTP or GDP) bind to the G_α subunit. The G_α and G_γ subunits have covalently attached fatty acids that insert into the inner leaflet of the plasma

membrane and allow the G protein subunits to be in close proximity to the GPCR. The β and γ subunits are always bound tightly together and are usually referred to as the $G_{\beta\gamma}$ *subunit*. As we see shortly, the G_α subunit dissociates and reassociates with the $G_{\beta\gamma}$ subunit, depending on its activation state.

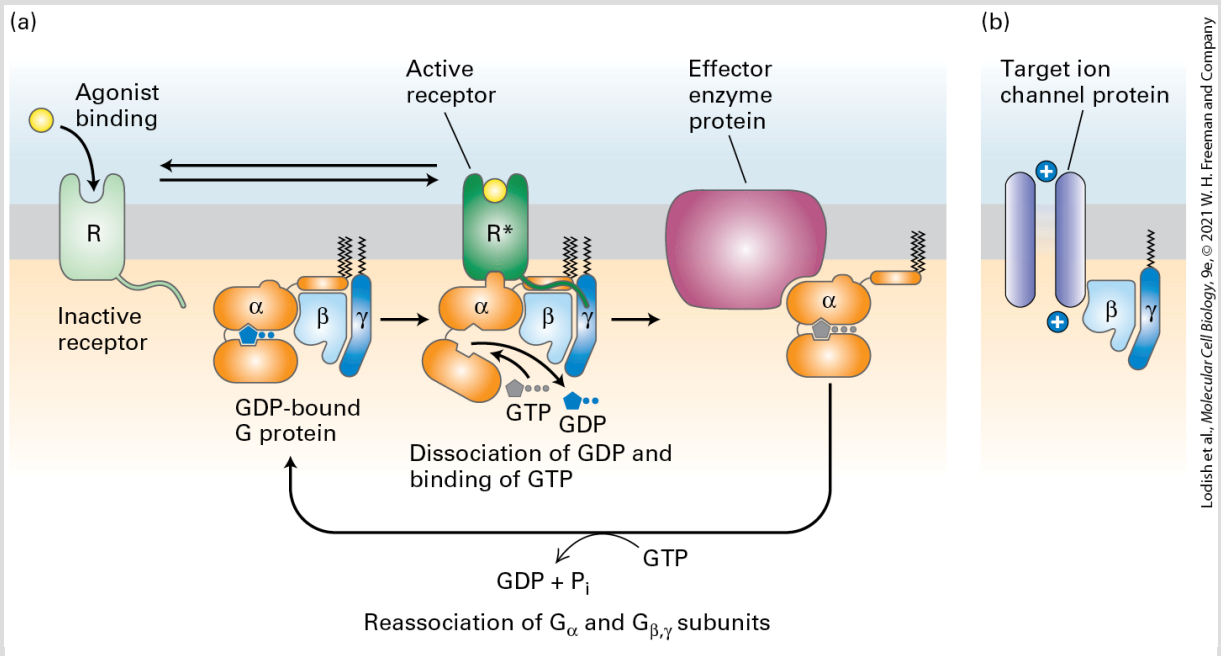


FIGURE 15-16 General mechanism of the activation of effector proteins associated with G protein–coupled receptors. Light green denotes the inactive and dark green the active conformations of the GPCR. The G_α and $G_{\beta\gamma}$ subunits of a heterotrimeric G protein are tethered to the membrane by covalently attached lipid molecules (wiggly black lines). Following ligand binding, the receptor undergoes a conformational change enabling it to bind to the G_α subunit of the G protein, triggering dissociation of the GDP and binding of GTP, concomitant with dissociation of the $G_{\beta\gamma}$ protein subunits. In some cases (a) the free $G_\alpha \cdot GTP$ subunit diffuses in the plasma membrane and binds to and activates an effector protein. In others (b), the free $G_{\beta\gamma}$ subunit binds to and activates an ion channel or other target protein. Hydrolysis of GTP terminates signaling and leads to reassembly of the heterotrimeric G protein, returning the system to the resting state. Binding of another ligand molecule causes repetition of the cycle.

[Data from W. Weis and B. Kobilka, 2018, *Annu. Rev. Biochem.* **87**:897.]

In the resting state, when no ligand is bound to the receptor, the G_α subunit has a GDP bound to its nucleotide binding site and the G_α GDP is complexed with $G_{\beta\gamma}$, forming a heterotrimer ($^{GDP}G_\alpha G_{\beta\gamma}$). Binding of a ligand to a G protein–coupled receptor alters the conformation of its transmembrane helices and enables the receptor to bind to the G_α subunit of the intact heterotrimeric G protein containing GDP (Figures 15-14 and 15-16). The binding of epinephrine to the β-adrenergic receptor provides an example of how the change in receptor conformation allows binding to the G_α subunit. When unbound by ligand and in the inactive conformation (R), the receptor (see Figure 15-14a) has no surface that is complementary to the G protein and thus cannot bind to it. The binding of the agonist induces major conformational changes ($L + R \rightarrow LR^*$) in this receptor: there are substantial movements of transmembrane helices 5 and 6 and changes in the structure of the C3 cytosol-facing loop that connects these helices. Together, these changes create a surface that can now bind to a segment of the G_α subunit of $^{GDP}G_\alpha G_{\beta\gamma}$ (see Figure 15-14b).

Binding of the GPCR to the G_α subunit of $^{GDP}G_\alpha G_{\beta\gamma}$ induces a change in the conformation of the GDP-binding site on G_α, allowing a lobe of the protein to move and thus release the GDP. Next because of its higher concentration in the cytosol, GTP rapidly binds to the empty guanine nucleotide site in the G_α subunit. Thus the activated ligand-bound receptor functions as a guanine nucleotide exchange factor for $^{GDP}G_\alpha$.

GTP binding to G_α induces a change in the conformation of its switch segments (see Figure 15-7), leading the $^{GTP}G_\alpha$ subunit to dissociate from the $G_{\beta\gamma}$ subunit and from the GPCR. All three elements can now diffuse independently in the plane of the cytosolic face of the plasma membrane. The freed, ligand-bound, activated GPCR (LR*) is then able to activate

another ${}^{GDP}G_{\alpha}G_{\beta\gamma}$, and thus amplify the signal from the receptor-ligand complex.

In most cases the released G_{α}·GTP then mediates the next step of the signal transduction pathway. While remaining anchored in the plasma membrane, G_{α}·GTP binds to and activates a downstream effector protein (Figure 15-16a). In some cases, G_{α}·GTP inhibits, rather than activates, the effector. As we see below, there are various types of G_{α} protein subunits that interact with different receptors and effectors. In some cases (Figure 15-16b), depending on the type of cell and G protein involved, it is the $G_{\beta\gamma}$ subunit, freed from the G_{α} subunit, that transduces a signal by interacting with an effector protein. For example, *muscarinic acetylcholine receptors* are a type of GPCR found in cardiac muscle; when activated by binding acetylcholine, these receptors slow the rate of heart muscle contraction. Following receptor activation, the released $G_{\beta\gamma}$ subunit binds to and opens a K^+ channel (the effector protein) in the plasma membrane (see Figure 15-16b). The subsequent efflux of K^+ ions out of the cytosol causes an increase in the magnitude of the usual inside-negative potential across the plasma membrane that lasts for several seconds, a change that reduces muscle contraction.

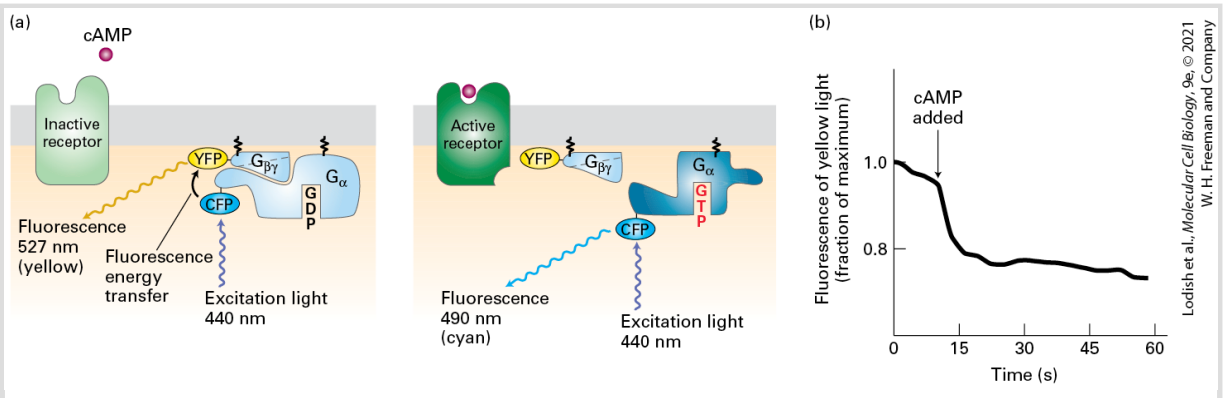
In all cases, the active ${}^{GTP}G_{\alpha}$ state is relatively short-lived because the bound GTP is hydrolyzed to GDP in seconds to minutes. The hydrolysis of GTP is catalyzed by the intrinsic GTPase activity of the G_{α} subunit itself. The G_{α} subunit thus switches from the active ${}^{GTP}G_{\alpha}$ to the inactive ${}^{GDP}G_{\alpha}$ conformation that can no longer interact with downstream effector proteins but can quickly reassociate with $G_{\beta\gamma}$ forming ${}^{GDP}G_{\alpha}G_{\beta\gamma}$. The cycle is complete and ${}^{GDP}G_{\alpha}G_{\beta\gamma}$ can interact with an activated receptor (LR*) and start the process all over again.

The rate of GTP hydrolysis by $^{GTP}G_{\alpha}$ is sometimes enhanced by its binding to an effector that also functions as a GTPase-activating protein (GAP). This feedback mechanism significantly reduces the duration of effector activation and prevents the cell from overreacting. In many cases, another type of GAP protein, called a *regulator of G protein signaling (RGS)*, can accelerate GTP hydrolysis by $^{GTP}G_{\alpha}$, further reducing the time during which the $^{GTP}G_{\alpha}$ can activate its downstream effector.

Thus built-in feedback mechanisms ensure that the effector protein remains activated for only a few seconds or minutes following receptor activation. Prolonged signaling is possible only if receptors are kept continually active by ligand binding, so that these receptors can continue to activate the corresponding G protein.

Early evidence supporting the model that hydrolysis of the GTP in $^{GTP}G_{\alpha}$ (see Figure 15-16) is important for GPCR signal transduction pathways came from studies using compounds that are structurally similar to GTP, called *GTP analogues*. These compounds can bind as well as GTP to the GTP binding sites in G_{α} subunits but cannot be hydrolyzed by the intrinsic GTPase. In some of these compounds, the P–O–P phosphodiester linkage connecting the β and γ phosphates of GTP is replaced by a nonhydrolyzable $P–CH_2–P$ or P–NH–P linkage. Such a GTP analog can be added to a plasma-membrane preparation containing a particular receptor in the presence of an agonist. The nonhydrolyzable GTP analog is exchanged for GDP on G_{α}, and thus both the G_{α} protein and its associated effector protein remain activated for much longer than occurs when GTP is bound. Many investigators use a nonhydrolyzable analogue called GTP-γS for such studies.

The dissociation of heterotrimeric G proteins after binding to a GPCR can be detected in live cells. These studies have exploited the phenomenon of Förster resonance energy transfer (FRET), which changes the wavelength of emitted fluorescence when two fluorescent molecules interact (see Figure 4-23). Figure 15-17 shows how this experimental approach has demonstrated the dissociation of the $G_\alpha \cdot G_{\beta\gamma}$ complex within a few seconds after ligand addition, providing evidence for the rapid dissociation of the heterotrimeric G protein. This general experimental approach can be used to follow the formation and dissociation of other protein-protein complexes in live cells.



EXPERIMENTAL FIGURE 15-17 Activation of a G protein occurs within seconds of ligand binding to its cell-surface G protein–coupled receptor. In the amoeba *Dictyostelium discoideum*, cAMP acts as an extracellular signaling molecule that binds to and signals via a G protein–coupled receptor; cAMP is not a second messenger. Amoeba cells were engineered to express two genes encoding two fusion proteins: a G_α fused to cyan fluorescent protein (CFP) and a G_β fused to yellow fluorescent protein (YFP). CFP normally fluoresces 490-nm light; YFP, 527-nm light. (a) When CFP and YFP are close to each other, as in the resting ${}^{GDP}G_\alpha \cdot G_{\beta\gamma}$ complex, resonance energy transfer can occur between them (*left*). As a result, irradiation of resting cells with 440-nm light (which directly excites CFP but not YFP) causes emission of 527-nm (yellow) light, characteristic of YFP, because of resonance energy transfer from CFP to YFP. However, if ligand binding leads to dissociation of the G_α and $G_{\beta\gamma}$ subunits, then resonance energy transfer cannot occur. In this case, irradiation of cells at 440 nm causes emission of 490-nm (cyan) light, characteristic of CFP (*right*). (b) Plot of the emission of yellow light (527 nm) from a single transfected amoeba cell before and after addition of extracellular cAMP (arrow). The drop in yellow fluorescence, which results from the dissociation of the G_α-CFP fusion protein from the G_β-YFP fusion protein, occurs within seconds of cAMP addition.

[Data from C. Janetopoulos et al., 2001, *Science* **291**:2408.]

Some bacterial toxins contain a subunit that penetrates the plasma membrane of cells, such as those that line the small intestine, and, once in the cytosol, catalyzes a chemical modification of $G_{\alpha s}$ proteins that prevents hydrolysis of bound GTP to GDP. An important example is the toxin produced by the bacterium *Vibrio cholerae*, which causes cholera. As a result, $G_{\alpha s}$ remains in the active state, continuously activating its effector protein adenylyl cyclase (described in Section 15.4) even though no hormone is present to activate the receptor. The resulting excessive rise in the intracellular second messenger cAMP leads to the loss of electrolytes and water from these intestinal cells into the intestinal lumen, producing the watery diarrhea characteristic of infection by these bacteria. ■

Different G Proteins Are Activated by Different GPCRs and in Turn Regulate Different Effector Proteins

All effector proteins in GPCR signal transduction pathways are either membrane-bound ion channels or membrane-bound enzymes that catalyze the formation of one or more of the second messengers shown in Figure 15-9. The variations on the theme of GPCR signaling that we examine in Sections 15.4, 15.5, and 15.6 and in later chapters arise because eukaryotic genomes encode multiple G proteins with distinct activities. Humans have 21 different G_α subunits: 6 G_β subunits and 12 G_γ subunits. The 21 G_α subunits are encoded by 16 genes, several of which undergo alternative splicing. So far as is known, the different $G_{\beta\gamma}$ subunits are interchangeable in their functions, while the different G_α subunits provide specificity to the various G proteins. Thus we can refer to the entire three-subunit G protein

by the name of its α subunit. The heterotrimeric G proteins ($G_\alpha G_{\beta\gamma}$) are usually assigned to one of four classes based on the sequences of their α subunits. Table 15-2 summarizes the functions of the major classes of G proteins with different G_α subunits and the effector proteins they activate.

TABLE 15-2 • Major Classes of Mammalian Heterotrimeric G Proteins and Their Effectors[i]

G_α Class	Associated Effector	Second Messenger	Receptor Examples
$G_{\alpha s}$	Adenylyl cyclase	cAMP (increased)	β-Adrenergic (epinephrine) receptor; receptors for glucagon, serotonin, vasopressin
$G_{\alpha i}$	Adenylyl cyclase K^+ channel ($G_{\beta\gamma}$ activates effector)	cAMP (decreased) Change in membrane potential	α_2-Adrenergic receptor Muscarinic acetylcholine receptor
$G_{\alpha olf}$	Adenylyl cyclase	cAMP (increased)	Odorant receptors in nose
$G_{\alpha q}$	Phospholipase C	IP_3, DAG (increased)	α_1-Adrenergic receptor
$G_{\alpha o}$	Phospholipase C	IP_3, DAG (increased)	Acetylcholine receptor in endothelial cells
$G_{\alpha t}$	cGMP phosphodiesterase	cGMP (decreased)	Rhodopsin (light receptor) in rod cells

[i]A given G_α subclass may be associated with more than one effector protein. To date, only one major $G_{\alpha s}$ has been identified, but multiple $G_{\alpha q}$ and $G_{\alpha i}$ proteins have been described. Effector proteins commonly are regulated by G_α but in some cases by $G_{\beta\gamma}$ or the combined action of G_α and $G_{\beta\gamma}$. $IP_3 = \text{inositol 1,4,5-trisphosphate}$; $DAG = \text{1,2-diacylglycerol}$.
SOURCES: L. Birnbaumer, 1992, *Cell* **71**:1069; Z. Farfel et al., 1999, *New Engl. J. Med*. **340**:1012; and K. Pierce et al., 2002, *Nat. Rev. Mol. Cell Biol*. **3**:639.

Analysis of GPCRs Has Identified Important Human Hormones

The GPCRs represent the largest protein family in the human genome. Approximately half of the ~800 genes encoding these proteins are thought to encode sensory receptors; of these, the majority are in the olfactory system and bind odorants (see Chapter 23).

The first GPCRs to be discovered were found by using their well-known hormone ligands as bait to fish for the corresponding receptor. Subsequently, many GPCRs were initially identified by searching genome sequences for genes whose protein coding sequences had high similarity to previously identified GPCRs (e.g., the β-adrenergic receptors and rhodopsin) and whose predicted proteins would have seven transmembrane α helices. The natural ligands for these putative receptors were not known, and so they were called *orphan receptors* — that is, putative GPCRs without known ligands. Over time, the natural ligands of some — but not all — of these have been identified. Many of the remaining orphan receptors are likely to bind signaling molecules that have not yet been identified, including novel peptide hormones.

One approach that has proved fruitful in identifying the ligands of orphan GPCRs involves expressing the gene encoding the receptor in transfected cells and using the cells as a reporter system to detect substances in tissue extracts that activate the receptor and its downstream signal transduction pathway. This approach led to the discovery of two novel peptides, termed *orexin-A* and *orexin-B* (from the Greek *orexis*, meaning "appetite"), that were identified as the ligands for two orphan GPCRs that are members of the B family of GPCRs that includes the glucagon receptor (see Figure 15-13b).

Further research showed that the *orexin* gene is expressed only in the hypothalamus, the part of the brain that regulates feeding. Injection of orexin into the brain ventricles of animals caused them to eat more, and expression of the *orexin* gene increased markedly during fasting. Both of these findings are consistent with orexin having a role in increasing appetite. Strikingly, mice deficient in orexins suffer from narcolepsy, a disorder characterized in humans by excessive daytime sleepiness (in mice, which are nocturnal, nighttime sleepiness). Moreover, recent reports suggest that the orexin system is dysfunctional in a majority of human narcolepsy patients: orexin peptides cannot be detected in their cerebrospinal fluid (although there is no evidence of mutation in their *orexin* genes). These findings firmly link orexin neuropeptides and their receptors to both feeding behavior and sleep in humans and other animals. They also illustrate the importance of GPCRs in human physiology.

KEY CONCEPTS OF SECTION 15.3

Structure and Mechanism of G Protein–Coupled Receptors

- G protein–coupled receptors (GPCRs) are a large and diverse group of proteins with a common structure of seven membrane-spanning α helices and an internal ligand-binding pocket that is specific for particular ligands.
- Different families of GPCRs bind their ligands in different manners, but binding always alters receptor conformation so that it is able to bind to and activate a G protein.
- GPCRs are coupled to heterotrimeric G proteins, which contain three subunits designated α, β, and γ. The G_α subunit is a GTPase switch protein that alternates between an active (“on”) state with bound GTP and inactive (“off”) state with bound GDP. The “on” form dissociates from the β and γ subunits and activates a membrane-bound effector. The β and γ subunits remain bound together and in some cases activate an effector.
- Ligand binding causes a conformational change in certain membrane-spanning helices and intracellular loops of the GPCR, allowing it to bind to its coupled G_α subunit and function as a guanine nucleotide exchange factor (GEF), catalyzing dissociation of GDP and allowing GTP to bind. The resulting change in the conformation of the switch region in G_α causes it to dissociate from the $G_{\beta\gamma}$ subunit and the receptor and interact with an effector protein.

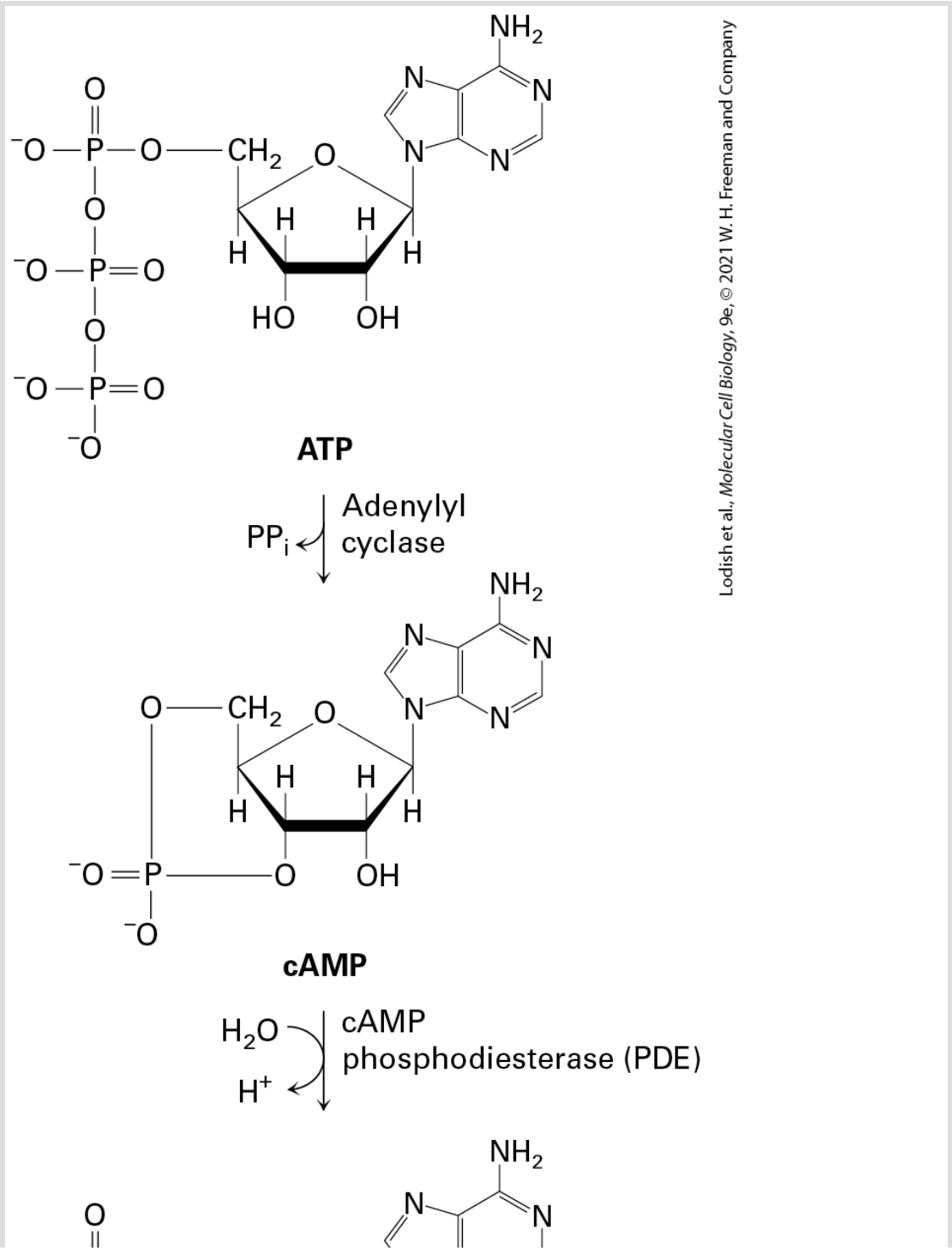
- Signaling is terminated when the GTP bound to the G_α subunit is hydrolyzed to GDP, triggering reassociation of G_α with $G_{\beta\gamma}$.
- FRET experiments demonstrate receptor-mediated dissociation of coupled G_α and $G_{\beta\gamma}$ subunits in live cells.
- The effector proteins activated (or inactivated) by heterotrimeric G proteins are either enzymes that form second messengers (e.g., adenylyl cyclase, phospholipase C) or ion channels. In each case, it is the G_α subunit that determines the function of the G protein and affords its specificity.
- GPCRs can have a range of cellular effects depending on the subtype of receptor that binds a ligand. The hormone epinephrine, for example, which mediates the fight-or-flight response, binds to multiple subtypes of GPCRs in multiple cell types, with varying physiological effects.
- Efforts to de-orphan GPCRs led to the discovery of orexins, hormones that regulate feeding behavior and sleep in both animals and humans.

15.4 Regulating Metabolism of Many Cells: G Protein–Coupled Receptors That Activate or Inhibit Adenylyl Cyclase

To illustrate the importance of GPCRs in human metabolism and physiology, in this section we consider a GPCR that is activated by the hormone **epinephrine**. Epinephrine is particularly important in mediating the body's response to stress, also known as the *fight-or-flight response*. During moments of fear or heavy exercise, when the need for energy may increase, epinephrine (or its derivative, norepinephrine) is released by the adrenal gland. This hormone binds to β_2-adrenergic receptors on the surfaces of hepatic (liver) and adipose (fat) cells, and induces these target cells to rapidly break down either glycogen to glucose (liver) or triacylglycerols to fatty acids (fat). Within seconds of their generation, these metabolic fuels are secreted into the blood and taken up by muscle and other cells, which use them to generate ATP.

The activated β_2-adrenergic receptor induces GDP $\rightarrow$ GTP exchange on a *stimulatory* G protein (G_s). The released GTP-bound alpha subunit ($^{GTP}G_{\alpha s}$) activates a membrane-bound effector enzyme called **adenylyl cyclase** (see Figures 15-16 and 15-18). Once activated, this enzyme catalyzes synthesis from ATP of the second messenger cAMP, which is able to diffuse throughout the cell

(Figure 15-18). cAMP activates a specific protein kinase, called *protein kinase A (PKA)*, that in turn phosphorylates effector proteins that mediate the diverse metabolic effects in different cells.



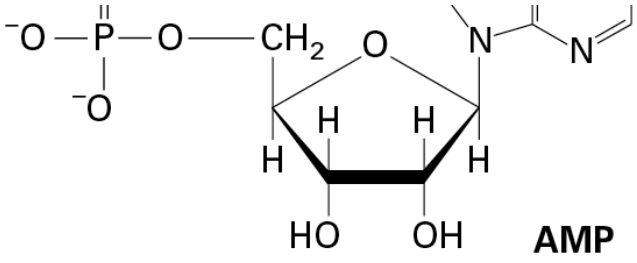


FIGURE 15-18 Synthesis of cAMP by adenylyl cyclase and hydrolysis of cAMP by cAMP phosphodiesterase (PDE). Similar reactions occur for production of cGMP from GTP and hydrolysis of cGMP.

In mammals, more than 30 different GPCRs activate G_s and adenylyl cyclase; most cell types express one or more such GPCRs. Early studies of the mechanism of activation of adenylyl cyclase led to the discovery of the role of the first GTP-binding protein in receptor signaling and to a Nobel Prize.

Adenylyl Cyclase Is Stimulated and Inhibited by Different Receptor-Ligand Complexes

To explore the GPCR $\rightarrow$ cAMP pathways, we focus on an important metabolic pathway: the generation of glucose-1-phosphate from *glycogen*, a storage polymer of glucose (Figure 15-19). The breakdown of glycogen (*glycogenolysis*) is a principal way in which glucose is made available to cells in need of energy. Glycogen breakdown occurs in muscle and liver cells in response to hormones

such as epinephrine and the polypeptide hormone glucagon. This example shows how activation of a GPCR can stimulate or inhibit several intracellular enzymes, all coordinated to carry out a physiologically important task: glycogen metabolism.

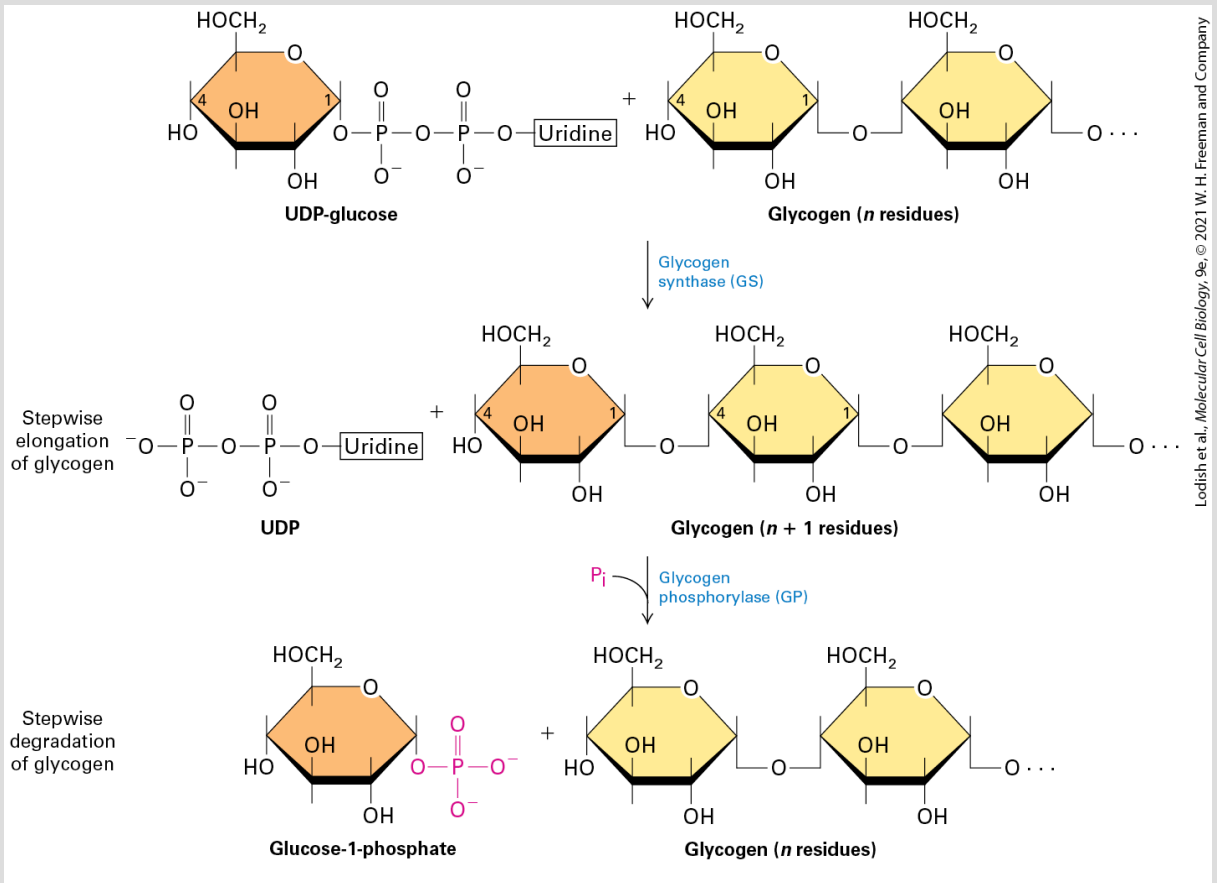


FIGURE 15-19 Synthesis and degradation of glycogen. Incorporation of glucose from UDP-glucose into glycogen is catalyzed by glycogen synthase. Removal of glucose units from glycogen is catalyzed by glycogen phosphorylase. Because two different enzymes catalyze the formation and the degradation of glycogen, the two reactions can be independently regulated.

In case of sudden danger, epinephrine is released by the adrenal glands. When body demand for glucose is high, such as when blood

sugar is low (hypoglycemia) or during exercise, glucagon is released by the α cells of the pancreatic islets. Epinephrine and glucagon bind to different G protein–coupled receptors — the β_2-adrenergic receptor and the glucagon receptor, respectively — but both receptors activate the same G_s protein, whose $^{GTP}G_{\alpha s}$ activates adenylyl cyclase. Hence both hormones induce the same metabolic responses.

Regulation of adenylyl cyclase activity, which occurs in many cell types, provides fine-tuned control of the cAMP level and thus of the downstream cellular response (Figure 15-20). Whereas epinephrine and glucagon activate adenylyl cyclase, other GCPR ligands may inhibit it. For example, binding of two other hormones, prostaglandin E_1 (PGE_1) and adenosine, to their respective G protein–coupled receptors activates the inhibitory G_i protein, whose active $^{GTP}G_{\alpha i}$ subunit binds to and inhibits adenylyl cyclase, lowering cAMP levels and shutting off cellular responses.

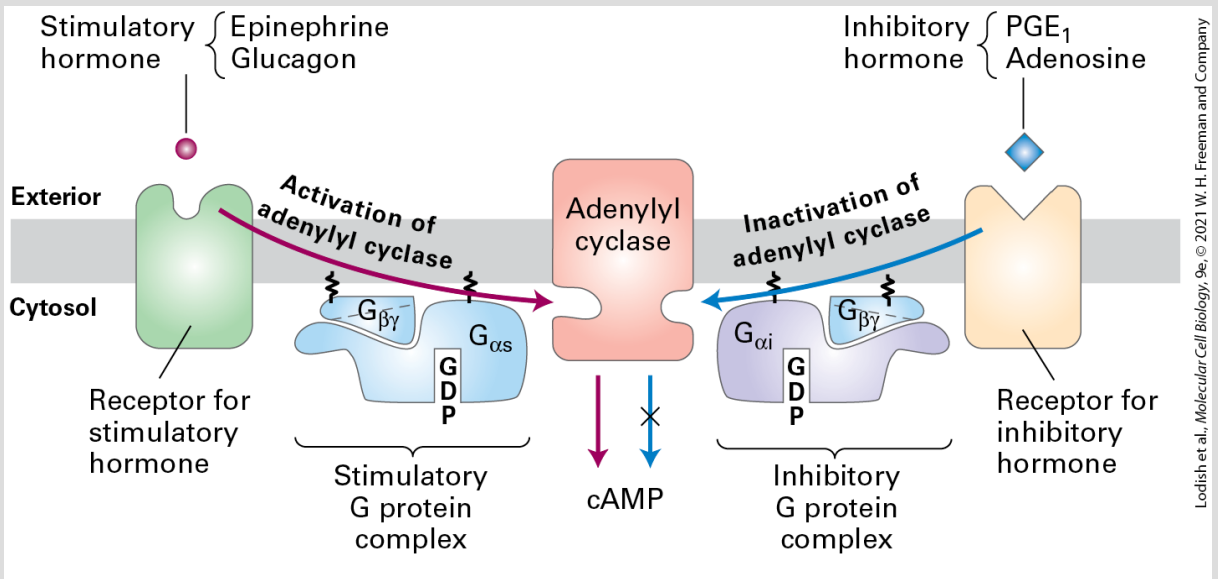


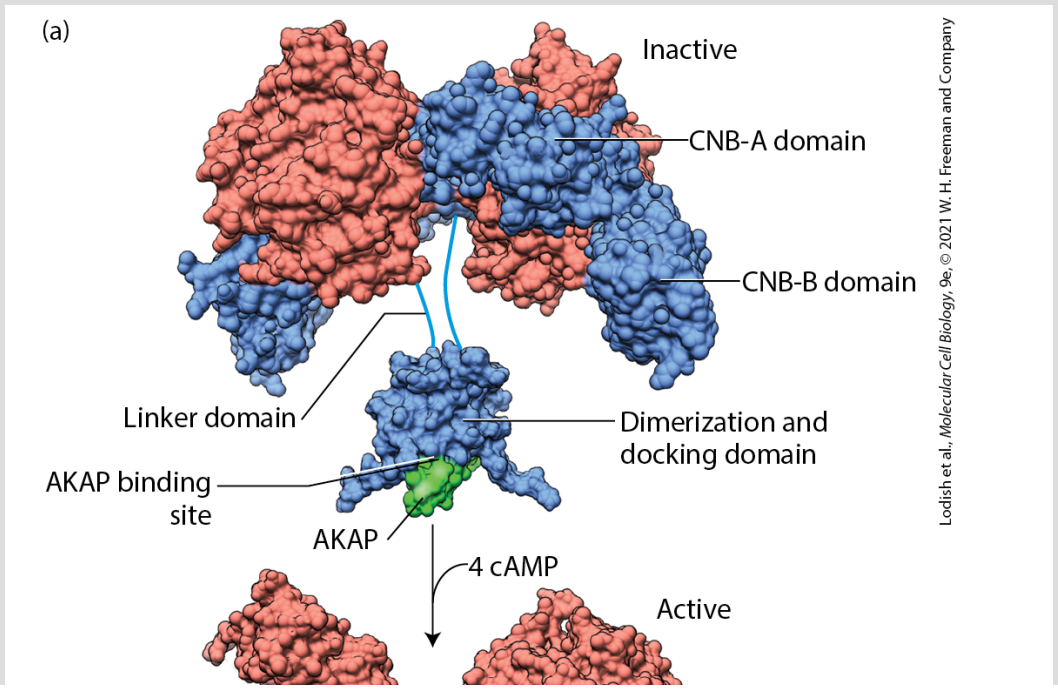
FIGURE 15-20 Hormone-induced activation and inhibition of adenylyl cyclase in adipose cells. Ligand binding to G_s-coupled receptors causes activation of adenylyl cyclase, whereas ligand binding to G_i-coupled receptors causes inhibition of the enzyme. The $G_{\beta\gamma}$ subunit in stimulatory and in inhibitory G proteins is identical; the G_α subunits and their corresponding receptors differ. Active $G_\alpha \cdot GTP$ complexes are formed by the same mechanism in both $G_{\alpha s}$ and $G_{\alpha i}$ proteins (see Figure 15-16). However, $G_{\alpha s} \cdot GTP$ and $G_{\alpha i} \cdot GTP$ interact differently with adenylyl cyclase, so that one stimulates and the other inhibits its catalytic activity. See A. G. Gilman, 1984, *Cell* **36**:577.

cAMP Activates Protein Kinase A by Releasing Its Inhibitory Subunits

The second messenger cAMP, synthesized by adenylyl cyclase, is a component of many different signal transduction pathways initiated by different receptor-ligand complexes and leading to distinct cellular responses in different cell types in multicellular animals. Virtually all of the diverse effects of cAMP are mediated through its

activation of *protein kinase A (PKA)*, also called *cAMP-dependent protein kinase*. This protein kinase phosphorylates multiple intracellular target proteins expressed in different cell types.

Inactive PKA is a tetramer consisting of two regulatory (R) subunits and two catalytic (C) subunits (Figure 15-21a). Each of the two R subunits in an R_2 dimer has two cAMP-binding domains, called *CNB-A* and *CNB-B*, as well as a linker domain of unknown structure. At the C-terminus there is a dimerization domain that binds an AKAP protein discussed later. A segment of the CNB-A domain of the R subunit is a *pseudosubstrate*; its sequence resembles that of a peptide substrate and binds to the active site in the catalytic domain but is not phosphorylated. In the absence of cAMP the pseudosubstrate segment binds to this catalytic site and inhibits its activity.



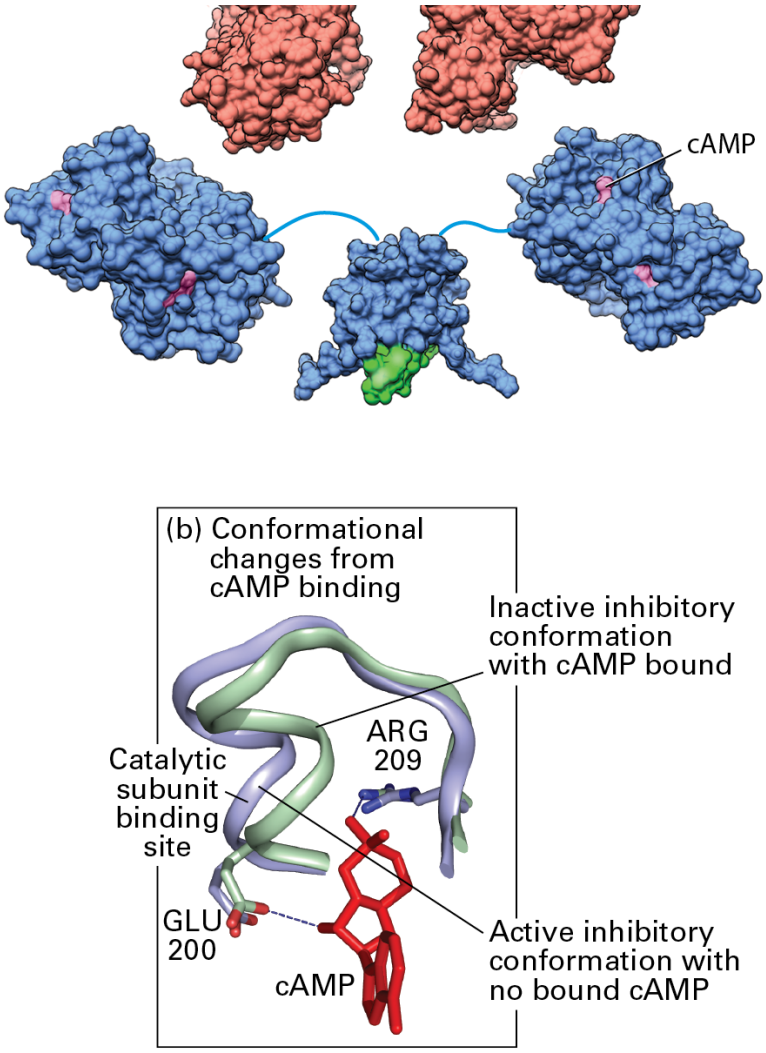


FIGURE 15-21 Structure of PKA and its activation by cAMP. (a) PKA consists of two regulatory (R) subunits (blue) and two catalytic (C) subunits (pink). The two R subunits form a dimer, joined by a flexible linker and a dimerization/docking domain to which an A kinase–associated protein (AKAP; see Figure 15-25) can bind. Each R subunit has two cAMP-binding domains, CNB-A and CNB-B. In the absence of cAMP, the CNB-A domain of an R subunit binds to the catalytic site of a C subunit, inhibiting its activity. Binding of cAMP causes a conformational change in the CNB-A domain so that it can no longer bind to and inhibit the catalytic subunit, instantly activating its kinase activity. (b) Binding of cAMP to the CNB-A domain causes a subtle conformational change that displaces the C subunit from the R subunit, leading to kinase activation. Without bound cAMP, one loop of the CNB-A domain (purple) is in a conformation that can bind the catalytic (C) subunit. A glutamate (E200) and arginine (R209) residue participate

in binding of cAMP (red), which causes the conformational change (green) in the loop that prevents binding of the loop to the C subunit.

[Data from C. Kim, N. H. Xuong, and S. S. Taylor, 2005, *Science* **307**:690, PDB ID 1u7e.]

Inactive PKA is turned on by elevations in the cAMP level. Binding of cAMP to the CNB-A domain causes a change in its conformation so that it can no longer bind to and inhibit the catalytic subunit (Figure 15-21b); this releases the C subunit, instantly activating its kinase activity. cAMP binds to an R subunit of PKA in a cooperative fashion; that is, binding of the first cAMP molecule to CNB-B lowers the K_d of CNB-A, and thus the binding of one cAMP stimulates the binding of the second cAMP and release of the R subunit from the kinase. Thus small changes in the level of cytosolic cAMP can cause relatively large changes in the number of dissociated active C subunits and, hence, in kinase activity. We will see that rapid activation of enzymes by signal-triggered dissociation of an inhibitor is a common feature of many signaling pathways.

Glycogen Catabolism Is Stimulated by Hormone-Induced Activation of PKA

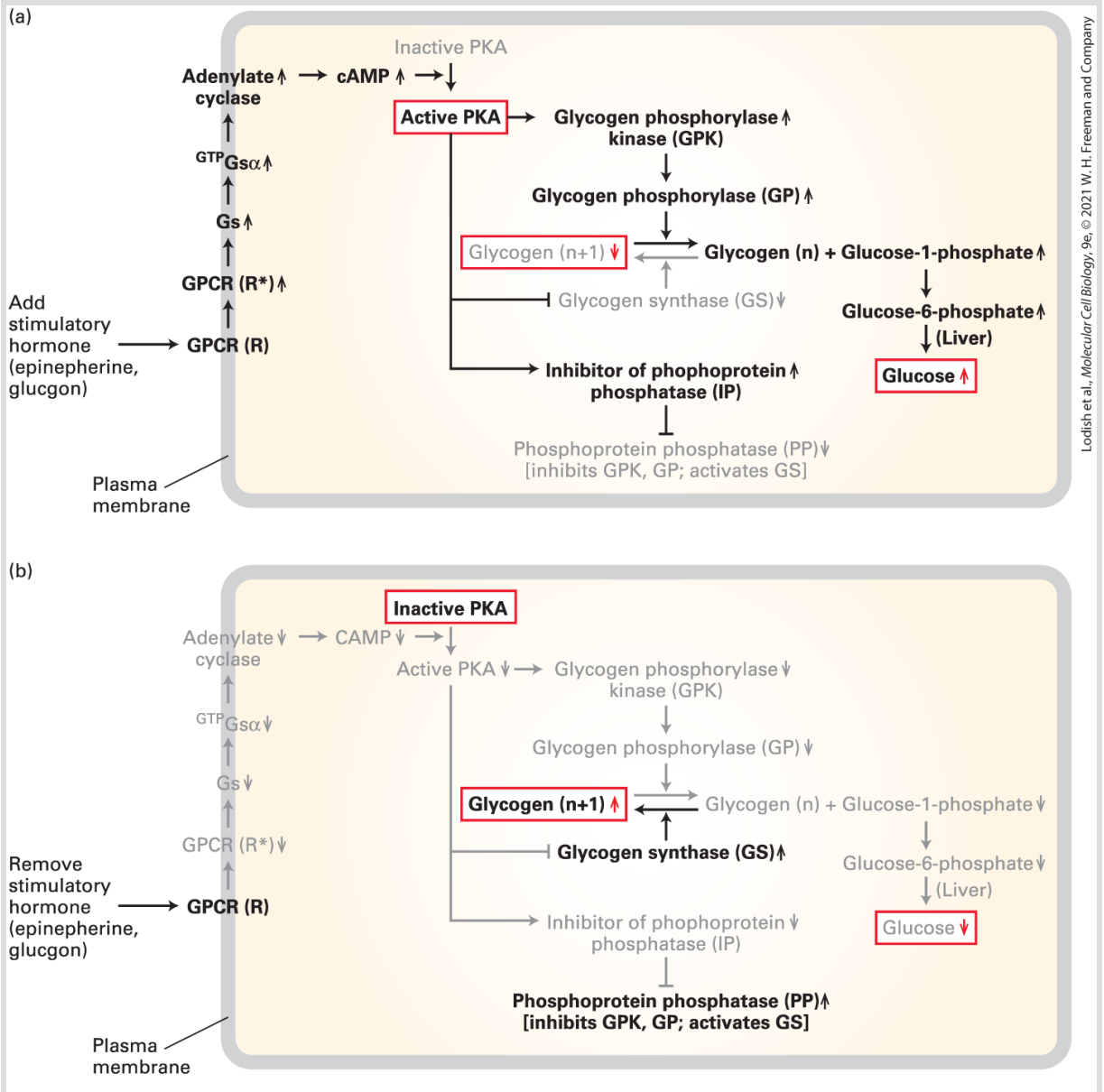
Like all biopolymers, glycogen is synthesized by and degraded by distinct enzymes (see Figure 15-19). Glycogen degradation (glycogenolysis) involves the stepwise removal of single glucose residues from one end of the polymer, whereas glycogen synthesis

(glycogenesis) involves the stepwise addition of single glucose residues. The enzyme *glycogen phosphorylase (GP)* catalyzes the degradation of glycogen via a phosphorolysis reaction yielding glucose-1-phosphate and a polymer shorter by one residue. The enzyme *glycogen synthase (GS)* uses UDP-glucose and glycogen as substrates to generate glycogen that is one residue longer $(n + 1)$.

In both muscle and liver cells, glucose-1-phosphate produced by glycogen degradation is converted by the enzyme *phosphoglucomutase* to glucose-6-phosphate. In muscle cells, glucose-6-phosphate enters the glycolytic pathway and is metabolized to generate ATP for use in powering muscle contraction (see Chapters 12 and 17). Unlike muscle cells, liver cells contain a phosphatase that hydrolyzes glucose-6-phosphate to glucose, which is exported from these cells in part by a glucose transporter (GLUT2) in the plasma membrane (see Chapter 11). Thus glycogen stores in the liver are primarily broken down to glucose, which is immediately released into the blood and transported to other tissues, particularly the muscles and brain, to nourish them. Glycogen breakdown is induced in both types of cells by increases in blood epinephrine, and hence activation of its GPCRs, as part of the fight-or-flight response.

Now we can see the final steps in the signaling pathway leading to the cellular response: glycogen breakdown. Activated PKA enhances the conversion of glycogen to glucose-1-phosphate in two ways: by

inhibiting glycogen synthesis and by *stimulating* glycogen degradation (Figure 15-22a). PKA directly phosphorylates and, in so doing, *inactivates* glycogen synthase (GS). PKA promotes glycogen degradation indirectly by phosphorylating and thus activating an intermediate kinase, glycogen phosphorylase kinase (GPK). In turn, active GPK phosphorylates and activates GP, the enzyme that degrades glycogen.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 15-22 Regulation of glycogen metabolism by cAMP and PKA. Active enzymes are highlighted in black; inactive forms, in gray. (a) An increase in cytosolic cAMP activates PKA, which phosphorylates glycogen synthase (GS), inhibiting its catalytic activity to synthesize glycogen. Active PKA also promotes glycogen degradation via a protein kinase cascade that activates glycogen phosphorylase, the enzyme that removes glucose moieties from glycogen (Figure 15-19). At high cAMP concentrations, PKA also phosphorylates an inhibitor of phosphoprotein phosphatase (PP). Binding of the phosphorylated inhibitor to PP prevents this phosphatase from de-phosphorylating the activated enzymes in the glycogen phosphorylase kinase cascade or the inactive glycogen synthase. (b) A decrease in cAMP

inactivates PKA, leading to release of the active form of PP. The activation of PP promotes glycogen synthesis and inhibits glycogen degradation.

The GPK in skeletal muscle is a huge protein with multiple subunits. One subunit contains the kinase catalytic activity that activates GP and the other subunits are regulatory. One of the regulatory subunits is the ubiquitous protein *calmodulin*, which has four calcium ion binding sites (see Figure 3-34). In muscle, rises in the level of cytosolic Ca^{2+} trigger contraction and also an enhancement of glycogen breakdown to generate the ATP to power contraction. Two of the regulatory subunits contain phosphorylation sites for PKA. GPK enzyme activity is increased either by phosphorylation of these two regulatory subunits by PKA or by Ca^{2+} binding to the calmodulin subunit; maximal activity requires both.

Activated PKA enhances the conversion of glycogen to glucose-1-phosphate in yet a third way. The activities of GPK and GP are inhibited by a phosphatase called *phosphoprotein phosphatase (PP)*, which removes the phosphates added by PKA. In contrast, the activity of GS is increased when it is de-phosphorylated by PP. To prevent these reactions from happening, at high cAMP levels PKA phosphorylates an inhibitor of phosphoprotein phosphatase (IP), which enables it to bind to and keep this phosphatase in its inactive state (see Figure 15-22a). By this means, GPK, GP, and GS remain in their phosphorylated states, promoting glycogen breakdown.

The entire process is reversed when epinephrine or another hormone activating $G_{\alpha s}$, such as glucagon, is removed and the level of cAMP drops, inactivating PKA. When PKA is inactive, it can no longer phosphorylate IP, the inhibitor of phosphoprotein phosphatase, so the inhibitor is released and PP becomes active (Figure 15-22b). PP removes the phosphate residues previously added by PKA to GS and GPK, as well as the phosphates on GP added by GPK. As a consequence, GS becomes activated and the synthesis of glycogen by GS is enhanced. Conversely, GP becomes inactivated, and the degradation of glycogen by GP is inhibited. Thus less glucose-6-phosphate or glucose is produced.

Whether induced by epinephrine or glucagon, glycogen breakdown thus exhibits dual regulation: activation of the enzymes catalyzing glycogen degradation and inhibition of enzymes promoting glycogen synthesis. Such dual regulation is an efficient mechanism for regulating a particular cellular response and is a common phenomenon in cell biology.

Signal Amplification Occurs in the cAMP-PKA Glycogen Degradation Pathway

Receptors such as the β-adrenergic receptor are low in abundance; typically only a few hundred or a thousand copies are present in a

cell. Yet the cellular responses induced by a hormone such as epinephrine can require production of large numbers of cAMP and activated enzyme molecules per cell. As an example, following activation of G_s-coupled receptors, the intracellular concentration of cAMP must rise to about 10^{-6} M to induce appropriate cellular responses. In a typical cell that is roughly a cube about 15 μm on a side, this concentration is equivalent to 2 million molecules of cAMP per cell. Thus substantial amplification (10^2–10^3-fold) of the signal is necessary if it is to induce a significant cellular response.

Signal amplification is possible in part because both receptors and G proteins can diffuse rapidly in the plane of the plasma membrane. A single epinephrine-GPCR complex converts up to a hundred inactive G_s molecules to the active ${}^{GTP}G_{\alpha s}$ form before epinephrine dissociates from the receptor. Each active ${}^{GTP}G_{\alpha s}$, in turn, activates a single adenylyl cyclase molecule, but that adenylyl cyclase molecule then catalyzes the synthesis of many cAMP molecules during the time ${}^{GTP}G_{\alpha s}$ is bound to it.

The degree of amplification that occurs in such a signal transduction cascade depends on the number of steps in it and on the relative concentrations of its various components. In the epinephrine-induced cascade shown in Figure 15-23, for example, blood levels of epinephrine as low as 10^{-10} M can stimulate glycogenolysis in the liver and release of glucose. An epinephrine stimulus of this magnitude generates an intracellular cAMP concentration of

10^{-6} M, an amplification of 10^4-fold. Because three more catalytic steps precede the release of glucose, another 10^4 amplification can occur, resulting in a 10^8 amplification of the epinephrine signal.

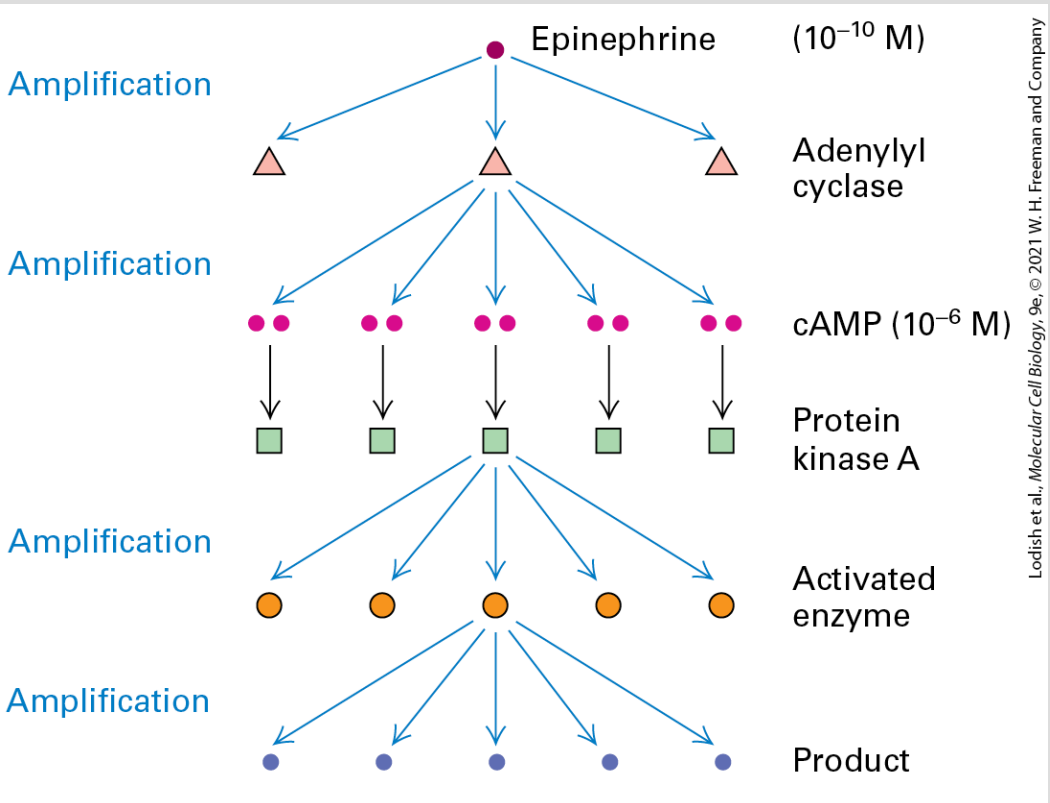


FIGURE 15-23 Amplification of an extracellular signal by a signal transduction pathway involving cAMP and PKA. Binding of a single epinephrine molecule to one G protein–coupled receptor induces activation of many molecules of adenylyl cyclase, the enzyme that catalyzes the synthesis of cyclic AMP, and each of these enzyme molecules synthesizes a large number of cAMP molecules, the first level of amplification. Two molecules of cAMP activate one PKA, but each activated PKA phosphorylates and activates multiple target proteins. This second level of amplification involves several sequential reactions in which the product of one reaction activates the enzyme catalyzing the next reaction. The more steps in such a cascade, the greater the signal amplification possible.

cAMP-Mediated Activation of PKA Produces Diverse Responses in Different Cell Types

Epinephrine also induces the activation of PKA in adipose cells. In these cells, PKA promotes phosphorylation and activation of the lipase that hydrolyzes stored triglycerides to yield free fatty acids and glycerol. These fatty acids are released into the blood and taken up as an energy source by cells in other tissues, such as the kidney, heart, and muscles (see Chapter 12). Therefore, activation of PKA by epinephrine in two different cell types, liver and adipose cells, has different effects. Indeed, cAMP and PKA mediate a large array of hormone-induced cellular responses in multiple tissues (Table 15-3).

TABLE 15-3 • Cellular Responses to Hormone-Induced Rise in cAMP in Various Tissues[ii]

Tissue	Hormone Inducing Rise in cAMP	Cellular Response
Adipose	Epinephrine; ACTH; glucagon	Increase in hydrolysis of triglyceride; decrease in amino acid uptake
Liver	Epinephrine; norepinephrine; glucagon	Increase in conversion of glycogen to glucose; inhibition of glycogen synthesis; increase in amino acid uptake; increase in gluconeogenesis (synthesis of glucose from amino acids)
Ovarian	FSH; LH	Increase in synthesis of estrogen, progesterone

follicle		
Adrenal cortex	ACTH	Increase in synthesis of aldosterone, cortisol
Cardiac muscle	Epinephrine	Increase in contraction rate
Thyroid gland	TSH	Secretion of thyroxine
Bone	Parathyroid hormone	Increase in resorption of calcium from bone
Skeletal muscle	Epinephrine	Conversion of glycogen to glucose-1-phosphate
Intestine	Epinephrine	Fluid secretion
Kidney	Vasopressin	Resorption of water
Blood platelets	Prostaglandin I	Inhibition of aggregation and secretion

[ii]Nearly all the effects of cAMP are mediated through PKA, which is activated by binding of cAMP.
SOURCE: E. W. Sutherland, 1972, *Science* **177**:401.

Although PKA acts on different substrates in different types of cells, it always phosphorylates a serine or threonine residue that is located within the common sequence motif: X-Arg-(Arg/Lys)-X-(Ser/Thr)-Φ, where X denotes any amino acid and Φ denotes a hydrophobic amino acid. The Ser or Thr whose side chain is phosphorylated is underlined. Thus many potential protein targets of PKA can be

identified simply by analyzing their amino acid sequence. Other serine/threonine kinases phosphorylate target residues within other sequence motifs, and thus their potential targets can often be predicted as well. Some kinases recognize additional structural and sequence motifs in their targets in addition to the sequence adjacent to the residue to be phosphorylated. In these cases, analysis of the primary amino acid sequence of the protein may not predict whether or not it is a substrate of such kinases.

CREB Links cAMP and PKA to Activation of Gene Transcription

Activation of PKA also stimulates the expression of many genes, leading to long-term effects on cells. For instance, in liver cells, PKA induces the expression of several enzymes involved in *gluconeogenesis* — the conversion of three-carbon compounds such as pyruvate (see Figure 12-3) to glucose. The expression of these genes thus leads to an increase in the concentration of glucose in the blood, thereby enhancing the rapid effects of activated PKA seen in Figure 15-22a.

The promoter or enhancer segments of all genes regulated by PKA contain a cis-acting DNA sequence, the *cAMP-response element (CRE)*, which contains the highly conserved nucleotide sequence, 5′-TGACGTCA-3′. Following the elevation of cAMP levels and the

release of active PKA catalytic subunits, some of the catalytic subunits move into the nucleus. There they phosphorylate serine-133 on a transcription factor called *CRE-binding (CREB) protein*. Phosphorylated CREB binds to CRE-containing target genes and also binds to the *co-activator* termed *CBP/P300* (see Figure 8-26). CBP/P300 links CREB to RNA polymerase II and other gene regulatory proteins, thereby stimulating gene transcription (Figure 15-24).

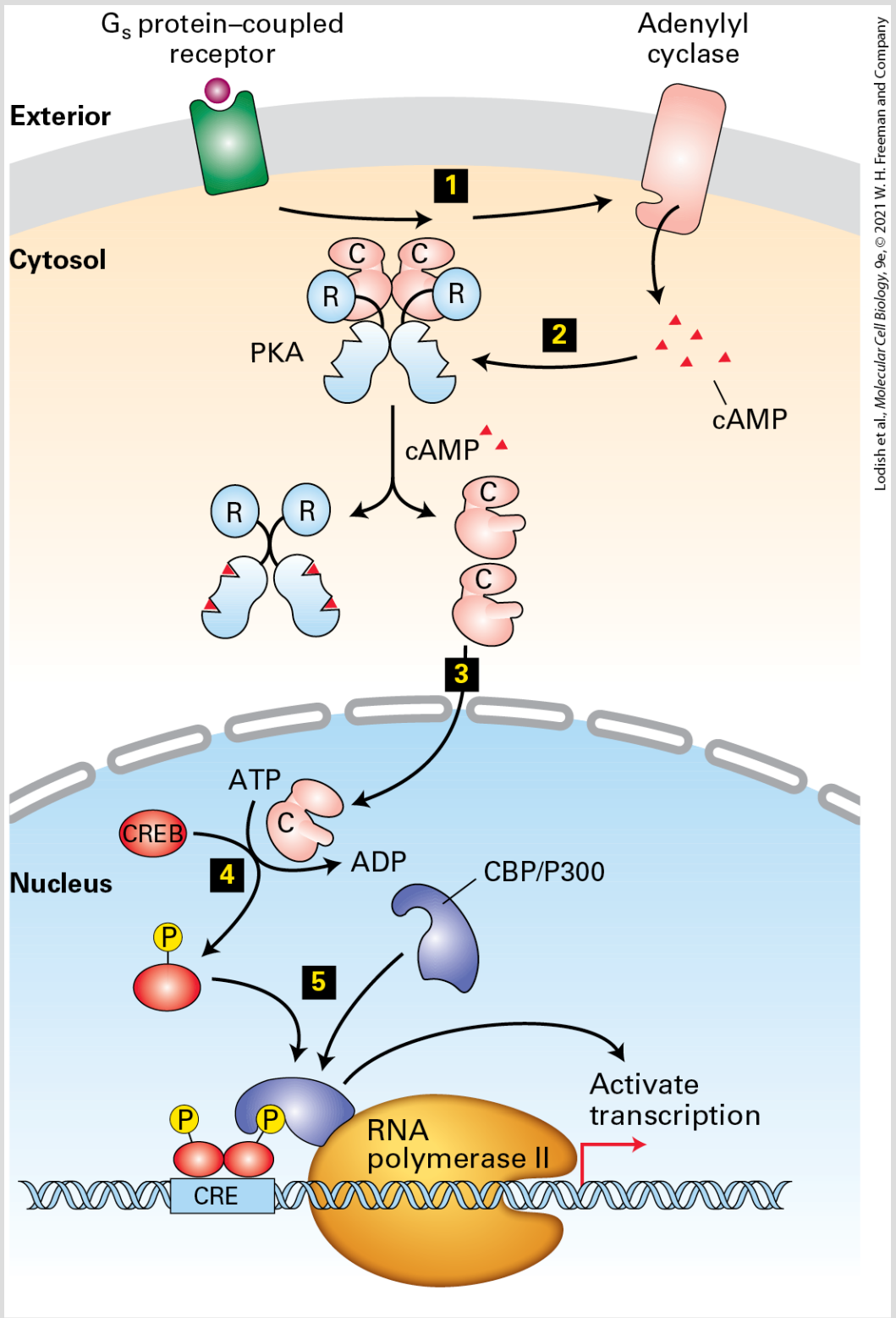


FIGURE 15-24 Activation of CREB transcription factor following ligand binding to $G_{\alpha s}$-coupled GPCRs. Receptor stimulation (step 1) leads to activation of PKA (step 2). Catalytic subunits of PKA translocate to the

nucleus (step 3) and there phosphorylate and activate the CREB transcription factor (step 4). Phosphorylated CREB associates with the co-activator CBP/P300 (step 5) and other proteins to stimulate transcription of the various target genes controlled by a CRE regulatory element. See K. A. Lee and N. Masson, 1993, *Biochim. Biophys. Acta* **1174**:221; and D. Parker et al., 1996, *Mol. Cell Biol.* **16**(2):694.

Anchoring Proteins Localize Effects of cAMP to Specific Regions of the Cell

In many cell types, a rise in the cAMP level may produce a response that is required in one part of the cell but is unneeded, or perhaps deleterious, in another. Anchoring proteins confine members of the PKA family to specific locations in the cell, thereby restricting cAMP-dependent responses to these locations. Each of the roughly 50 anchoring proteins, referred to as *A kinase–associated proteins (AKAPs)*, has a two-domain structure: one domain confers a specific subcellular location by binding to a protein found only in a specific subcompartment, and the other binds to the regulatory (R) subunit of PKA. AKAPs regulate cAMP and PKA signaling within the cell both spatially and temporally.

One AKAP in heart muscle anchors PKA to the outer nuclear membrane (Figure 15-25). From this location, the released, catalytically active C subunits readily enter the nucleus, where they

phosphorylate and activate the CREB transcription factor (see Figure 15-24).

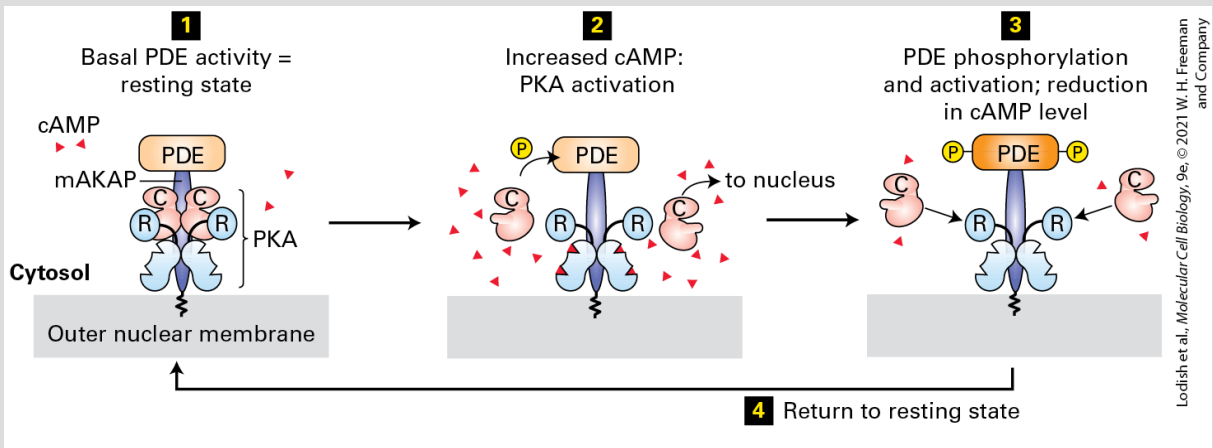


FIGURE 15-25 Localization of PKA and cAMP phosphodiesterase (PDE) to the nuclear membrane in heart muscle by an A kinase–associated protein (AKAP). This member of the AKAP family, designated mAKAP, anchors both PDE and the regulatory subunit (R, see Figure 15-21a) of PKA to the cytosolic surface of the outer nuclear membrane, maintaining them in a negative feedback loop that closely controls the cAMP level and PKA activity in this local segment of the cell. Step **1**: The basal level of PDE activity in the absence of hormone (resting state) keeps cAMP levels below those necessary for PKA activation. Steps **2** and **3**: Activation of β-adrenergic receptors causes an increase in cAMP to a level in excess of that which can be degraded by PDE. The resulting binding of cAMP to the R subunits of PKA releases the active catalytic (C) subunits into the cytosol. Some C subunits enter the nucleus, where they phosphorylate and thus activate the transcription factor CREB (see Figure 15-24). C subunits also phosphorylate PDE, stimulating its catalytic activity. Active PDE hydrolyzes cAMP, thereby driving cAMP levels back to basal levels and causing re-formation of the inactive PKA C-R complex. Step **4**: Subsequent de-phosphorylation of PDE returns the complex to the resting state. See K. L. Dodge et al., 2001, *EMBO J.* **20**:1921.

Also anchored to the same AKAP is cAMP phosphodiesterase (PDE), the enzyme that hydrolyzes cAMP to AMP (see Figure 15-18). The anchoring of PKA and PDE to the same AKAP keeps these two

proteins in close proximity, where the action of PDE confines the activity of PKA to short local bursts. The cAMP concentration and hence local PKA activity is tightly controlled by negative feedback. As cAMP levels rise in response to hormone stimulation, PKA becomes activated and then phosphorylates and activates several target proteins, including PDE. The now active PDE, in turn, hydrolyzes cAMP, thus quickly returning PKA to its inactive state and allowing only a short burst of PKA activity.

In certain heart muscle cells, a different AKAP is tethered to the cytosolic face of the plasma membrane near a particular type of gated Ca^{2+} channel. In the heart, these channels open in response to activation of β-adrenergic receptors by epinephrine (as part of the fight-or-flight response). The resulting influx of Ca^{2+} increases the rate of heart muscle contraction. These Ca^{2+} channels open when phosphorylated by PKA. The binding of AKAP to PKA localizes the kinase next to these channels, thereby reducing the time that would otherwise be required for PKA catalytic subunits to diffuse from their sites of generation to their Ca^{2+}-channel substrates.

Multiple Feedback Mechanisms Suppress Signaling by the GPCR/cAMP/PKA Pathway

In Section 15.1 we stressed that, for cells to respond effectively to a hormone or to other changes in their environment, they must not only activate a signal transduction pathway, but also reduce or terminate the response once it is no longer needed; otherwise, signal transduction pathways would remain "on" too long, or at too high a level, and the cell would become overstimulated.

The activity of most G protein–coupled receptors is reduced or even totally inhibited by multiple feedback mechanisms, as is exemplified by β-adrenergic and other receptors that are coupled to G_s and that activate adenylyl cyclase.

- First, the intrinsic GTPase activity of $G_{\alpha s}$ converts bound GTP to GDP, thereby terminating the ability of $G_{\alpha s}$ to activate its downstream target adenylyl cyclase. Importantly, the rate of hydrolysis of GTP bound to $G_{\alpha s}$ is enhanced when ${}^{GTP}G_{s\alpha}$ binds to adenylyl cyclase. Thus the downstream effector adenylyl cyclase also functions as a GAP for ${}^{GTP}G_{s\alpha}$, increasing the rate of inactivation of ${}^{GTP}G_{s\alpha}$ and so lessening the duration of cAMP production. More generally, many effector proteins of G_α subunits are also GAPs — the binding of most, if not all, ${}^{GTP}G_\alpha$'s to their respective effector proteins accelerates the rate of GTP hydrolysis.
- Second, PDE (see Figure 15-18) acts to hydrolyze cAMP to 5′-AMP, terminating the cellular response. Thus the hormone must be continuously present at a sufficient concentration in order to keep adenylyl cyclase activated and maintain an

elevated cAMP level. Once the extracellular hormone concentration falls sufficiently, the cAMP level falls due to PDE action and most cAMP-dependent rapid cellular responses quickly wane.

- Most GPCRs are down-regulated by *feedback repression*, in which an end product of a signaling pathway blocks an early step in that pathway. For instance, when a G_s protein–coupled receptor is exposed to hormonal stimulation for even short periods of time — seconds to minutes depending on the receptor — active PKA phosphorylates several serine and threonine residues typically in the third cytoplasmic loop or the C-terminal cytosol-facing tail domain of the receptor. The phosphorylated receptor can bind its ligand, but because these cytosolic domains participates in binding to the $G_{\alpha s}$ subunit, the receptor is less able to activate G_s. In other words, ligand bound to the phosphorylated receptor is less efficient in activating adenylyl cyclase than is ligand bound to the nonphosphorylated receptor. Because the activity of PKA is enhanced by the high cAMP level induced by any hormone that activates G_s, prolonged exposure to one such hormone — say, epinephrine — desensitizes not only β-adrenergic receptors, but also other $G_{\alpha s}$–coupled receptors that are phosphorylated by PKA, even though they bind different ligands (e.g., glucagon receptors in the liver). This cross-regulation is called *heterologous desensitization*.

A second type of phosphorylation-mediated feedback control occurs only when the GPCR is in its ligand-bound, activated conformation (R*). In the R*, but not R, conformation, serine and threonine

residues in the third cytoplasmic loop or the C-terminal cytosol-facing domain of the receptor can be phosphorylated by a special group of kinases called *G-protein–coupled receptor kinases (GRKs)* (Figure 15-26a, b). GRKs recognize the subtle conformational differences between R and R* and phosphorylate only the active R* conformation. The longer the receptor remains in its active state, the greater the extent of its phosphorylation by GRKs. Importantly, the greater the extent of the phosphorylation by a GRK, the lower the ability of the receptor to activate the downstream G protein. This process is termed *homologous desensitization* because a ligand can only induce the deactivating phosphorylation of receptors to which it is bound — those in the R* conformation. Unlike what happens in heterologous desensitization, a ligand cannot induce the phosphorylation by GRK of a receptor to which it cannot bind.

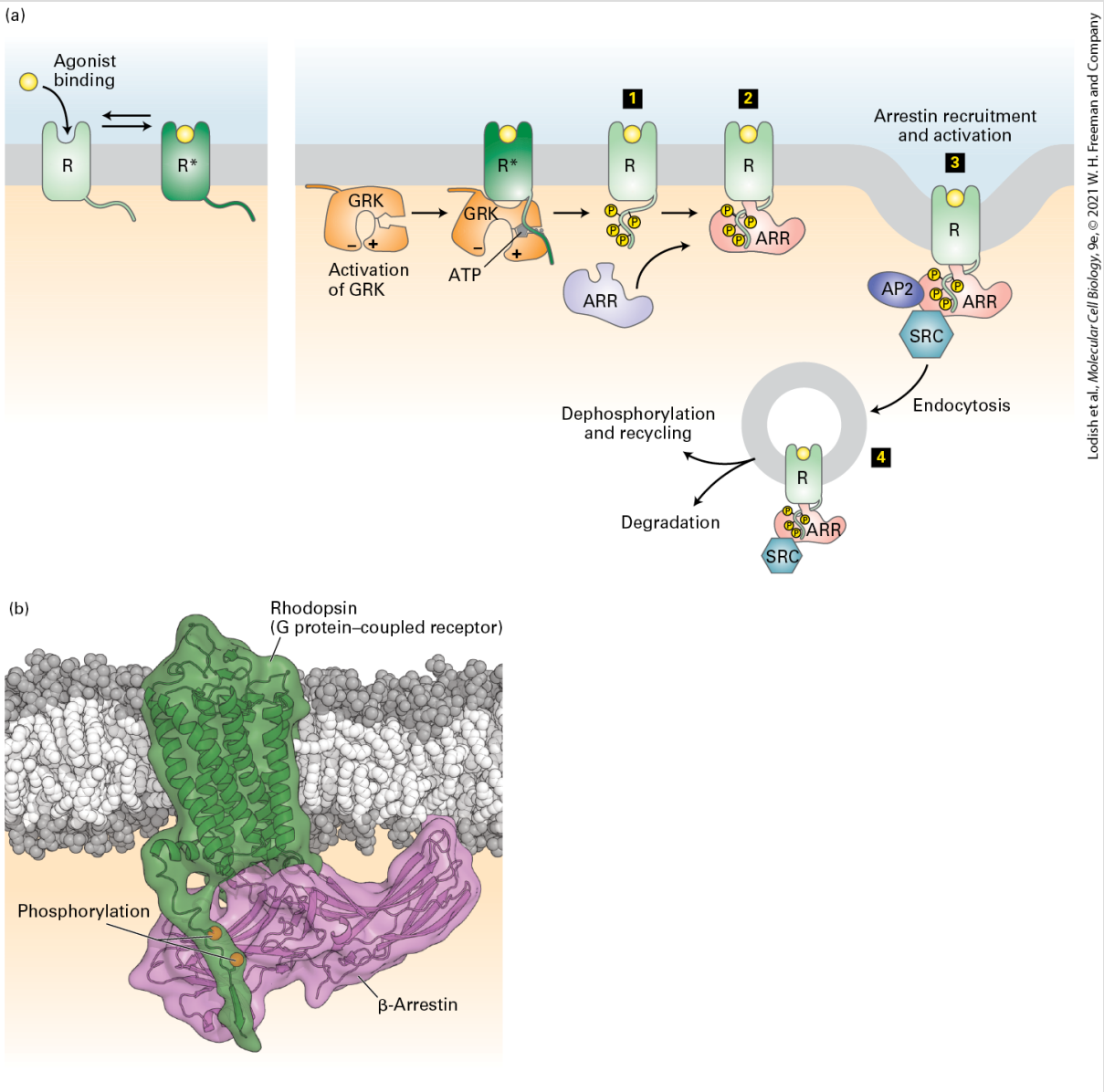


FIGURE 15-26 Binding of β-arrestin to phosphorylated GPCRs triggers receptor desensitization as well as activation of different signal transduction proteins. (a) A G-protein–coupled receptor kinase (GRK) phosphorylates serine and threonine residues in the cytoplasmic segments of activated G protein–coupled receptors (GPCRs), those that have bound agonists (step 1). β-Arrestin (Arr) then binds to these phosphorylated receptors (step 2). Steps 3 and 4: Clathrin and AP2, two other proteins bound by β-arrestin, promote endocytosis of the receptor (see Figure 14-29). β-Arrestin also functions in transducing signals from activated receptors by binding to and activating several cytosolic protein kinases. One such kinase, Src, activates the MAP kinase pathway that, as detailed in the next chapter, leads to phosphorylation and activation of key transcription factors. (b)

Three-dimensional structure of the GPCR rhodopsin bound to arrestin. Arrestin binds to segments of the C-terminal cytosolic alpha helix of activated rhodopsin that includes the two phosphorylated resides as well as to parts of transmembrane helix 7.

[Part (a) Data from W. Weis and B. Kobilka, 2018, *Annu. Rev. Biochem.* **87**:897. Part (b) Data from Y. Kang et al., 2015, *Nature* **523**:561–567, PDB ID 4zwj, and custom PDB.]

An important breakthrough in understanding the control of GPCR-mediated signal transduction came with the discovery of adapter proteins called *arrestins*. Arrestins bind tightly to GPCRs only when these receptors have been phosphorylated by GRKs on two or three of the serine residues in the cytosol facing C-terminal domain (see Figure 15-26a, step **2**). Arrestin binding can dramatically speed up the inactivation process. Arrestin bound to the phosphorylated receptor completely prevents its interaction with a G protein, blocking formation of the active $^{GTP}G_{\alpha}$ and preventing further downstream signaling. Eventually, phosphatase enzymes remove the phosphates added by a GRK and prevent arrestin binding, restoring the receptor to its original, hormone-sensitive state.

Another function of arrestin is to help terminate the signal by reducing the numbers of activated GPCRs on the cell surface. Subsequent studies revealed that β-arrestin binds not only to phosphorylated GPCRs, but also to clathrin and an associated protein termed *AP2*, two key components of the coated vesicles that are involved in endocytosis from the plasma membrane (see Figure 15-26a, steps **3** and **4**; see also Chapter 14). These interactions promote the endocytosis of the associated receptors, thereby

decreasing the number of receptors exposed on the cell surface. Eventually some of the receptors that have been moved inside the cell are degraded, and some are de-phosphorylated in endosomes. Following dissociation of β-arrestin, the resensitized, de-phosphorylated GPCRs are recycled to the cell surface in a manner similar to the recycling of the LDL receptor (see Chapter 14).

In addition to their role in regulating receptor activity, arrestins bound to the phosphorylated GPCR can function as adapter proteins that recruit additional downstream signaling proteins to the GPCR. As a consequence, the GPCR can initiate a second signal transduction pathway that is independent of G proteins. The GPCR-arrestin complex acts as a scaffold for the binding and activation of several cytosolic kinases (see Figure 15-26a, step **4**), which we discuss in detail in subsequent chapters. These kinases include Src, a protein tyrosine kinase that activates the MAP kinase pathway and other pathways leading to the transcription of genes needed for cell division (see Chapters 16 and 19).

Thus the phosphorylation of GPCRs by GRK followed by arrestin binding can function as a switch, turning off signaling by G proteins and turning on other signaling pathways. The multiple functions of arrestins illustrate the importance of adapter proteins in both regulating signaling and transducing signals from cell-surface receptors.

KEY CONCEPTS OF SECTION 15.4

Regulating Metabolism of Many Cells: G Protein–Coupled Receptors That Activate or Inhibit Adenylyl Cyclase

- Ligand binding by G protein–coupled receptors that activate $G_{\alpha s}$ results in the activation of the membrane-bound enzyme adenylyl cyclase, which converts ATP to the second messenger cAMP. Ligand binding of G protein–coupled receptors that activate $G_{\alpha i}$ results in the inhibition of adenylyl cyclase and lower levels of cAMP.
- $G_{\alpha s}\cdot GTP$ and $G_{\alpha i}\cdot GTP$ bind to the catalytic domain in adenylyl cyclase to activate or inhibit the enzyme, respectively.
- cAMP binds cooperatively to a regulatory subunit of PKA, releasing the active kinase catalytic subunit.
- In liver and muscle cells, activation of PKA induced by epinephrine and other hormones exerts a dual effect, inhibiting glycogen synthesis and stimulating glycogen breakdown via a kinase cascade. The result is an increase in glucose for production of ATP.
- PKA mediates the diverse effects of cAMP in most cells. The substrates for PKA, and thus the cellular responses to hormone-induced activation of PKA, vary among cell types.
- The signal that activates the GPCR/adenylyl cyclase/cAMP/PKA signaling pathway is amplified tremendously by second messengers and kinase cascades.
- Activation of PKA often leads to phosphorylation of nuclear CREB protein, which, together with the CBP/P300 co-activator, stimulates transcription of genes, thus initiating a long-term change in the cell's protein composition.
- Localization of PKA to specific regions of the cell by anchoring proteins restricts the effects of cAMP to particular locations in the cell.
- Signaling from G_s-coupled receptors is down-regulated by multiple mechanisms: first, the intrinsic GTPase activity of $G_{\alpha s}$ that converts the bound GTP to GDP is enhanced when $G_{\alpha s}$ binds to adenylyl cyclase (this occurs when many $G_{\alpha}\cdot GTP$ complexes bind to their respective effector proteins); and second, PDE acts to hydrolyze cAMP to $5'\text{-AMP}$, terminating the cellular response.
- Most GPCRs are regulated by *feedback repression*. They are deactivated by kinases in the G-protein–receptor kinase (GRK) family, which phosphorylate cytosolic residues of the receptor in its active R^* conformation. GRK phosphorylation of ligand-bound β-adrenergic receptors also leads to the binding of arrestin and endocytosis of the

receptors. The consequent reduction in the number of cell-surface receptors renders the cell less sensitive to additional hormone.

- The GPCR-arrestin complex functions as a scaffold that activates several cytosolic kinases, initiating cascades that lead to transcriptional activation of many genes controlling cell growth.

15.5 Regulating Protein Secretion and Muscle Contraction: Ca^{2+} Ions as Second Messengers in Multiple Signal Transduction Pathways

A small rise in cytosolic Ca^{2+} induces a variety of cellular responses, depending on cell type: endocrine cells secrete hormones, pancreatic acinar cells secrete digestive enzymes, neurons release neurotransmitter at the synapse, and muscle fibers contract (Table 15-4). For example, in response to acetylcholine stimulation of GPCRs, secretory cells of the pancreas and of the parotid (salivary) glands secrete digestive proteins. In this case, a rise in cytosolic Ca^{2+} triggers the fusion of secretory vesicles with the plasma membrane and release of their protein contents into the extracellular space. Thrombin, a proteolytic enzyme in the blood-clotting cascade, binds to a GPCR on blood platelets. The binding of thrombin (which activates the GPCR by proteolytic cleavage of the receptor extracellular domain) triggers a rise in cytosolic Ca^{2+}, which in turn causes a shape change in the platelets and other modifications that lead to their aggregation, an important step in blood clotting to prevent blood from leaking out of damaged blood vessels. Thrombin inhibitors are drugs that block blood coagulation.

TABLE 15-4 • Cellular Responses to Hormone-Induced Rise in Cytosolic Ca^{2+} in Various Tissues[iii]

Tissue	Hormone Inducing Rise in Ca^{2+}	Cellular Response
Pancreas (acinar cells)	Acetylcholine	Secretion of digestive enzymes, such as amylase and trypsinogen
Parotid (salivary) gland	Acetylcholine	Secretion of amylase
Vascular or stomach smooth muscle	Acetylcholine	Contraction
Liver	Vasopressin	Conversion of glycogen to glucose
Blood platelets	Thrombin	Aggregation, shape change, secretion of hormones
Mast cells	Antigen	Histamine secretion
Fibroblasts	Peptide growth factors	DNA synthesis, cell division (e.g., bombesin and PDGF)
Nerve cells	Many	Secretion of neurotransmitters

[iii]Hormone stimulation leads to production of inositol 1,4,5-trisphosphate (IP_3), a second messenger that promotes release of Ca^{2+} stored in the endoplasmic reticulum.
SOURCE: M. J. Berridge, 1987, *Annu. Rev. Biochem.* **56**:159; and M. J. Berridge and R. F. Irvine, 1984, *Nature* **312**:315.

The total calcium present in a subcellular compartment is the sum of the free Ca^{2+} ions and bound Ca^{2+}. Free Ca^{2+} ions are Ca^{2+} ions that are not tightly bound to proteins and that can be quantified by the fluorescence techniques mentioned previously. In the ER and mitochondrial matrix space, the amounts of bound Ca^{2+} are thought

to be much greater than those of free Ca^{2+} because these organelles contain abundant Ca^{2+}-binding proteins. As an example, the ER lumen contains several Ca^{2+}-binding proteins, including the chaperones *calreticulin* and *calnexin* (Chapter 13), which have a low affinity and large capacity for Ca^{2+} binding and that buffer the amount of calcium in the lumen of that organelle, ${Ca^{2+}}_{ER}$. Measurements of free $[Ca^{2+}]$ have indicated that, while the concentrations can vary considerably among different types of resting cells, typically $[{Ca^{2+}}_{cytosol}] = \sim 100\ nM$, $[{Ca^{2+}}_{ER}] = \sim 400\ \mu M$, and $[{Ca^{2+}}_{mitochondria}] = \sim 100\ nM$.

The free Ca^{2+} concentration in organelles can vary in response to stimulation by hormones or neuronal signals, but $[{Ca^{2+}}_{cytosol}]$ often increases to about $1\ \mu M$ and $[{Ca^{2+}}_{mitochondria}]$ to approximately $1 - 10\ \mu M$, whereas $[{Ca^{2+}}_{ER}]$ decreases to about $100\ \mu M$. We learned in Chapter 12 that increases in the concentration of free Ca^{2+} in the mitochondrial matrix accelerate pyruvate oxidation and ATP production. Thus, in muscle, increases in the concentration of Ca^{2+} are used both to induce contractions and to increase mitochondrial ATP synthesis to provide the energy to fuel those contractions.

Products of the Hydrolysis of the Membrane Lipid Phosphatidylinositol

4,5-Bisphosphate by Phospholipase C Elevate Cytosolic Ca^{2+} Levels

A number of important second messengers, used in several signal transduction pathways, are derived from the membrane lipid *phosphatidylinositol* (*PI*; Figure 15-27). One of these, IP_3, triggers the elevation of Ca^{2+} in the cytosol and mitochondrial matrix. IP_3 is generated in several steps from the precursor PI.

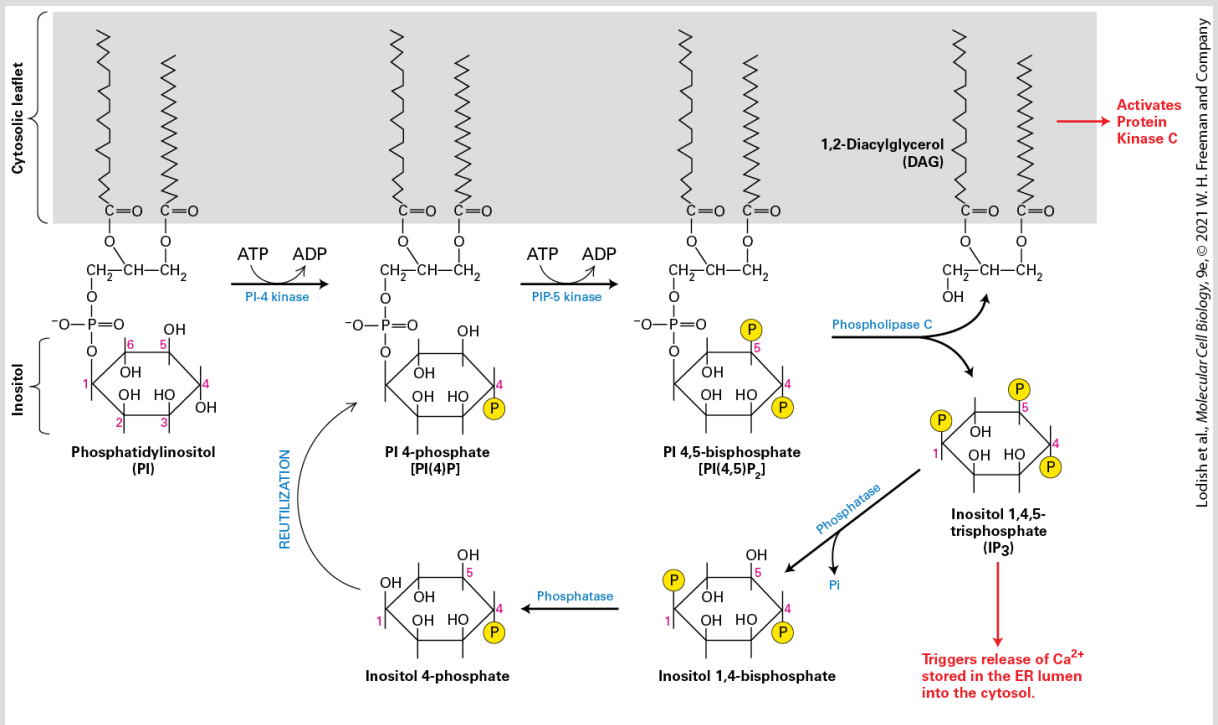


FIGURE 15-27 Synthesis of second messengers DAG and IP_3 from phosphatidylinositol (PI). Each membrane-bound PI kinase places a phosphate (yellow circles) on a specific hydroxyl group on the inositol ring, producing the phosphorylated derivatives PI(4)P and $PI(4,5)P_2$. Cleavage of $PI(4,5)P_2$ by phospholipase C yields the two important second messengers DAG and IP_3. Signaling is terminated when a phosphatase removes the 5-phosphate from IP_3; a second phosphatase removes the 1-phosphate, and the inositol 4-

phosphate is reused to synthesize PI 4-phosphate. See A. Toker and L. C. Cantley, 1997, *Nature* **387**:673; and C. L. Carpenter and L. C. Cantley, 1996, *Curr. Opin. Cell Biol.* **8**:153.

In the plasma membrane this phospholipid resides in the cytoplasmic leaflet with its inositol head group always facing the cytosol. Inositol can be phosphorylated on one or more hydroxyl groups by the combined actions of different kinases, and the phosphates can be removed by phosphatases, as discussed in Chapter 16. One derivative of PI, the lipid phosphatidylinositol 4,5-bis phosphate [$PI(4,5)P_2$], is made by stepwise addition of two phosphates to PI to the hydroxyl groups attached to the #4 and #5 carbon atoms of the inositol ring.

When the enzyme *phospholipase C (PLC)* is activated by a signal transduction pathway, it catalyzes the cleavage of $PI(4,5)P_2$ in the plasma membrane into two important second messengers: 1,2-diacylglycerol (DAG), a lipophilic molecule that remains associated with the plasma membrane, and inositol 1,4,5-trisphosphate (IP_3), which can freely diffuse in the cytosol (see Figures 15-27 and 15-28). We refer to downstream events involving these two second messengers collectively as the *IP_3/DAG pathway*. IP_3, in turn, triggers the release of Ca^{2+} from the ER into the cytosol and mitochondrial matrix.

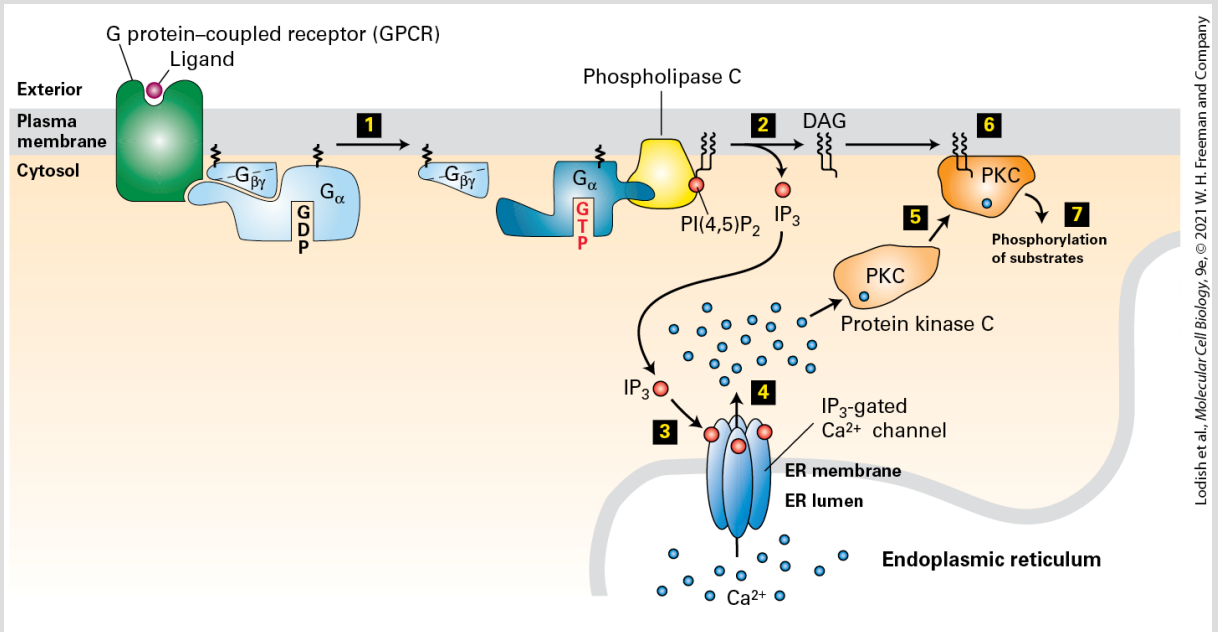


FIGURE 15-28 The IP_3/DAG pathway and the elevation of cytosolic Ca^{2+}. Opening of endoplasmic reticulum Ca^{2+} channels can be triggered by ligand binding to GPCRs that primarily activate members of the G_q family, leading to activation of phospholipase C (step **1**). Cleavage of $PI(4,5)P_2$ by phospholipase C yields IP_3 and DAG (step **2**). After diffusing through the cytosol, IP_3 interacts with and opens IP_3-gated Ca^{2+} channels, often termed *IP_3 receptors* or *IP_3R's* in the membrane of the endoplasmic reticulum (step **3**), causing release of Ca^{2+} ions stored in the ER lumen into the cytosol (step **4**). One of several cellular responses induced by a rise in cytosolic Ca^{2+} is recruitment of protein kinase C (PKC) to the plasma membrane (step **5**), where it is activated by binding DAG (step **6**). The activated membrane-associated kinase can phosphorylate various cellular enzymes and receptors, thereby altering their activity (step **7**). See J. W. Putney, 1999, *Proc. Nat'l. Acad. Sci. USA* **96**:14669; Y. Zhou, 2010, *Proc. Nat'l. Acad. Sci. USA* **107**:4896; M. Cahalan, 2010, *Science* **330**:43; and Q. C. Wang, X. Wang, and T.-S. Tang, 2018, *J. Cell Biol.* **217**(6):1899–1900, https://doi.org/10.1083/jcb.201805037.

Phospholipase C is activated by binding either ${}^{GTP}G_{\alpha o}$ or ${}^{GTP}G_{\alpha q}$ subunits released after GPCR activation (Figure 15-28, step **1**). In turn, activated phospholipase C cleaves $PI(4,5)P_2$ into DAG and IP_3 (see Figure 15-28, step **2**). The two second messengers trigger

distinct, but sometimes interrelated, downstream pathways. Some phospholipase C's are activated by receptors other than GPCRs and are described in the following chapter.

Ca^{2+} Release from the ER Triggered by IP_3

Ligand binding to G protein–coupled receptors that activate phospholipase C induces an elevation in cytosolic Ca^{2+} even when Ca^{2+} ions are absent from the surrounding extracellular fluid. In this case, Ca^{2+} is released into the cytosol from the ER lumen (where, as noted above, Ca^{2+} can accumulate to almost millimolar concentrations by binding to Ca^{2+}-binding proteins) through the operation of *IP_3-gated Ca^{2+} channels* in the ER membrane, as depicted in Figure 15-28 (steps **3** and **4**). Each of these large IP_3-gated channels is composed of four identical subunits, and each subunit contains an IP_3-binding site in its N-terminal cytosol-facing domain. IP_3 binding induces opening of the channel, allowing Ca^{2+} to flow down its concentration gradient from the ER lumen into the cytosol. This family of IP_3-gated Ca^{2+} channel proteins is similar in structure to the family of voltage-sensitive Ca^{2+} channels that are called *ryanodine receptors* and that reside in a specialized ER in muscle cells called the *sarcoplasmic reticulum* (see Chapter 17).

The rise in the cytosolic Ca^{2+} level is transient because Ca^{2+} pumps located in the plasma membrane and in the ER membrane actively transport Ca^{2+} from the cytosol to the cell exterior or back into the ER lumen, respectively. Furthermore, within a second of its generation, the phosphate linked to carbon 5 of IP_3 is hydrolyzed, yielding inositol 1,4-bisphosphate (IP_2) (see Figure 15-27). IP_2 cannot open IP_3-gated Ca^{2+} channels.

Once released, the Ca^{2+} ions trigger further steps leading to the cellular response. As detailed in Chapter 3, the ubiquitous small cytosolic protein calmodulin functions as a multipurpose switch protein that mediates many cellular effects of Ca^{2+} ions. Binding of Ca^{2+} to four sites on calmodulin triggers a major conformational change that allows calmodulin to bind to and modulate the activity of many enzymes and other proteins (see Figure 3-34). Because four Ca^{2+} ions bind to calmodulin in a cooperative fashion, a small change in the level of cytosolic Ca^{2+} leads to a large change in the level of active calmodulin. One well-studied enzyme activated by the Ca^{2+}-calmodulin complex is myosin light-chain kinase, which regulates the activity of myosin and thus contraction in smooth muscle cells (see Chapter 17). Another is cAMP phosphodiesterase (PDE), the enzyme that degrades cAMP to 5′-AMP and terminates its effects (see Figure 15-18). This reaction thus links Ca^{2+} and cAMP, one of many examples in which two second messenger–mediated pathways interact to fine-tune a cellular response.

IP_3-triggered Ca^{2+} Transport from the ER to the Mitochondrial Matrix

In some cells, a cellular response to IP_3 takes place in the mitochondrial matrix in addition to the cytosol. In Chapter 12, we learned that there are direct contacts between specialized regions of the ER membrane, called *mitochondria-associated membranes (MAMs)*, and the outer mitochondrial membrane. These contacts affect mitochondrial structure, dynamics, and function. In response to elevated IP_3, Ca^{2+} is able to move from the ER lumen through the MAMs into the mitochondrial matrix (Figure 15-29). Once in the matrix, the elevated Ca^{2+} levels increase ATP synthesis as well as enhance mitochondrial activity in other ways.

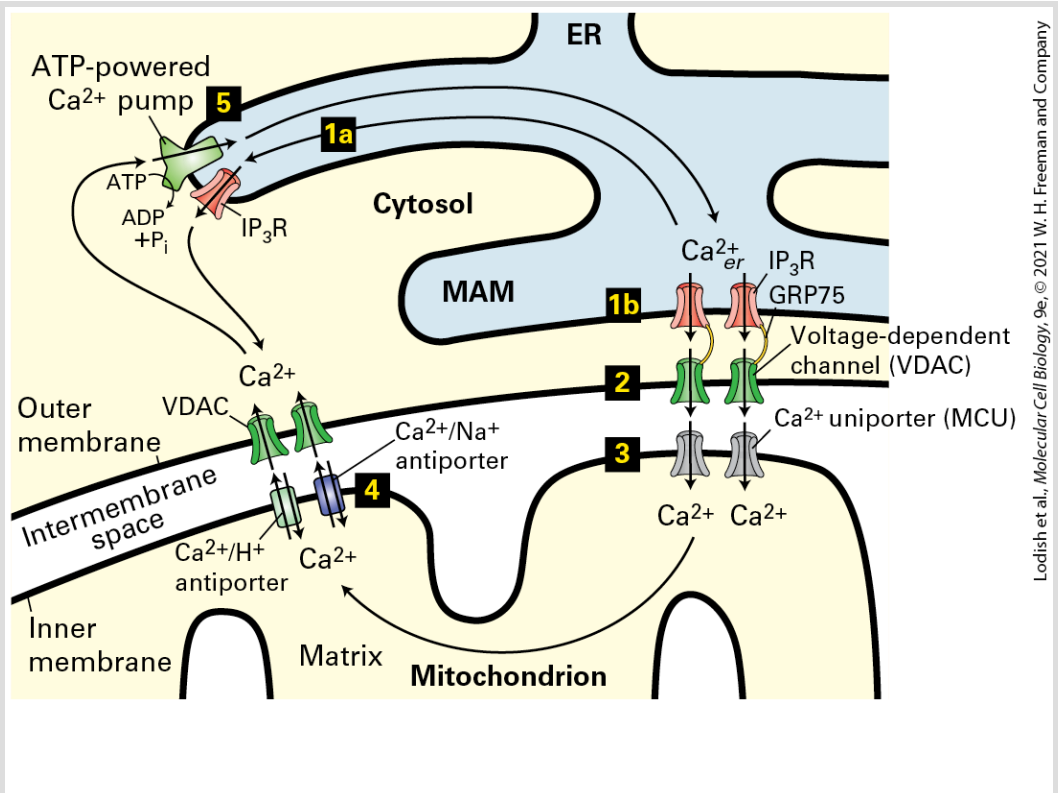


FIGURE 15-29 Movement of Ca^{2+} between the cytosol, mitochondrion, and endoplasmic reticulum. The ER is the main intracellular storage depot for Ca^{2+}. Binding of IP_3 to IP_3-gated Ca^{2+} channels (IP_3R) in the membrane of the endoplasmic reticulum releases Ca^{2+} into the cytosol (step **1a**); the IP_3-gated Ca^{2+} channels in the mitochondria-associated membranes (MAMs) of the ER are also opened by binding of IP_3 (step **1b**). Step **2**: VDAC channels in the outer mitochondrial membrane segments adjacent to MAMs are physically linked to IP_3Rs by the GRP75 protein; they directly and efficiently pass the Ca^{2+} released from the MAMs into the mitochondrial intermembrane space. Step **3**: The high concentration of Ca^{2+} in the intermembrane space induces the opening of MCU Ca^{2+} channels and other Ca^{2+} channels in the inner membrane, resulting in the flow of Ca^{2+} into the mitochondrial matrix. Step **4**: Over time, Ca^{2+} is released from the mitochondrial matrix into the intermembrane space by Ca^{2+}/Na^+ and Ca^{2+}/H^+ antiporters in the inner membrane, then transferred into the cytosol through VDAC or other Ca^{2+} channels in the outer membrane. Finally, pumping of Ca^{2+} from the cytosol by ATP-powered Ca^{2+} pumps in the ER membrane (step **5**) or plasma membrane restores the high ER Ca^{2+} and low cytosolic Ca^{2+} levels. See M. Schäfer et al., 2014, *Cell Tissue Res.* **357**:395; and K. Kamer and V. Mootha, 2015, *Nat. Rev. Mol. Cell Biol.* **16**:545.

The Ca^{2+} has to cross three separate membranes to reach the mitochondrial matrix. As detailed above, IP_3-gated Ca^{2+} channels in the ER membrane (step **1a**), including those in the MAMs (step **1b**), open in response to a rise in cytosolic IP_3. Voltage-dependent anion channels (VDACs) in the outer mitochondrial membrane adjacent to the MAMs, and which are physically linked by the GRP75 protein to these IP_3-gated Ca^{2+} channels, efficiently pass the Ca^{2+} released from the ER lumen into the intermembrane space (step **2**). A mitochondrial calcium uniporter (MCU) in the inner

mitochondrial membrane then transports Ca^{2+} from the intermembrane space into the mitochondrial matrix (step **3**).

MCUs open only in the presence of a high Ca^{2+} concentration in the intermembrane space. Regulatory subunits of the MCU that face the intermembrane space have Ca^{2+}-binding EF hands (see Chapter 3) with relatively low binding affinity for Ca^{2+}; these subunits must bind calcium for the MCU to open. Individuals with mutations in the gene encoding one of these subunits have skeletal muscle defects and learning disabilities, symptoms that can accompany other mitochondrial disorders. The consequences of these mutations attest to the importance of these uniporters in mitochondrial metabolism.

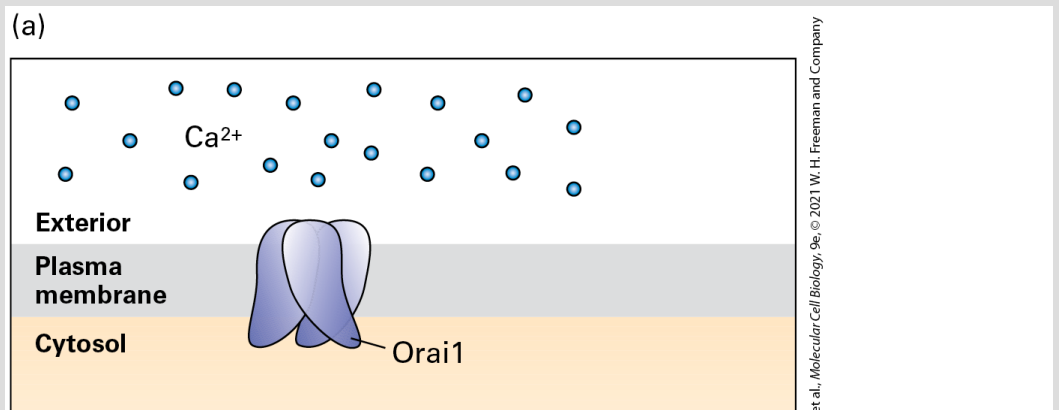
Too much calcium in the mitochondrial matrix is potentially toxic. To avoid the buildup of excess calcium, the mitochondrial matrix gradually releases Ca^{2+} into the cytosol. The calcium first moves across the inner mitochondrial membrane into the intermembrane space via Na^+/Ca^{2+} and H^+/Ca^{2+} antiporters, then crosses the outer mitochondrial membrane into the cytosol, most likely via VDACs (step **4**). The calcium transport cycle is completed when cytosolic calcium enters the ER via an ATP-powered Ca^{2+} pump (step **5**; see Figure 11-10) or is pumped out of the cell by ATP-powered Ca^{2+} pumps in the plasma membrane.

The Store-Operated Ca^{2+} Channel in the Plasma Membrane

If the IP_3-gated Ca^{2+} channel in the ER membrane stays open long enough, coupled with operation of the plasma-membrane Ca^{2+} export pump, the outflow of Ca^{2+} would eventually deplete the stores of Ca^{2+} in both the ER and mitochondrial matrix, and a cell would soon be unable to increase the cytosolic Ca^{2+} level in response to hormone-induced generation of IP_3. Using patch-clamping studies (see Figure 11-22), investigators established that cells must have a Ca^{2+} channel in the plasma membrane, called the *store-operated channel (SOC)*, that opens and admits extracellular Ca^{2+} into the cytosol when Ca^{2+} stores in the ER are depleted. The SOC thereby provides Ca^{2+} to replenish the ER and mitochondrial stores. To function, a SOC would require some additional ER-associated protein that could sense Ca^{2+} depletion in the ER and signal the SOC to open.

Initial attempts to identify a SOC were not successful because investigators focused on potential Ca^{2+} channel proteins with typical channel sequences and structures. Eventually investigators established the identity of this channel protein as *Orai1*; the sequence and three-dimensional structure of Orai1 is unlike that of any other known ion-channel protein.

The protein that senses Ca^{2+} depletion in the ER is called *STIM1*. It is a transmembrane protein in the ER membrane (Figure 15-30a) that binds Ca^{2+} when its concentration in the lumen of the ER is high. To bind Ca^{2+}, STIM1 uses an EF hand motif, similar to that in calmodulin (see Figure 3-34), located on the luminal side of the ER membrane. STIM1 is also bound to several cytosolic proteins and, via the EB1 protein, to cytosolic microtubules, all of which keep STIM1 away from the plasma membrane. As stores of Ca^{2+} in the ER are depleted, the STIM proteins lose their bound Ca^{2+}, oligomerize, and dissociate from EB1 and other cytosolic proteins. No longer anchored by these proteins, STIM proteins diffuse in the ER membrane to segments closely juxtaposed to the plasma membrane (Figure 15-30b). There portions of the STIM, called *CAD domains*, bind to and trigger the opening of Orai1, allowing the influx of extracellular Ca^{2+}; these two proteins are the key components of the store-operated Ca^{2+} pathway. Once the level of cytosolic Ca^{2+} rises to normal, a complex of Ca^{2+} bound to the protein calmodulin (Chapter 3) binds to Orai1, preventing STIM1 from binding to the ion channel and shutting down Ca^{2+} entry.



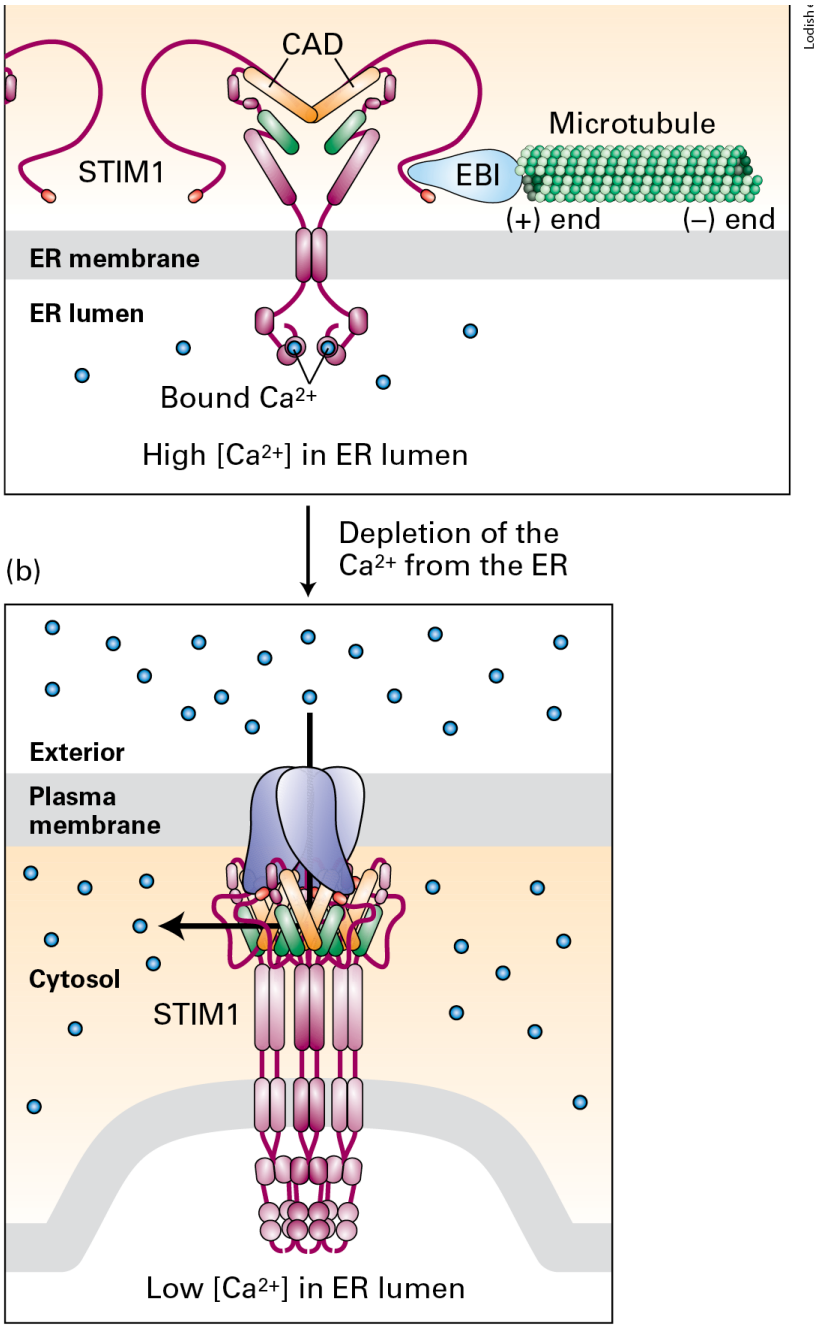


FIGURE 15-30 Opening of store-operated Ca^{2+} channels in the plasma membrane. (a) In the resting cell, Ca^{2+} levels in the ER lumen are high. Ca^{2+} ions (blue circles) bind to the luminal EF hand domains of the transmembrane STIM1 proteins, which are distributed throughout the ER. STIM1 binds to EB1, a protein that also binds to the $(+)$ ends of cytosolic microtubules. Binding to EB1 traps STIM1 molecules from translocating to sites where the ER and plasma membrane are juxtaposed. (b) As Ca^{2+} stores in the ER are depleted, Ca^{2+}

ions dissociate from the EF hands. STIMs undergo oligomerization, weakening their binding to EB1, and relocate to areas of the ER membrane near the plasma membrane. There the STIM CAD domains (orange) bind to and trigger the opening of the store-operated Ca^{2+} channels (Orai1) in the plasma membrane, allowing influx (black arrow) of extracellular Ca^{2+} into the cytosol.

Feedback Loops in ER and Cytosol Cycling of Ca^{2+} Trigger Oscillations in the Cytosolic Ca^{2+} Concentration

In the case of certain G protein–coupled receptors, their continuous activation induces rapid, repeated spikes in the level of cytosolic Ca^{2+} (Figure 15-31) rather than a prolonged elevation in cytosolic Ca^{2+}. These oscillations are caused by complex feedback interactions between the cytosolic Ca^{2+} concentration and the IP_3-gated Ca^{2+} channels in the ER membrane. In unstimulated, resting cells, activation of a GPCR triggers a rise in IP_3, and the rise in IP_3 in turn leads to a rapid rise in cytosolic Ca^{2+}. However, the high cytosolic Ca^{2+} level reached at the peak of the spike inhibits the further release of Ca^{2+} from ER stores by binding to the IP_3-gated Ca^{2+} channels and decreasing their affinity for IP_3. As a result, these Ca^{2+} channels close, and the cytosolic Ca^{2+} level drops rapidly as Ca^{2+} is pumped back into the ER lumen or out of the cell. Thus cytosolic Ca^{2+} is a feedback inhibitor of the IP_3-gated Ca^{2+} channels that, when open, trigger a rise in cytosolic Ca^{2+}.

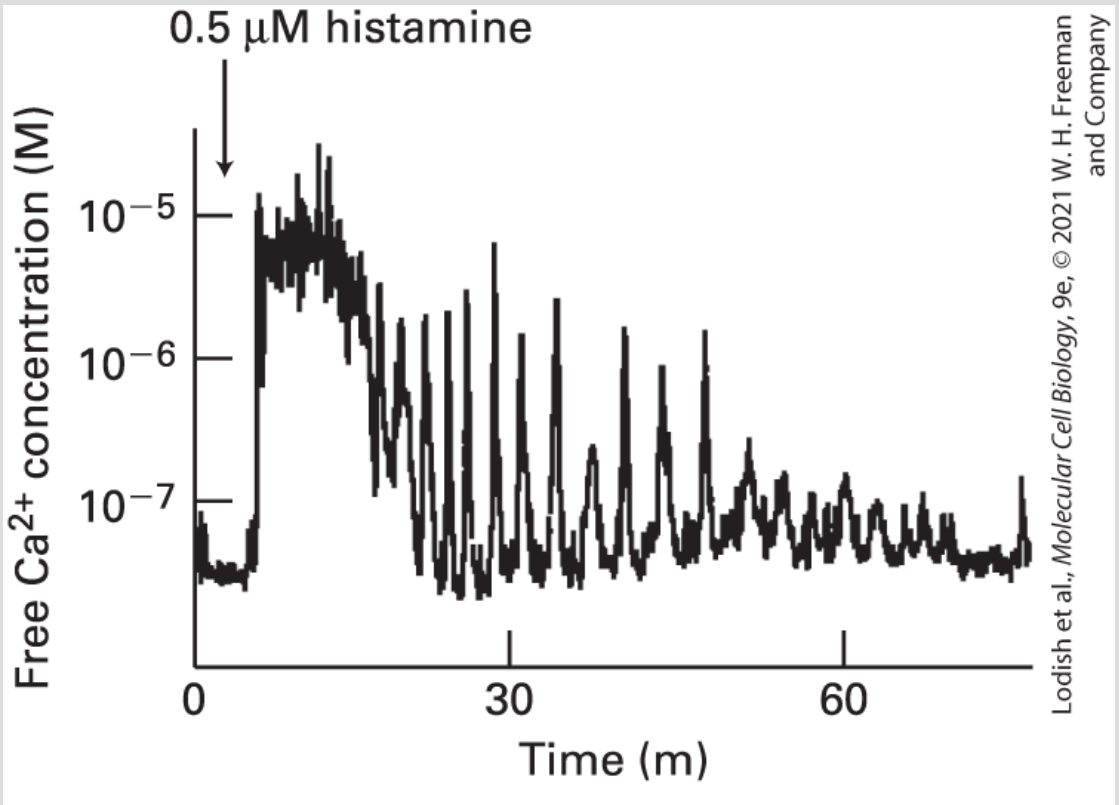


FIGURE 15-31 Oscillations in the cytosolic Ca^{2+} concentration following treatment of human HeLa cells with histamine. Like the LH receptors described in the text, the histamine GPCR activates the IP_3-DAG signaling pathway. The feedback loops generating the spikes in the cytosolic Ca^{2+} concentration are detailed in the text.

[Data from A. Miyawaki et al., 1997, *Nature* **388**:882.]

As an example, this mechanism produces calcium ion oscillations in the pituitary gland cells that secrete luteinizing hormone (LH), which plays an important role in controlling ovulation and thus female fertility. LH secretion is induced by the binding of luteinizing hormone–releasing hormone (LHRH) to its G protein–coupled receptors on the surfaces of these cells; LHRH binding induces repeated Ca^{2+} spikes. Each Ca^{2+} spike induces exocytosis of a few

LH-containing secretory vesicles, presumably those close to the plasma membrane, increasing the level of LH in the circulation.

DAG Activates Protein Kinase C

Recall that the hydrolysis of $PI(4,5)P_2$ by phospholipase C produces two products, IP_3 and DAG. DAG is also a second messenger, and its principal function is to activate a family of protein kinases collectively termed *protein kinase C (PKC)*.

After its formation, the hydrophobic DAG (see Figure 15-28) remains associated with the inner leaflet of the plasma membrane. In the absence of hormone stimulation, its target protein kinase C is present in the cytosol as a soluble protein that is catalytically inactive. A rise in the cytosolic Ca^{2+} level causes protein kinase C to translocate to the cytosolic surface of the plasma membrane, where it can bind to and become activated by phosphatidylserine and DAG (see Figure 15-28, steps **5** and **6**). Activation of PKC thus depends on an increase of both Ca^{2+} ions and DAG, suggesting an interaction between the two branches of the IP_3/DAG pathway.

The activation of PKC in different cells results in a varied array of cellular responses, indicating that it plays a key role in many aspects of cellular growth and metabolism. In many cells, PKC phosphorylates transcription factors that are localized in the

cytosol, triggering their movement into the nucleus, where they activate genes necessary for cell division. In liver cells, PKC helps regulate glycogen metabolism by phosphorylating and so inhibiting glycogen synthase.

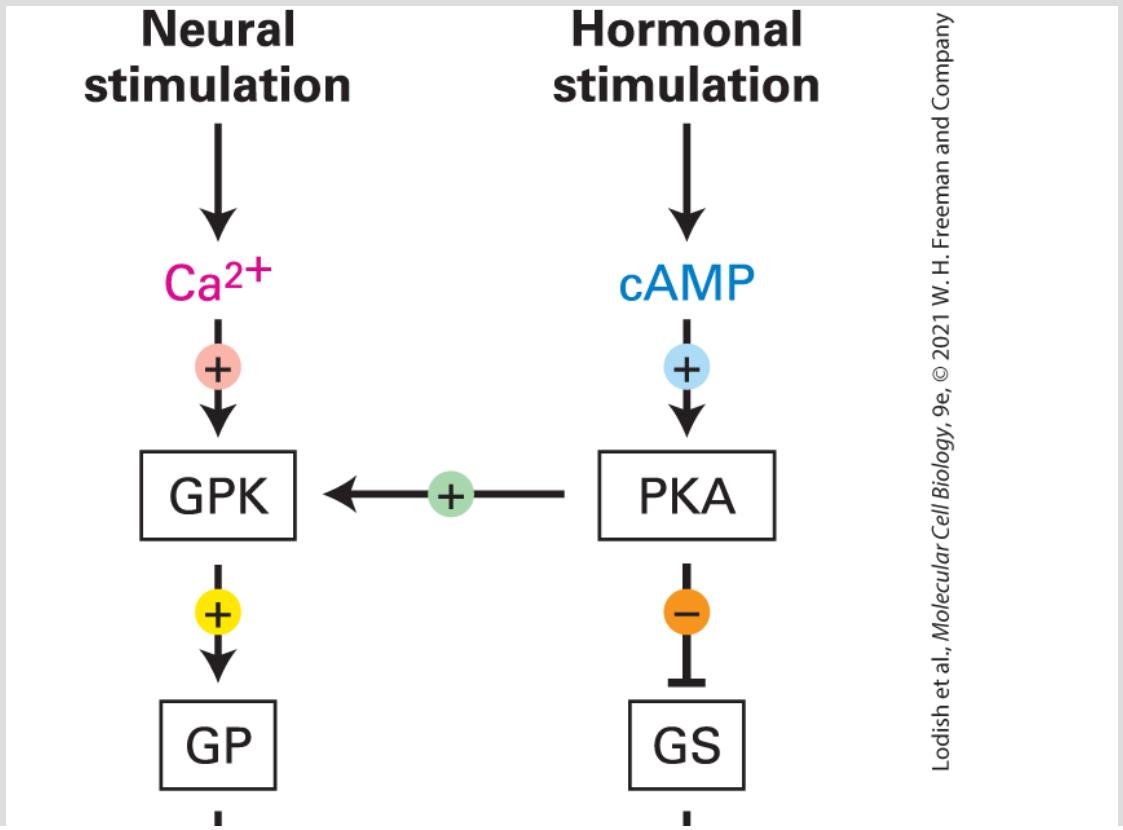
Integration of Ca^{2+} and cAMP Second Messengers Regulates Glycogenolysis

All cells constantly receive multiple signals from their environment, including changes in levels of hormones and metabolites, and all of these signals must be integrated. The breakdown of glycogen to glucose (glycogenolysis) provides an excellent example of how cells can integrate their responses to more than one signal. As discussed in Section 15.4, stimulation of muscle and liver cells by epinephrine leads to a rise in the second messenger cAMP, which promotes glycogen breakdown (see Figure 15-22a). In both muscle and liver cells, other signaling pathways produce the same cellular response of enhanced glycogenolysis.

In striated muscle cells (see Figure 17-29), stimulation by nerve impulses causes the release of Ca^{2+} ions from the sarcoplasmic reticulum and an increase in the cytosolic Ca^{2+} concentration, which triggers muscle contraction. With the rise in cytosolic Ca^{2+}, Ca^{2+} also binds to the calmodulin subunit of glycogen phosphorylase kinase (GPK), thus partially stimulating its kinase

catalytic activity. GPK then phosphorylates and thus activates its target glycogen phosphorylase. That enzyme stimulates the degradation of glycogen to glucose-1-phosphate, which provides fuel for prolonged muscle contraction.

As blood epinephrine concentrations rise, adenylyl cyclase is activated and cAMP concentrations increase, activating PKA. Recall that phosphorylation by PKA also activates GPK (see Figure 15-22). We see that maximal activation of GPK, and therefore a maximal rate of glycogenolysis, requires both phosphorylation and Ca^{2+}. Thus, in muscle, a key regulator of glycogenolysis, *GPK*, is subject to regulation by both nerve signals and hormones and integrates both signal transduction pathways (Figure 15-32).



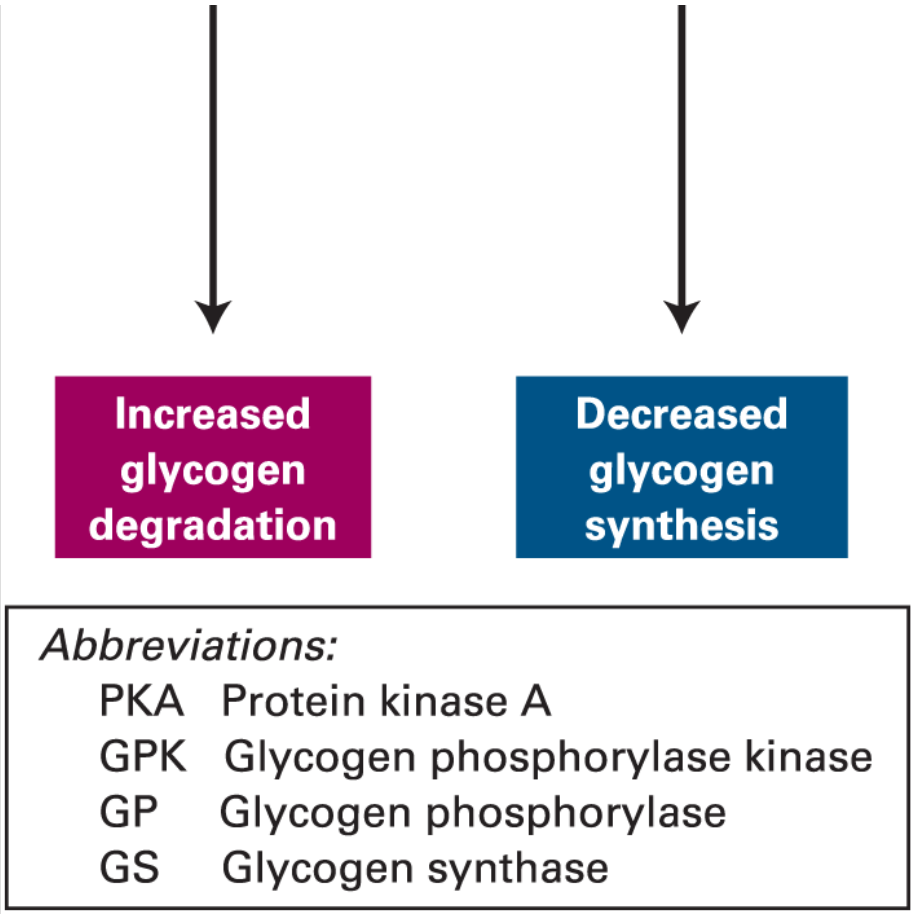


FIGURE 15-32 Integrated regulation of glycogenolysis by Ca^{2+} and cAMP/PKA pathways. Neuronal stimulation of striated muscle cells or the binding of epinephrine to β-adrenergic receptors on their surfaces leads to an increase in the cytosolic concentrations of the second messengers Ca^{2+} and cAMP, respectively. The key regulatory enzyme glycogen phosphorylase kinase (GPK) is activated both by binding Ca^{2+} ions and by phosphorylation by cAMP-dependent PKA. Enzymes are marked by white boxes.
$(+)=$ activation of enzyme activity; $(-)=$ inhibition.

KEY CONCEPTS OF SECTION 15.5

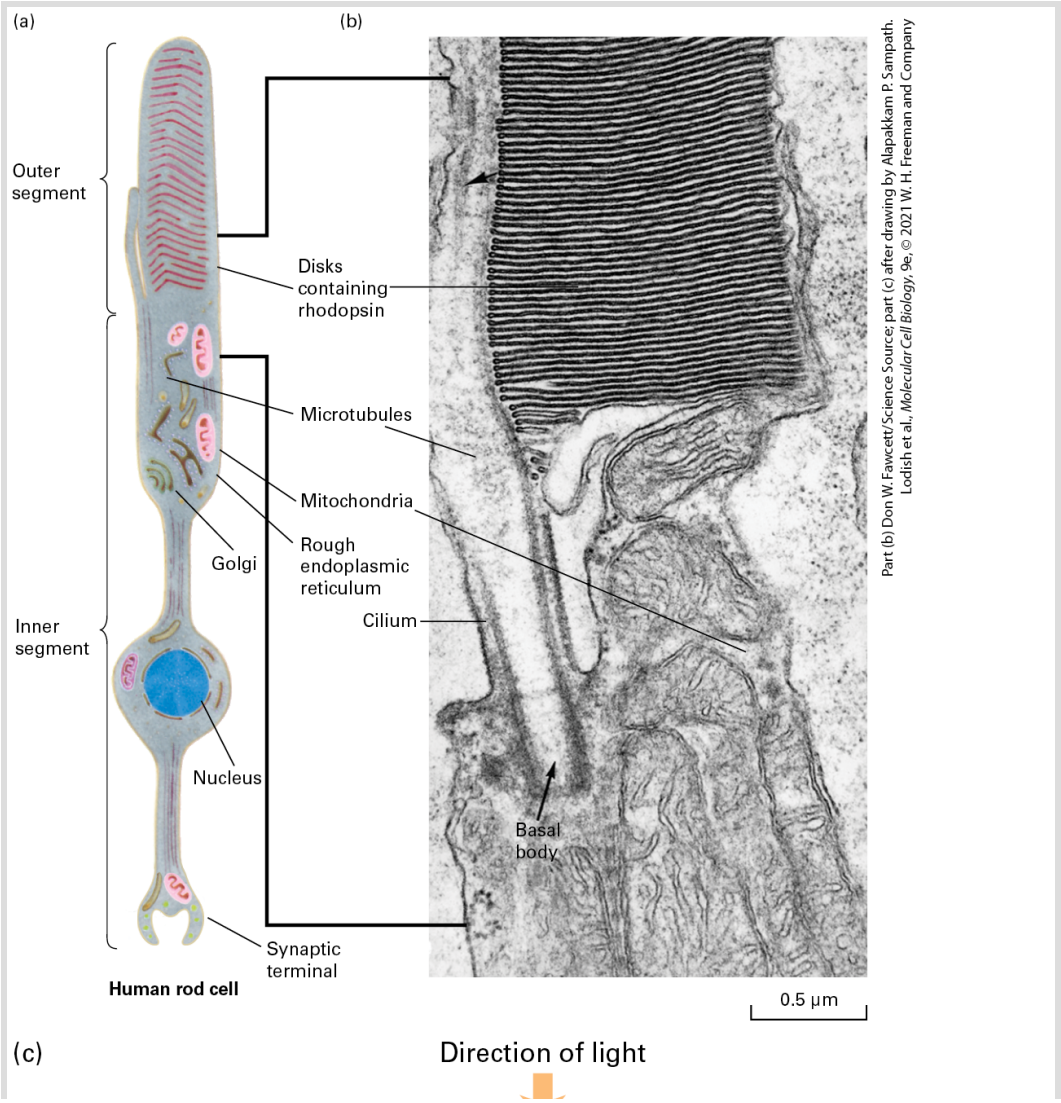
Regulating Protein Secretion and Muscle Contraction: Ca^{2+} Ions as Second Messengers in Multiple Signal Transduction

Pathways

- A small rise in cytosolic Ca^{2+} induces a variety of responses in different cells, including hormone secretion, contraction of muscle, and platelet aggregation.
- Many hormones bind GPCRs coupled to G proteins containing a $G_{\alpha q}$ subunit. The effector protein activated by GTP-bound $G_{\alpha q}$ is a phospholipase C enzyme.
- Phospholipase C cleaves a phospholipid in the cytosolic leaflet of the plasma membrane known as $PI(4,5)P_2$, generating two second messengers: water-soluble IP_3 and membrane-bound DAG.
- IP_3 triggers the opening of IP_3-gated Ca^{2+} channels in the endoplasmic reticulum and elevation of cytosolic free Ca^{2+}.
- Opening of IP_3-gated Ca^{2+} channels in the endoplasmic reticulum also leads to an increase in Ca^{2+} in the mitochondrial matrix and an acceleration of ATP synthesis.
- Depletion of ER Ca^{2+} stores leads to the opening of plasma-membrane store-operated Ca^{2+} channels and an influx of Ca^{2+} from the extracellular medium.
- In response to elevated cytosolic Ca^{2+}, protein kinase C is recruited to the plasma membrane, where it is activated by DAG.
- Glycogen breakdown and synthesis in muscle is coordinately regulated by the second messengers Ca^{2+} and cAMP, whose levels are regulated by neural and hormonal stimulation, respectively.

15.6 Vision: How the Eye Senses Light

Animals have a remarkable ability to perceive the world thanks to their ability to sense and respond to physical stimuli. One of the most important of these stimuli is light. To understand light-initiated signal transduction, we will focus on one of the two types of photoreceptor cells in the human retina, the *rods* (Figure 15-33).



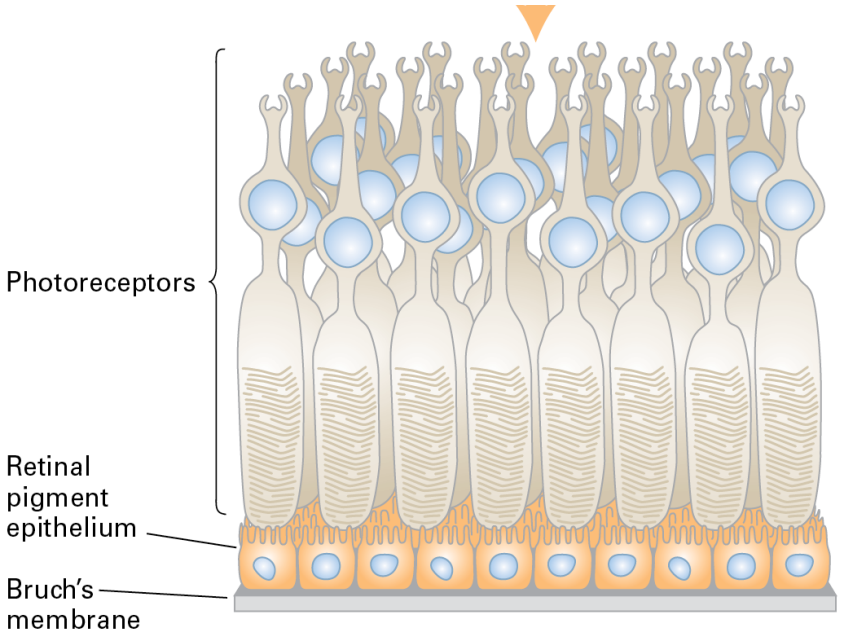


FIGURE 15-33 Human rod cell. (a) Schematic diagram of an entire rod cell. At the synaptic terminal (also called the *spherule*), the rod cell forms synapses with one or more interneurons. Rhodopsin, a light-sensitive G protein–coupled receptor, is located in the flattened membrane disks of the cell's outer segment. (b) Electron micrograph of the region of the rod cell indicated by the bracket in (a). This region includes the junction of the inner and outer segments. (c) Structure of the human retina, illustrating the layer of pigment epithelial cells that underlie the rods and cones and in which all-*trans*-retinol is converted to 11-*cis*-retinal before being shuttled back to the rod and cone cells.

[Part (b) Don W. Fawcett/Science Source. Part (c) after drawing by Alapakkam P. Sampath.]

Light Activates Rhodopsin in Rod Cells of the Eye

Rod cells are stimulated by weak light such as moonlight and are almost entirely responsible for night vision; they have little role in

color vision. The other type of human photoreceptor cells, *cones*, are involved in color vision and use a similar signal transduction pathway. These photoreceptor cells signal to (synapse with) layer upon layer of interneurons at the back of the eye that are innervated by different combinations of photoreceptor cells. All these signals are processed and relayed through the visual thalamus to the part of the brain called the *visual cortex*, where they are interpreted.

Rod cells sense light with the aid of a light-sensitive GPCR called *rhodopsin*. Rhodopsin has two components: the protein opsin, a typical member of the GPCR A family with seven transmembrane α-helical segments, and a light-absorbing pigment called 11-*cis*-retinal that is covalently bound to opsin (Figure 15-34a). The bound 11-*cis*-retinal locks the rhodopsin into its inactive (R) conformation, preventing any signaling in the absence of light. Rhodopsin, found only in rod cells (cone cells have homologous photoreceptor molecules), is localized to the thousand or so flattened membrane disks that make up the outer segment of each of these rod-shaped cells (see Figure 15-33). There are about 90 million rod cells in the human retina, and each contains about 4×10^7 molecules of rhodopsin that are very densely packed in the membrane disks. The heterotrimeric G protein coupled to rhodopsin, called *transducin* (G_t), contains a G_α subunit referred to as $G_{\alpha t}$; like rhodopsin, $G_{\alpha t}$ is found only in rod cells.

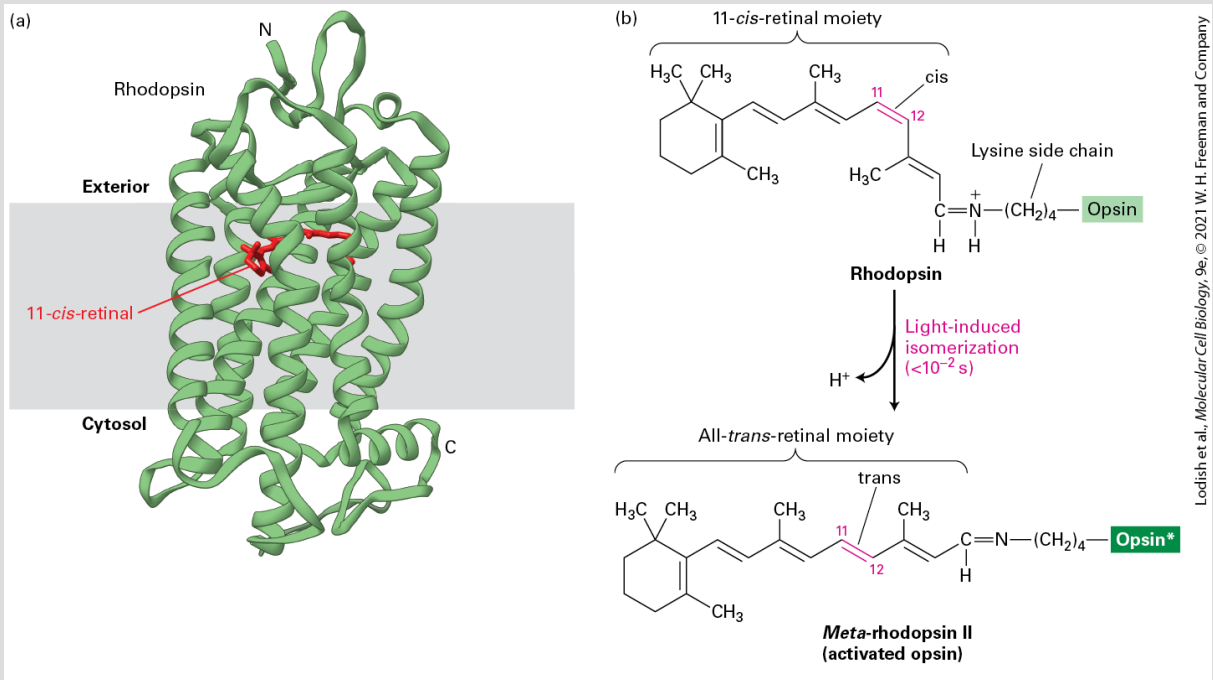


FIGURE 15-34 Vision depends on the light-triggered isomerization of the retinal moiety of rhodopsin. (a) Rhodopsin consists of the light-absorbing pigment 11-*cis*-retinal covalently attached to the amino group of lysine residue 296 in the opsin protein. (b) Absorption of light causes rapid photoisomerization of the bound *cis*-retinal to the all-*trans* isomer, triggering a conformational change in opsin that enables it to activate the G protein transducin.

Rhodopsin differs from other GPCRs in that binding of a ligand does not activate the receptor. Rather, the activating signal is the absorption of a photon of light by the bound retinal (Figure 15-34b). On absorption of a photon, the retinal moiety of rhodopsin is rapidly converted from the *cis* isomeric form (11-*cis*-retinal) to the all-*trans* isomeric form. This change of shape of the retinal induces a change in the protein from its inactive (R) to active (R*) conformation, equivalent to the activating conformational change that occurs upon ligand binding to other GPCRs. As a consequence, the R* form of

rhodopsin binds to transducin, triggering exchange of GTP for GDP on its $G_{\alpha t}$ subunit. At the same time, transducin also dissociates into its $G_{\alpha t}$·GTP and $G_{\beta\gamma}$ subunits. A single activated rhodopsin protein activates ~13 transducin molecules in the few seconds before the covalent bond between the opsin protein and all-*trans*-retinal is spontaneously cleaved; once the all-*trans*-retinal has dissociated opsin can no longer activate transducin.

The free all-*trans*-retinal is converted back to 11-*cis*-retinal in multiple steps. First, all-*trans*-retinal is converted to an intermediate, all-*trans*-retin*ol*, by an enzyme in the outer segment of the rod. This intermediate is then shuttled by a series of transport proteins to the adjacent *pigment epithelial cells* in the retina (Figure 15-33c), where it is converted to 11-*cis*-retinal. The 11-*cis*-retinal is then shuttled back to the rod cells where it binds to opsin, forming rhodopsin and completing the rhodopsin visual cycle.

RPE65 is the enzyme in the pigment epithelial cells of the retina that catalyzes the isomerization of the intermediate esterified all-*trans*-retinol to 11-*cis*-retinol. Autosomal recessive mutations in the RPE65 gene are a frequent cause of blindness in children. In 2017, the U.S. Food and Drug Administration approved the first in vivo human gene therapy in the United States, one to treat blindness based on a defective RPE65 enzyme. In this treatment, an adeno-associated virus vector injected directly into the retina is used to introduce a complete copy of the *RPE65* gene into the pigment

epithelial cells. All young recipients reported an increase in visual acuity and improved ability to negotiate mazes. ■

Activation of Rhodopsin by Light Leads to Closing of cGMP-Gated Cation Channels

In the dark, the potential across the plasma membrane of a rod cell is about $-30\ \mathrm{mV}$. That is considerably more positive (less negative) than the resting potential (-60 to $-90\ \mathrm{mV}$) typical of neurons and other electrically active cells (see Chapter 11). This state of the membrane, called the **depolarized state**, causes rod cells in the dark to release neurotransmitters constantly. Thus the neurons with which they synapse are constantly being stimulated; the signal they send to the brain through a series of neurons is perceived as darkness. The plasma membrane of a resting rod cell remains in this depolarized state because a large number of open ion channels admit both Na^+ and Ca^{2+} into the cell (hence they are called *nonselective*). Recall from Chapter 11 that movement of cations (positively charged ions) such as Na^+ and Ca^{2+} from the outside of the cell to the inside will reduce the magnitude of the inside-negative membrane potential, or depolarize the membrane.

These nonselective cation channels open in response to (are "gated" by) binding of the second messenger **cyclic guanosine**

monophosphate, or **cGMP** (see Figure 15-5). The outer segments of resting rod cells contain an unusually high concentration ($\sim$0.07 mM) of cGMP, which is generated from GTP in a reaction catalyzed by guanylyl cyclase. Thus in the dark the high cGMP concentration holds many of the cGMP-gated channels in the open state, depolarizing the membrane.

Light acts by lowering the concentration of cGMP and thereby closing these nonselective cation channels (Figure 15-35). Absorption of light by rhodopsin leads to receptor and then G_t activation. Almost instantaneously, $G_{\alpha t}$·GTP triggers the activation of *cGMP phosphodiesterase (PDE)*, which hydrolyzes cGMP to 5′-GMP. As the cGMP concentration drops, the nonselective cation channels close and the membrane potential becomes *more* negative inside (see Figure 15-35, step **6**). This repolarization of the plasma membrane, in turn, causes a reduction in neurotransmitter release. The more photons are absorbed by rhodopsin, the more cGMP is hydrolyzed, and the more channels are closed. The fewer Na^+ and Ca^{2+} ions cross the membrane from the outside, the more inside negative the membrane potential becomes, and so the less neurotransmitter is released. The reduction in neurotransmitter release is transmitted to the brain by a series of neurons, where it is perceived as light.

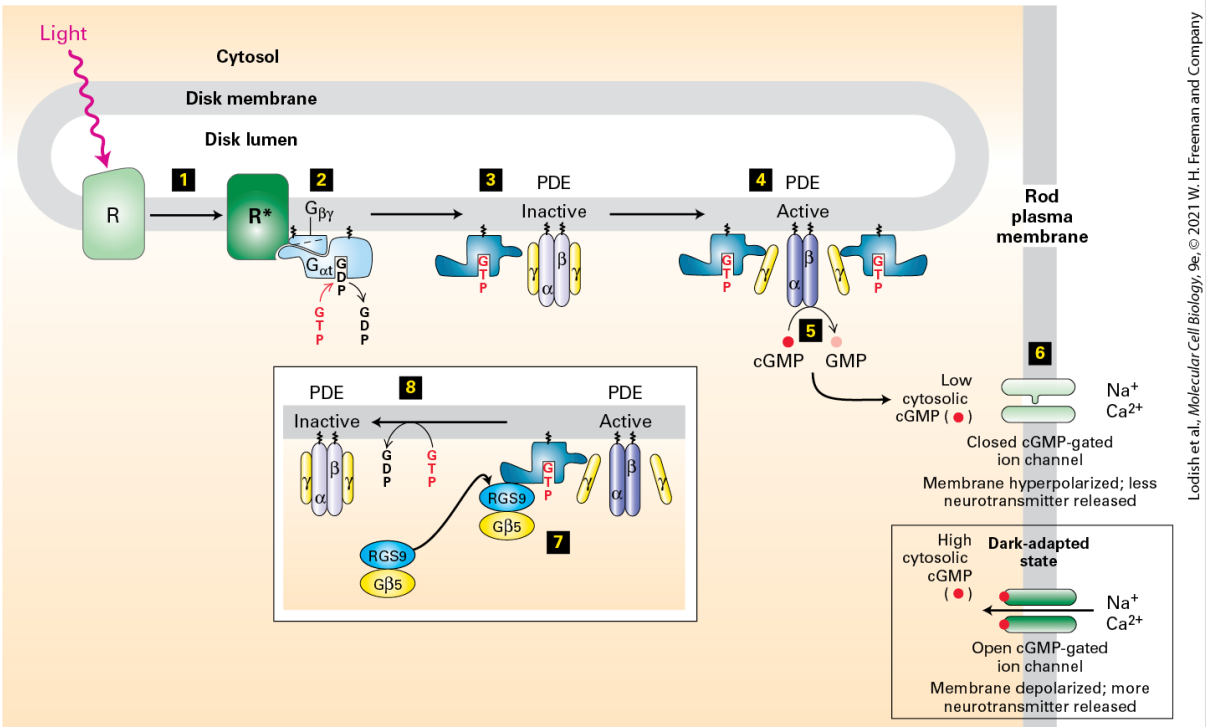


FIGURE 15-35 The light-activated rhodopsin pathway and the closing of cation channels in rod cells. In dark-adapted rod cells, a high level of cGMP keeps cGMP-gated plasma membrane nonselective cation channels open, leading to influx of Na^+ and Ca^{2+}, depolarization of the plasma membrane, and neurotransmitter release. Light absorption generates activated rhodopsin, R^* (step 1), which binds inactive, GDP-bound $G_{\alpha t}$ protein and mediates the exchange of GDP for GTP (step 2). The free $G_{\alpha t}{\cdot}GTP$ generated then activates PDE by binding to its inhibitory γ subunits (step 3) thus dissociating them from the catalytic α and β subunits (step 4). Relieved of their inhibition, the α and β subunits of PDE hydrolyze cGMP to GMP (step 5). The resulting decrease in cytosolic cGMP leads cGMP to dissociate from the cation channels in the plasma membrane and causes these channels to close (step 6). As a result the plasma membrane then becomes transiently hyperpolarized, and neurotransmitter release is reduced. The complex of $G_{\alpha t}{\cdot}GTP$ and the PDE γ subunits binds a GTPase-activating complex formed of the components RGS9 and Gβ5 (step 7); by hydrolyzing the bound GTP, this complex triggers the physiologically important rapid inactivation of the PDE (step 8). See V. Arshavsky and E. Pugh, 1998, *Neuron* **20**:11; and V. Arshavsky, 2002, *Trends Neurosci.* **25**:124.

As depicted in Figure 15-35, lipid anchors confine both $G_{\alpha t}$·GTP and its effector, PDE, to the cytosolic face of the disk membranes of the rod cell. The $G_{\alpha t}$·GTP complexes that are generated through activation of rhodopsin are able to move laterally along the membrane surface and bind to the two inhibitory γ subunits of PDE (see Figure 15-35); note that the stoichiometry is 1:1, that is, one $G_{\alpha t}$·GTPbinds to one γ subunit. The binding of $G_{\alpha t}$·GTP to the γ subunits releases the catalytically active αβ dimer, which then converts cGMP to GMP. This is another example of how removal of an inhibitory subunit can quickly activate an enzyme, a common mechanism in signaling pathways.

Direct support for the role of cGMP in rod-cell activity has been obtained in patch-clamping studies using isolated patches of plasma membrane taken from the rod outer segment, which contains abundant cGMP-gated cation channels. When cGMP is added to the cytosolic surface of these patches, the number of open cation channels increases rapidly; cGMP binds directly to a site on the cytosolic surface of the channel proteins to keep them open. Like the K^+ channels discussed in Chapter 11, the cGMP-gated channel protein contains four subunits. In this case, each of the subunits is able to bind a cGMP molecule. Three or four cGMP molecules must bind to each channel in order to open it; because of this cooperative allosteric interaction, channel opening is very sensitive to small changes in cGMP levels.

Signal Amplification Makes the Rhodopsin Signal Transduction Pathway Exquisitely Sensitive

Remarkably, a single photon absorbed by a resting rod cell produces a measurable response in the form of a small hyperpolarization in the membrane potential of about 1 mV, which in amphibians lasts a second or two. With their 90 million rod cells, dark-adapted human eyes are able to perceive as light the absorption of as few as five photons; the ability to respond to a very few photons is critical to night vision.

The light-detecting system is very sensitive because the signal is greatly amplified by the signal transduction pathway. During the time it is active, each rhodopsin molecule in the disk membrane of a rod cell can activate $\sim$13 $G_{\alpha t}$ molecules, two of which in turn activate a single phosphodiesterase molecule (see Figure 15-35). Each PDE molecule hydrolyzes hundreds of cGMP molecules during the fraction of a second it remains active. The absorption of a single photon — yielding a single activated rhodopsin molecule — can reduce the cGMP concentration enough to trigger the closing of thousands of cation channels (or about 5 percent of the open channels) in the plasma membrane, causing a measurable change in the membrane potential of the cell.

Rapid Termination of the Rhodopsin Signal Transduction Pathway Is Essential for the Temporal Resolution of Vision

In many organisms a fast response to rapidly changing light conditions (dark-to-light and light-to-dark) is critical for survival. As in all GPCR-mediated signaling pathways, after a light-to-dark transition all the activated intermediates must be inactivated rapidly to shut down the signaling pathway, restoring the system to its basal, dark state, ready for signaling by the reappearance of light. Thus the three protein intermediates generated during periods of light — activated rhodopsin (R*), $G_{\alpha t}$·GTP, and activated PDE — must all be inactivated when the light goes off. Once these intermediates are inactivated, the concentration of cytoplasmic cGMP is restored rapidly to its dark level as cGMP is resynthesized by the guanylyl cyclase enzyme.

Termination of the Signal from Light-Activated Rhodopsin (R*) by Rhodopsin Phosphorylation and Binding of Arrestin

As noted above, the spontaneous cleavage of the covalent bond between an opsin and its all-*trans*-retinal and the release of the all-*trans*-retinal terminates signaling from that activated rhodopsin (R*) molecule. Refilling the retinal binding site with 11-*cis*-retinal restores the protein to the inactive (R) state. This process, called *dark adaptation*, is relatively slow compared to potentially very rapid changes in external illumination. Thus there are additional mechanisms that suppress the activity of activated rhodopsin more rapidly.

One important process that helps to terminate the response to light quickly is the phosphorylation of rhodopsin when it is in its activated (R*) form, but not its inactivated, or dark (R), form. The enzyme that catalyzes this phosphorylation is *rhodopsin kinase*, a member of the GRK family of kinases (see Figure 15-26a). Each opsin molecule has multiple principal serine and threonine phosphorylation sites on its cytosol-facing C-terminal segment; the more sites that are phosphorylated by rhodopsin kinase, the less R* is able to bind to and activate G_t.

The binding of arrestin to rhodopsin dramatically speeds up the inactivation process. Arrestin binds tightly to rhodopsin only when rhodopsin kinase has phosphorylated a sufficient number of C-terminal serines and threonines. Arrestin binding to the phosphorylated rhodopsin completely prevents its interaction with G_t, thus blocking formation of the active $G_{\alpha t}$·GTP complex and

further activation of PDE. The entire process of rhodopsin phosphorylation and its inactivation by arrestin is fast — it is completed within 50 milliseconds in mammals. The phosphorylated rhodopsin is restored to its original, light-sensitive state (R) during dark adaptation. During the dark, a specific rhodopsin phosphatase slowly removes the phosphates and 11-*cis*-retinal binds to opsin.

Termination of the Signal from Activated $G_{\alpha t}$·GTP by GTP Hydrolysis

$G_{\alpha t}$ automatically shuts itself off after a brief period: it contains an intrinsic GTPase activity that hydrolyzes the bound GTP to GDP, producing an inactive $G_{\alpha t}$·GDP complex. That inactive $G_{\alpha t}$·GDP reassociates with a $G_{\beta\gamma}$ and can no longer bind to a PDE γ subunit. The bound PDE γ subunits are released (see Figure 15-35, step **7**), and then rejoin the PDE α and β subunits to regenerate an inactive PDE $\alpha\beta\gamma_2$ tetramer (step **8**). Thus PDE rapidly loses its activity, and the cGMP concentration begins to rise to its levels before light induction. Through this process, the eye can respond quickly to changing patterns of light created by moving objects or a change in the direction of gaze.

The intrinsic GTPase activity of $G_{\alpha t}$ in rod cells is accelerated by the action of a specific heterodimeric GTPase-activating protein (GAP) that is recruited to the complex of the γ subunit of PDE and

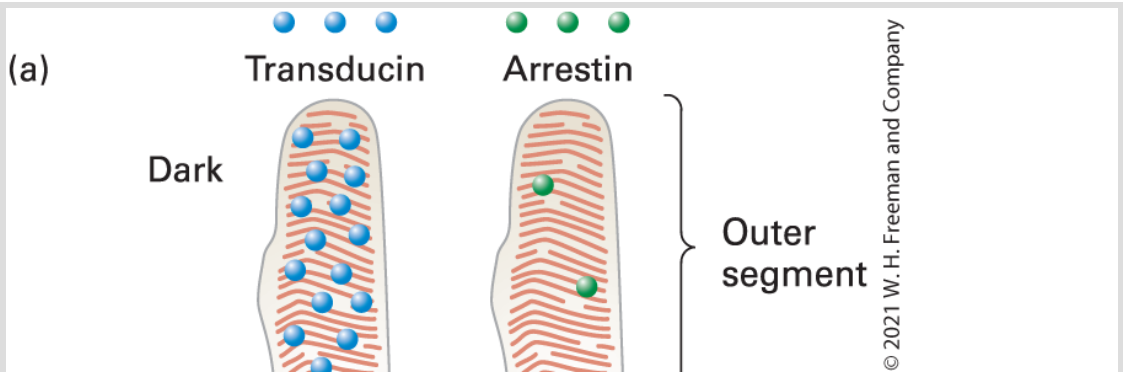
$G_{\alpha t}$·GTP. As a result, the $G_{\alpha t}$ normally remains in the active GTP-containing state for only a fraction of a second. The components of this GAP are RGS9 and Gβ5 (see Figure 15-35, step **7**). Experiments with mice in which the gene encoding RGS9 had been knocked out showed that this protein is essential for normal inactivation of the signaling cascade in vivo. In individual mouse rod cells, the time required for recovery from a single flash of light increased from 0.2 second in the normal mouse to about 9 seconds in the RGS9-deficient mouse. This represents a 45-fold increase, attesting to the importance of RGS9 as a component of the $G_{\alpha t}$·GTP-GAP complex.

Rod Cells Adapt to Varying Levels of Ambient Light by Intracellular Trafficking of Arrestin and Transducin

When we move from bright sunlight into a dimly lit room, we are temporarily insensitive to low levels of light and initially cannot easily see the objects around us. However, after some time spent in the low-light conditions, our eyes adjust and we are better able to see and distinguish objects; we become more sensitive to lower levels of light. On the other hand, when we move quickly from a dimly lit room into bright sunshine, we are initially "blinded" by the bright light (hypersensitivity), but with time we become less sensitive (desensitized) and can once again see clearly. As the result of this process, called *light adaptation*, rod cells are able to perceive

clearly objects over a 100,000-fold range of ambient light levels, all the way from a dimly lit room to bright sunlight. This wide range of sensitivities is possible because it is the *difference* in light levels in the visual field, rather than the absolute *amount* of light absorbed, that is ultimately used by the brain to form visual images. Much of the visual adaptation occurs in the photoreceptor cells, which are able to adjust the rhodopsin-initiated signal transduction pathway (see Figure 15-34) to respond to low or high ambient light intensities.

In dark-adapted rod cells, 80–90 percent of the $G_{\alpha t}$ and $G_{\beta\gamma}$ transducin subunits are located in the outer segments, while fewer than 10 percent of the molecules of the inhibitor arrestin are located there (Figure 15-36). This spatial distribution allows maximal activation of the downstream effector PDE and thus maximal sensitivity to small changes in light. Exposure for 10 minutes to moderate daytime intensities of light causes a redistribution of these proteins such that over 80 percent of the $G_{\alpha t}$ and $G_{\beta\gamma}$ subunits move out of the outer segment into other parts of the cell, while over 80 percent of arrestin moves into the outer segment.



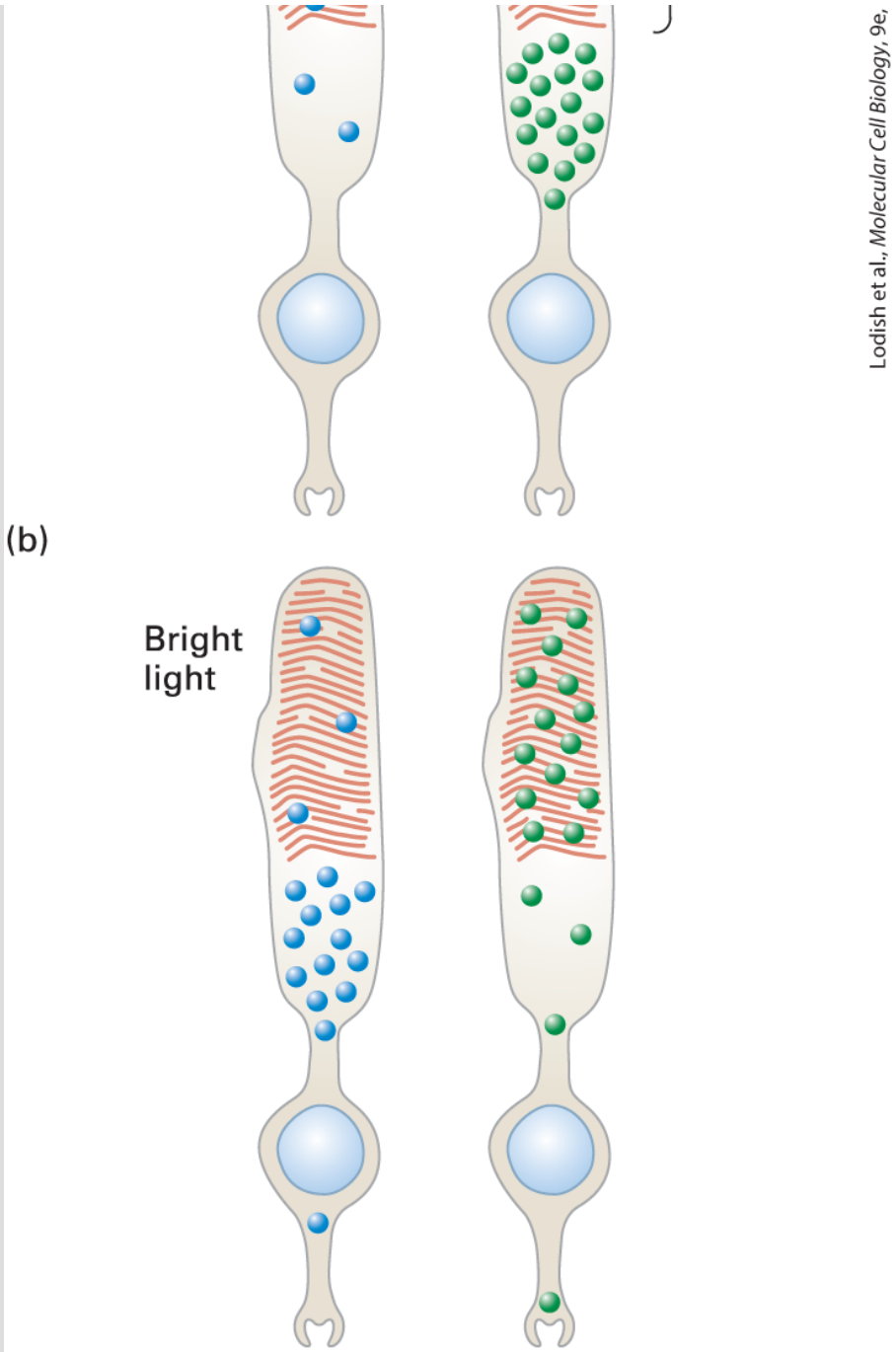


FIGURE 15-36 Schematic illustration of transducin and arrestin distribution in dark-adapted and light-adapted rod cells. (a) In the dark, most transducin (blue circles) is localized to the outer segment, while most arrestin (green circles) is found in other parts of the cell; in this condition, vision is most sensitive to very low light levels. (b) In bright light, little transducin is found in the outer segment, and abundant arrestin is found there; in this condition, vision is relatively insensitive to

small changes in light. The coordinated movement of these proteins contributes to our ability to perceive images over a 100,000-fold range of ambient light levels. See P. Calvert et al., 2006, *Trends Cell Biol.* **16**:560.

The mechanism by which these proteins move within the cell is not yet known, but most likely microtubule-attached motors carry the proteins into and out of the outer segment (see Chapter 18); note the microtubules in the diagram and micrograph in Figure 15-33. In bright light, the level of transducin, with its $G_{\alpha t}$ and $G_{\beta\gamma}$ subunits, is so reduced in the outer segment that G_t is simply not available for binding to and activation by rhodopsin. As a result, less PDE is activated. At the same time, the increase in the arrestin level in the outer segment means that any activated rhodopsin that has been phosphorylated by rhodopsin kinase will become rapidly inactivated. These protein movements are reversed when the ambient light level is lowered. Thus our visual system is desensitized in bright light by the drop in transducin level and the increase in arrestin level in the outer segment.

KEY CONCEPTS OF SECTION 15.6

Vision: How the Eye Senses Light

- Rhodopsin, the photosensitive GPCR in rod cells, comprises the protein opsin linked to 11-*cis*-retinal. The presence of light induces isomerization of the 11-*cis*-retinal moiety to all-*trans*-retinal. This changes the opsin conformation so that it activates the coupled heterotrimeric G protein transducin (G_t) by catalyzing exchange of free GTP for bound GDP on the $G_{\alpha t}$ subunit (see Figures 15-34 and 15-35).

- The effector protein in the rhodopsin pathway is PDE, which is activated by $G_{\alpha t} \cdot GTP$ binding to and releasing its inhibitory γ subunits. This enzyme reduces the cGMP level, and, as a result, cGMP-gated Na^{+}/Ca^{2+} channels close, the membrane becomes hyperpolarized, and the release of neurotransmitter decreases (see Figure 15-35).
- Several mechanisms act to terminate visual signaling: GAP proteins inactivate $G_{\alpha t} \cdot GTP$, and rhodopsin phosphorylation and binding of arrestin inhibit activation of transducin.
- Adaptation to a wide range of ambient light levels is mediated by movements of transducin and arrestin into and out of the rod-cell outer segment. Together, these two proteins modulate the ability of small increases in light levels to activate the downstream effector PDE, and thus the sensitivity of the rod cell to different ambient levels of light.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Classic Experiment 15-1: The Infancy of Signal Transduction Studies: GTP Stimulation of cAMP Synthesis
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

adenylyl cyclase
agonist
antagonist
β-adrenergic receptor
cell-surface receptors
cellular communication
cyclic guanosine monophosphate (cGMP)
depolarized state
dissociation constant
effector protein
epinephrine

G protein–coupled receptor (GPCR)
growth factor
GTPase superfamily of proteins
heterotrimeric G protein
hormone
ligand
monomeric G protein
protein kinase (PK)
protein phosphatase
receptor
second messenger
signal transduction pathway
target cell

Review the Concepts

1. What common features are shared by most cell-signaling systems?
2. Signaling by soluble extracellular molecules can be classified as endocrine, paracrine, or autocrine. Describe how these three types of cellular signaling differ. Growth hormone is secreted from the pituitary, which is located at the base of the brain and acts through growth hormone receptors located on the liver. Is this an example of endocrine, paracrine, or autocrine signaling? Why?
3. A ligand binds two different receptors with a K_d value of 10^{-7} M for receptor 1 and a K_d value of 10^{-9} M for receptor

2. For which receptor does the ligand show the greater affinity? Calculate the fraction of receptors that have a bound ligand ($[RL]/R_T$) in the case of receptor 1 and receptor 2 if the concentration of free ligand is 10^{-8} M.

4. To understand how a signaling pathway works, it is often useful to isolate the cell-surface receptor and to measure the activity of downstream effector proteins under different conditions. How could you use affinity chromatography to isolate a cell-surface receptor? With what technique could you measure the amount of activated G protein (the GTP-bound form) in ligand-stimulated cells? Describe the approach you would take.
5. How do seven–transmembrane domain G protein–coupled receptors transmit a signal across the plasma membrane? In your answer, include the conformational changes that occur in the receptor in response to ligand binding.
6. Signal-transducing heterotrimeric G proteins consist of three subunits designated α, β, and γ. The G_α subunit is a GTPase switch protein that cycles between active and inactive states depending on whether it is bound to GTP or to GDP. Review the steps for ligand-induced activation of effector proteins mediated by the heterotrimeric G proteins. Suppose that you have isolated a mutant G_α subunit that has an increased GTPase activity. What effect would this mutation have on the G protein and the effector protein?
7. Explain how FRET could be used to monitor the association of $G_{\alpha s}$ and adenylyl cyclase following activation of the

epinephrine receptor.

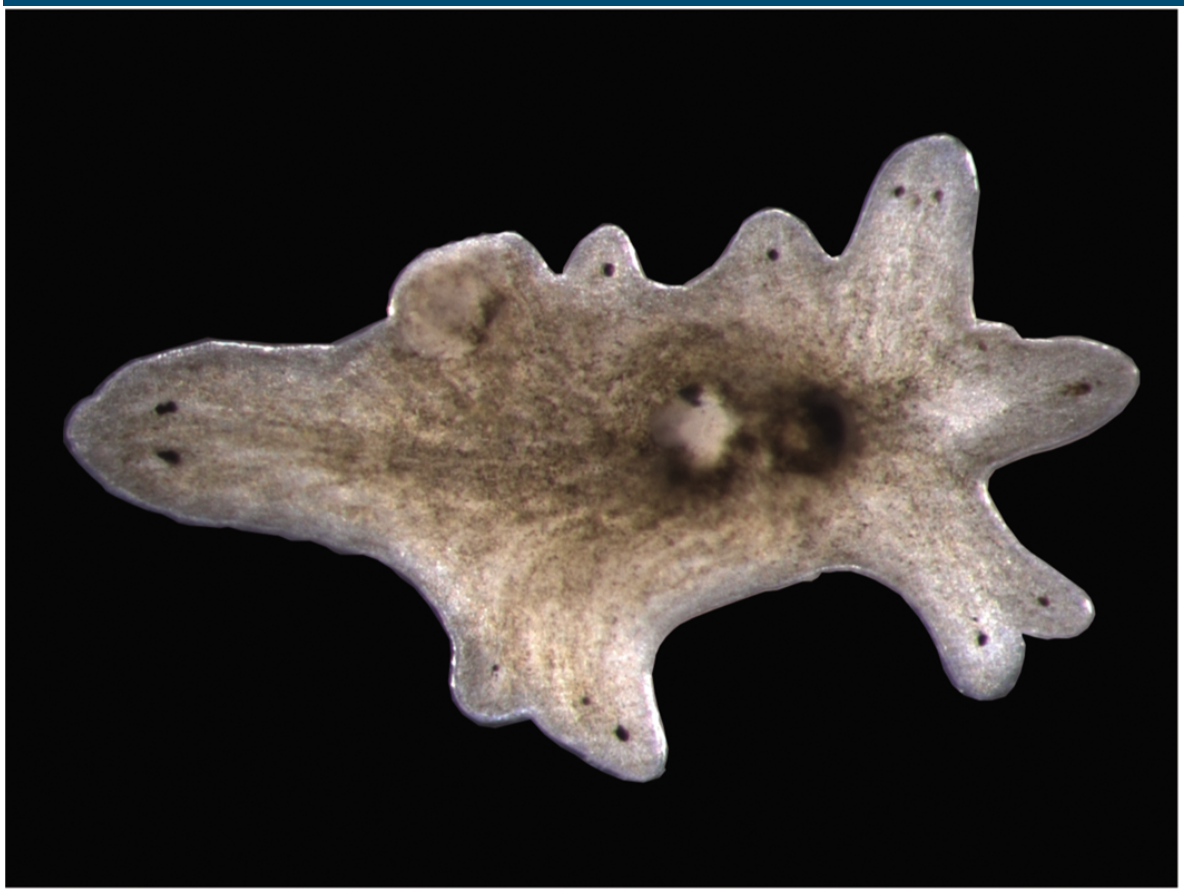
8. Which of the following steps amplify the epinephrine signal response in cells: receptor activation of G protein, G protein activation of adenylyl cyclase, cAMP activation of PKA, or PKA phosphorylation of glycogen phosphorylase kinase (GPK)? Which change will have a greater effect on signal amplification: an increase in the number of epinephrine receptors or an increase in the number of $G_{\alpha s}$ proteins?
9. The cholera toxin, produced by the bacterium *Vibrio cholerae*, causes a watery diarrhea in infected individuals. What is the molecular basis for this effect of cholera toxin?
10. Both rhodopsin in vision and the muscarinic acetylcholine receptor in cardiac muscle are coupled to ion channels via G proteins. Describe the similarities and differences between these two systems.
11. In liver and muscle, epinephrine stimulation of the cAMP pathway activates glycogen breakdown and inhibits glycogen synthesis, whereas in adipose tissue, epinephrine activates hydrolysis of triglycerides, and in other cells, it causes a diversity of other responses. What step in the cAMP signaling pathways in these cells specifies the cell response?
12. Continuous exposure of a G protein–coupled receptor to its ligand leads to a phenomenon known as *desensitization*. Describe several molecular mechanisms for receptor desensitization. How can a receptor be reset to its original sensitized state? What effect would a mutant receptor

lacking serine or threonine phosphorylation sites have on a cell?

13. What is the purpose of A kinase–anchoring proteins (AKAPs)? Describe how AKAPs work in heart muscle cells.
14. Inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG) are second messenger molecules derived from the cleavage of phosphatidylinositol 4,5-bisphosphate [$PI(4,5)P_2$] by activated phospholipase C. Describe the role of IP_3 in causing a rise in cytosolic Ca^{2+} concentration. How do cells restore resting levels of cytosolic Ca^{2+}? What is the principal function of DAG?
15. In Chapter 3, the Ca^{2+} of calmodulin EF hands for binding Ca^{2+} is given as 10^{-6} M. Many proteins have much higher affinities for their respective ligands. Why is the specific affinity of calmodulin important for Ca^{2+} signaling processes such as that initiated by production of IP_3?
16. Most of the short-term physiological responses of cells to cAMP are mediated by activation of PKA. Another common second messenger is cGMP. What are the targets of cGMP in rod cells?

CHAPTER 16

Growth Factor and Cytokine Signaling Pathways That Control Gene Expression



Courtesy of Christian Petersen and Peter Reddien, MIT, Whitehead Institute.

A planarian with multiple heads. Several hormones and their receptors regulate the regeneration of complex body parts in wounded planaria; the extracellular signaling protein Wnt promotes tail regeneration and inhibits head regeneration. In an experiment, the gene encoding *beta-catenin-1*, an essential protein in the Wnt signal transduction pathway, was inhibited by feeding animals an inhibitory double-stranded RNA. During the subsequent

month, normal uninjured animals developed heads around their periphery in a process induced by small wounds caused by defects in stem-cell mediated tissue turnover and repair; the complete absence of a Wnt signal at these wounds promotes formation of a head.

OUTLINE

16.1 Growth Factors and Their Receptor Tyrosine Kinases
16.2 The Ras/MAP Kinase Signal Transduction Pathway
16.3 Phosphoinositide Signal Transduction Pathways
16.4 Cytokines, Cytokine Receptors, and the JAK/STAT Signaling Pathway
16.5 The TGF-β Family of Growth Factors, Their Receptor Serine Kinases, and the Smad Transcription Factors They Activate
16.6 Signal Transduction Pathways That Utilize Regulated, Site-Specific Protein Cleavage: Notch/Delta and EGF Precursors
16.7 Signal Transduction Pathways That Utilize Proteasomal Degradation of Signaling Components: Wnt, Hedgehog, and the Many Hormones That Activate NF-κB

Hundreds of extracellular signaling molecules regulate the growth and fate of cells and in this way control the development of all metazoans; these signaling molecules can have both short- and long-term effects on cells. Short-term effects are usually triggered by modification of existing enzymes or other proteins, as we saw in Chapter 15. Long-term changes in cell function, however, require changes in gene expression. Thus many extracellular signaling molecules alter the expression of genes and in doing so can alter cell division and cell differentiation. For example, signaling molecules called *cytokines* induce the body's production of red blood cells, white blood cells, and platelets. Once differentiated, cells respond to their environment by changing their shape, metabolism, or movement. Here again, extracellular signaling molecules may lead

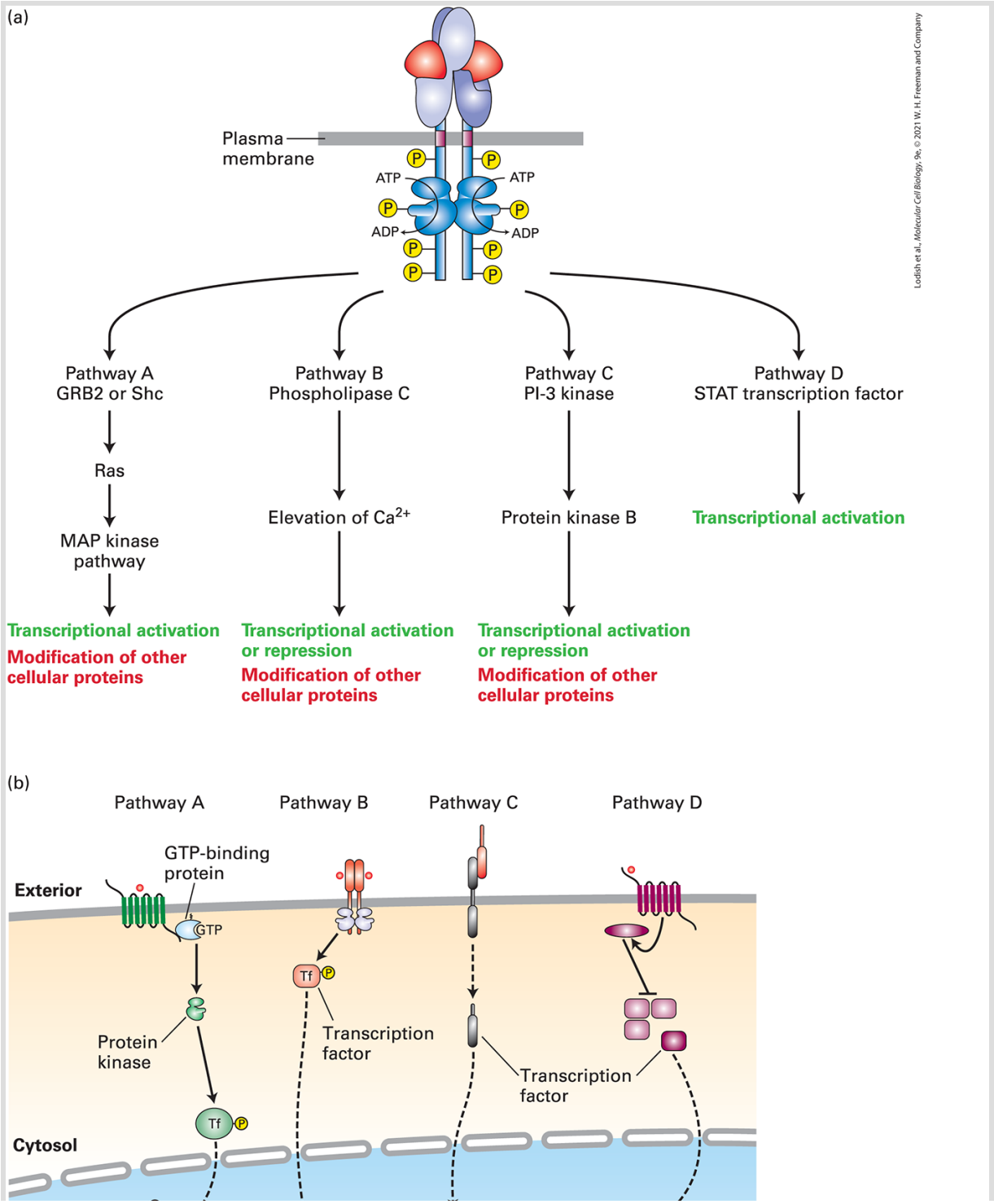
to changes in gene expression. In response to infection, for example, several hormones activate or repress the expression of several hundred genes in immune system cells that carry out multiple responses to infection. Given the extensive role of gene transcription in mediating critical aspects of development, metabolism, and movement, it is not surprising that mutations in such signaling pathways cause many human diseases, including cancer, diabetes, and immune-system disorders.

In this chapter, we explore the main groups of hormones, receptors, and their associated intracellular signal transduction pathways that cells use mainly to influence gene expression. In eukaryotes, there are about a dozen classes of highly conserved cell-surface receptors, and these receptors activate several types of highly conserved and often complex intracellular signal transduction pathways, leading to both short-term and long-term effects. These pathways have been conserved throughout evolution and operate in much the same manner in flies, worms, planaria, mice, and humans; their commonalities have enabled researchers to study them in a variety of experimental systems. For instance, the signaling protein Hedgehog (Hh) and its receptor were first identified in fruit fly *(Drosophila)* mutants that had impaired development. Subsequently, the human and mouse homologs of these proteins were cloned and shown to participate in a number of important signaling events during cell differentiation, resulting in the discovery that abnormal activation of the Hh pathway occurs in several human tumors.

Despite the complexity of many of these signal transduction pathways, there are features common to many of them. We learned about some in our study of GPCR signaling, and we will encounter others in the following sections. Signal transduction pathways that regulate gene expression act through transcription factors that switch genes on or off. In non-stimulated cells, most regulated transcription factors are sequestered in the cytosol, unable to move into the nucleus, bind to DNA, or affect gene expression. Activation of the relevant signal transduction pathway triggers transport of the transcription factor into the nucleus, in some cases by releasing it from an inhibitory complex in the cytosol, in others by unmasking a nuclear localization sequence. Alternatively, a kinase in the signal transduction pathway may phosphorylate the transcription factor, converting it into an active form that is able to bind to DNA. Another common theme is negative feedback regulation through changes in gene expression; a regulated transcription factor may induce synthesis of proteins that repress or dampen a signal transduction pathway — preventing overstimulation of the cell.

We begin this chapter with a large class of receptors — the Receptor Tyrosine Kinases (RTKs; Section 16.1) — that are activated by a large group of protein growth factors. Ligand binding to the extracellular domain induces two monomers to bind together as a dimer, activating a protein kinase contained in its cytosolic domain (Figure 16-1a). One important RTK signaling pathway includes the activation of the small GTP-binding protein termed Ras; Ras in turn activates a cascade of protein kinases in which one kinase activates another —

the Ras/MAP Kinase Pathway (Figure 16-1a, Pathway A; Section 16.2). The final kinase phosphorylates and activates one or more transcription factors (Tfs).



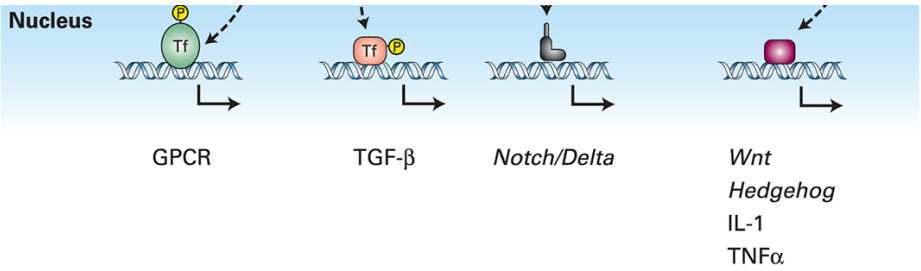


FIGURE 16-1 Common types of cell-surface receptors and signal transduction pathways that regulate gene expression. (a) Signal transduction pathways activated by receptor tyrosine kinases (RTKs) and cytokine receptors. The cytosolic domain of RTKs contains an intrinsic protein tyrosine kinase domain; the cytosolic domain of cytokine receptors bind a separate protein tyrosine kinase. In the absence of ligand these receptors generally exist as monomers that have poor kinase activity. Binding of ligands to the extracellular domains of these receptors triggers dimerization of two monomers and activation of the tyrosine kinase and phosphorylation of several tyrosine residues in the receptor's cytosolic domain. The resulting phosphotyrosines function as docking sites for binding and activating several downstream signal-transducing proteins. Pathway A: Binding of one type of adapter protein (e.g., GRB2) to an activated receptor leads to activation of the small GTP-binding "switch" protein Ras (see Section 16.2). Signal transduction pathways downstream of Ras involve several kinases; in these MAP kinase pathways, one kinase phosphorylates and thus activates the activity of another. This leads to phosphorylation and activation of transcription factors, which are often different in different cells. Pathways B and C: Two signal transduction pathways that involve phosphoinositides are triggered by recruitment of phospholipase C and PI-3 kinase, respectively, to the plasma membrane (see Section 16.3). Elevated levels of Ca^{2+} and activated protein kinase B modulate the activity of transcription factors as well as of cytosolic proteins that are involved in metabolic pathways or cell movement or shape change. Pathway D: In a pathway mainly employed by cytokine receptors, a STAT transcription factor binds to the activated receptor, becomes phosphorylated, dimerizes, moves to the nucleus, and directly activates transcription. (b) Other common types of cell-surface receptors and signal transduction pathways. Pathway A: Many GPCRs activate the heterotrimeric GTP-binding $G_{\alpha,s}$ protein, ultimately leading to activation of protein kinase A, and phosphorylation and activation of transcription factors such as CREB. Pathway B: The cytosolic domains of receptors for the TGF-β family of signaling proteins contain a serine-threonine kinase, which directly phosphorylates and activates a member of the Smad class of transcription factors, unmasking a nuclear localization signal. Pathway C: Binding of a Delta ligand to the extracellular domain of a

member of the Notch family of receptors triggers proteolytic cleavage of the receptor, releasing its cytosolic domain, which moves into the nucleus and regulates gene expression. Pathway D: Signal transduction pathways activated by binding of members of the Wnt, Hedgehog, or Interleukin 1 (IL-1) families of ligands to their respective receptors lead to ubiquitination and degradation of components of multiprotein complexes in the cytosol, releasing a transcription factor that then translocates into the nucleus.

Alternatively, RTKs can initiate signal transduction pathways containing phosphorylated inositols that often although not always lead to changes in gene expression. In Chapter 15 we saw one that leads to an elevation in the concentration of Ca^{2+} in the cytosol (Figure 16-1a, Pathway B). Depending on the type of cell, this pathway triggers processes as diverse as protein secretion and cell migration. RTKs also induce other signaling pathways involving phosphorylated inositols that activate yet other protein kinases, including Protein Kinase B, that in turn alter gene expression to regulate cellular metabolism and cell death (Figure 16-1a, Pathway C; Section 16-3).

Cytokines activate a distinct class of signaling receptors. The cytokines form a large family of protein hormones that regulate, for example, the formation and function of all types of blood cells, including all of the many kinds of cells that compose our immune system (Chapter 24). Cytokine receptor proteins are similar to RTKs in that ligand binding leads to the activation of a protein tyrosine kinase (JAK), although in this case the kinase is a separate protein tightly associated with the receptor's cytosolic domain. Cytokine

receptors and RTKs activate many of the same signal transduction pathways (see Figure 16-1a), and in Section 16-4 we learn how a JAK kinase directly phosphorylates a STAT transcription factor (Figure 16-1a, Pathway D). Phosphorylation leads to the exposure of a nuclear localization signal that guides the transcription factor to the nucleus.

The large TGF-β family of extracellular signaling proteins acts to regulate many developmental pathways. Like RTKs, these structurally distinct receptors employ activation of a kinase that is part of their cytosolic domain. The activated kinase directly phosphorylates transcription factors in the Smad family of DNA-binding proteins (Figure 16-1b, Pathway B, and Section 16.5). In this case, phosphorylation unmasks a nuclear localization signal that allows the transcription factor to move directly to the nucleus.

Not all of the signaling receptors described in this chapter lead to the activation of protein kinases or GTP-binding proteins in the earliest steps of their signal transduction pathways. For example, binding of a ligand to the extracellular domain of certain receptors, such as Notch, triggers proteolytic cleavage of the receptor itself, releasing the cytosolic domain. The released domain is a transcription factor that travels into the nucleus and interacts with other transcription factors to alter gene expression (Figure 16-1b, Pathway C, and Section 16-6).

Finally, we discuss several different pathways in which binding of the ligand to its receptor ultimately leads to ubiquitination and degradation of one or more proteins in a multiprotein cytosolic complex, consequently releasing a transcription factor that is part of the complex and that then translocates into the nucleus and affects gene expression (Figure 16-1b, Pathway D). These include signaling pathways activated by the Wnt and Hedgehog signaling proteins, and others downstream of several types of receptors that all lead to activation of the NF-κB transcription factor.

As with GPCRs, many receptors discussed in this chapter are expressed in multiple types of body cells, but activation of these receptors by the same hormone triggers induction (or repression) of very different sets of genes in different cell types. As we learned in Chapter 8, the epigenetic state of the cell is determined by its developmental history. Whether a transcription factor activated by a cell-surface receptor induces (or represses) a particular gene in a particular cell depends, first (Figure 16-2), on the epigenetic state of the gene: whether the gene is in an active "open" chromatin conformation, and therefore accessible to binding by the transcription factor, or in a silenced "closed" state that is not accessible by the transcription factor. In other words, a given transcription factor can potentially bind to multiple gene regulatory sites in chromosomal DNA, but in any given cell type only a fraction of these sites will be accessible for binding.

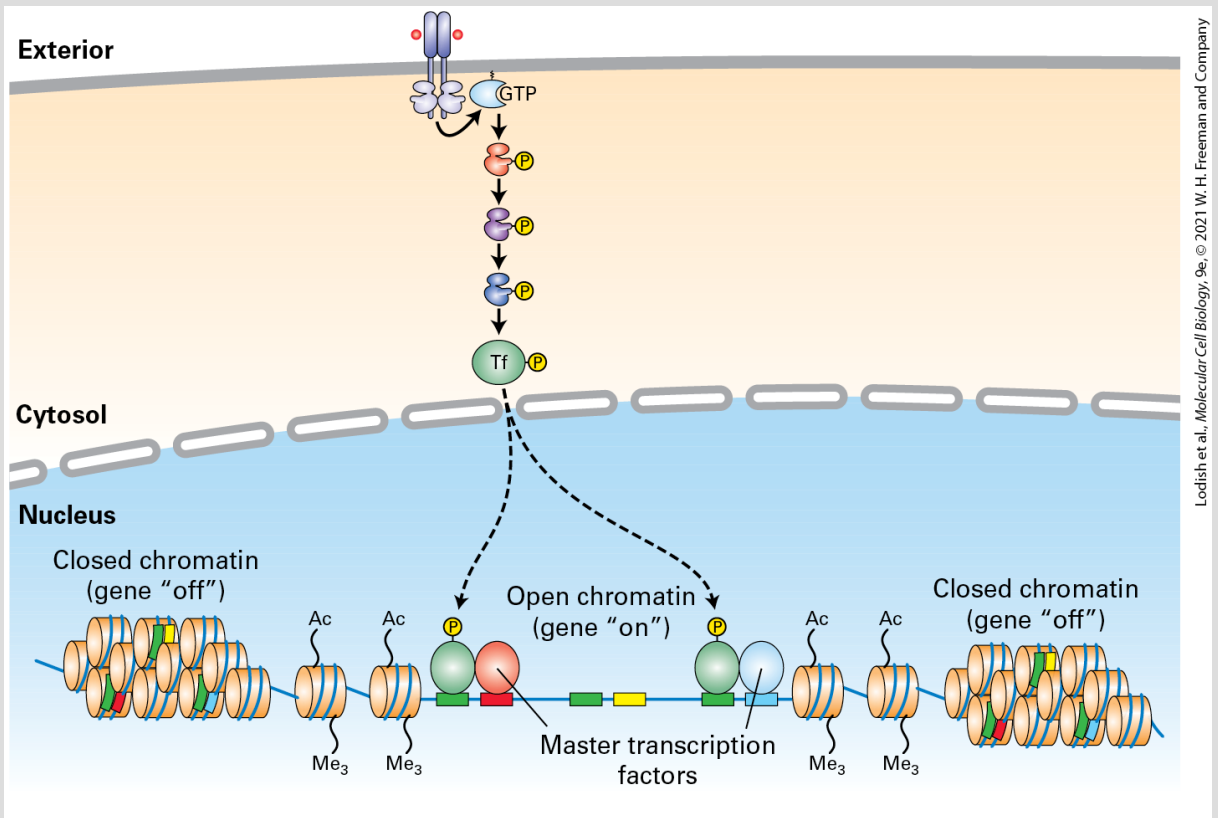


FIGURE 16-2 Induction of a particular gene by a transcription factor depends on DNA-binding sites for the factor, on the gene's epigenetic state, and on the presence of master transcription factors and other nuclear proteins. Any given activated transcription factor has multiple sites on the chromosomal DNA to which it can potentially bind (green), but in any given cell it will bind only to those sites that are in an "open" chromatin conformation and in which specific master transcription factors or other cell-specific proteins (blue and red, respectively) are bound to adjacent sites on the DNA; together these proteins activate (or repress) expression of the adjacent gene. Other potential transcription-factor binding sites are adjacent to binding sites for other master transcription factors (yellow) that are not expressed in this cell type, and thus the activated transcription factor will not bind to those sites.

Additionally, many cell types express one or more *master transcription factors* that determine the identity and developmental fate of the cell; a recent finding is that many transcription factors

activated by cell-surface receptors bind to chromosomal DNA at regulatory sites — mainly enhancers — adjacent to these master factors, and together these transcription factors induce (or repress) cell-specific genes (see Figure 16-2).

No cell-surface receptor, and no signal transduction pathway, works in isolation. Expression of many if not most genes in metazoans are regulated by multiple transcription factors, and these in turn are activated or repressed by different intracellular signaling pathways, each of which can be regulated by multiple extracellular signals. An emerging subfield of molecular cell biology — systems analysis — uses a combination of experimental and computational analyses to understand how cells of different types integrate these signals over time. For example, in Chapter 21 we consider how multiple hormones and multiple signal transduction pathways in different types of cells regulate the body's needs for the key metabolites glucose and fatty acids.

16.1 Growth Factors and Their Receptor Tyrosine Kinases

The extracellular signaling molecules that activate receptor tyrosine kinases (RTKs) constitute a large class of soluble and membrane-bound protein hormones, including many that were initially identified because they stimulate the growth of specific types of cells (hence the name *growth factor*). Those RTK ligands, such as nerve growth factor (NGF), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), and epidermal growth factor (EGF), trigger proliferation of cells that bear a cognate receptor. Others, such as insulin, regulate expression of multiple genes that control sugar and lipid metabolism in liver, muscle, and adipose (fat) cells. The large family of ephrins play critical roles in regulating the migration of many nerve cell axons to their target destinations (see Chapter 23). Many RTKs and their ligands were first identified in studies of human cancers, in which mutant forms of growth-factor receptors stimulate cell proliferation even in the absence of the growth-factor ligand. The mutation locks the receptor conformation in a state such that the kinase is permanently active, even if the ligand is not present (*constitutively* active; see Chapter 25). Other RTKs were uncovered during analysis of developmental mutations that lead to blocks in differentiation of specific cell types in *Caenorhabditis elegans, Drosophila*, and mice.

Binding of Ligand to the Extracellular Domain of an RTK Leads to Dimerization and Activation of Its Intrinsic Cytosolic Tyrosine Kinase

All RTK proteins have four essential components: an extracellular domain containing a ligand-binding site, a single hydrophobic transmembrane α helix, a cytosolic segment that includes a domain with protein tyrosine kinase activity, and a C-terminal segment that contains tyrosine residues that become phosphorylated by the receptor's own kinase (Figure 16-3). Most RTKs are monomeric, and ligand binding to the extracellular domain induces two monomers to join together as a dimer.

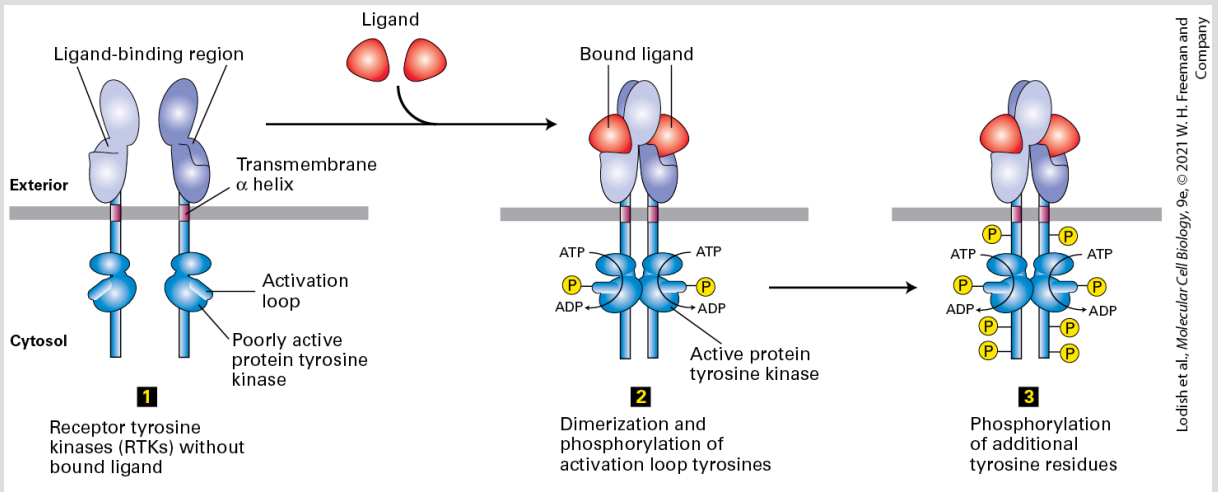


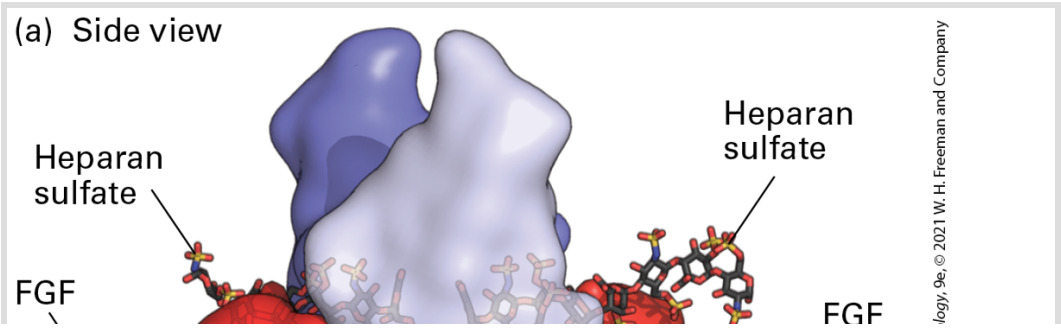
FIGURE 16-3 General structure and activation of receptor tyrosine kinases (RTKs). The cytosolic domain of RTKs contains an intrinsic protein tyrosine kinase domain. In the absence of ligand (step 1), RTKs generally exist as monomers with poor kinase activity. Binding of two ligands to the extracellular domains of two RTKs forms or stabilizes an activated dimeric receptor. This brings together two poorly active kinases such that each

one phosphorylates the other on a tyrosine residue in the activation loop (step **2**). Phosphorylation increases the activity of the kinase by causing the loop to move out of the kinase catalytic site, thus increasing the ability of ATP and/or the protein substrate to bind (see Chapter 3). The activated kinase then phosphorylates several tyrosine residues in the receptor's cytosolic domain (step **3**). The resulting phosphotyrosines function as docking sites for SH2 and other binding domains on downstream signal-transducing proteins, as summarized in Figure 16-1a.

RTK activation and signaling can be summarized as follows: in the resting, unstimulated (no ligand bound) state, the intrinsic kinase activity of an RTK is very low (Figure 16-3, step **1**). Like most other kinases, the cytosolic kinase domain of an RTK contains a flexible domain termed the **activation loop**. In the resting state, the activation loop is not phosphorylated and assumes a conformation that blocks kinase activity. In some receptors (e.g., the insulin receptor), this loop prevents binding of ATP. In others (e.g., the FGF receptor), it prevents binding of substrate. Binding of ligand causes a conformational change that promotes the dimerization of the RTK. The extracellular domains of two RTK monomers join, which brings their transmembrane segments — and therefore their cytosolic domains — close together and results in activation of the tyrosine kinases. For most, but not all, RTKs, the weakly active kinase in each subunit is thought to phosphorylate a key tyrosine residue in the activation loop of the other subunit (Figure 16-3, step **2**). This phosphorylation leads to a conformational change in the activation loop that unblocks the kinase active site, reducing the K_d for either ATP or the substrate to be phosphorylated and greatly enhances kinase activity (see Chapter 3). The strongly active kinase then

phosphorylates additional tyrosine residues in the cytosolic domain of the receptor (Figure 16-3, step **3**). These phosphotyrosine residues serve as binding sites for proteins that then activate several signal transduction pathways (see Figure 16-1a), each of which will be detailed in subsequent sections of this chapter.

Although dimerization is a necessary step in the activation of all RTKs, functional dimers can be formed in multiple ways, as we shall see. For example, many receptors are dimerized in a manner similar to the fibroblast growth factor (FGF) receptor (Figure 16-4), in which each of two FGF ligands binds simultaneously to the extracellular domains of both subunits of the dimerized receptor, stabilizing the dimer. An additional feature is that FGF binding, and thus FGF receptor dimerization, is enhanced by the additional binding to both the ligand FGF and its receptor of the negatively charged polysaccharide heparan sulfate (see Figure 16-4), which is a component of some cell-surface proteins and of the extracellular matrix (see Chapter 20). The participation of the heparan sulfate is essential for efficient receptor activation; we see later (Section 16.5) that binding of other hormones to components of the extracellular matrix is also essential for their signaling functions.



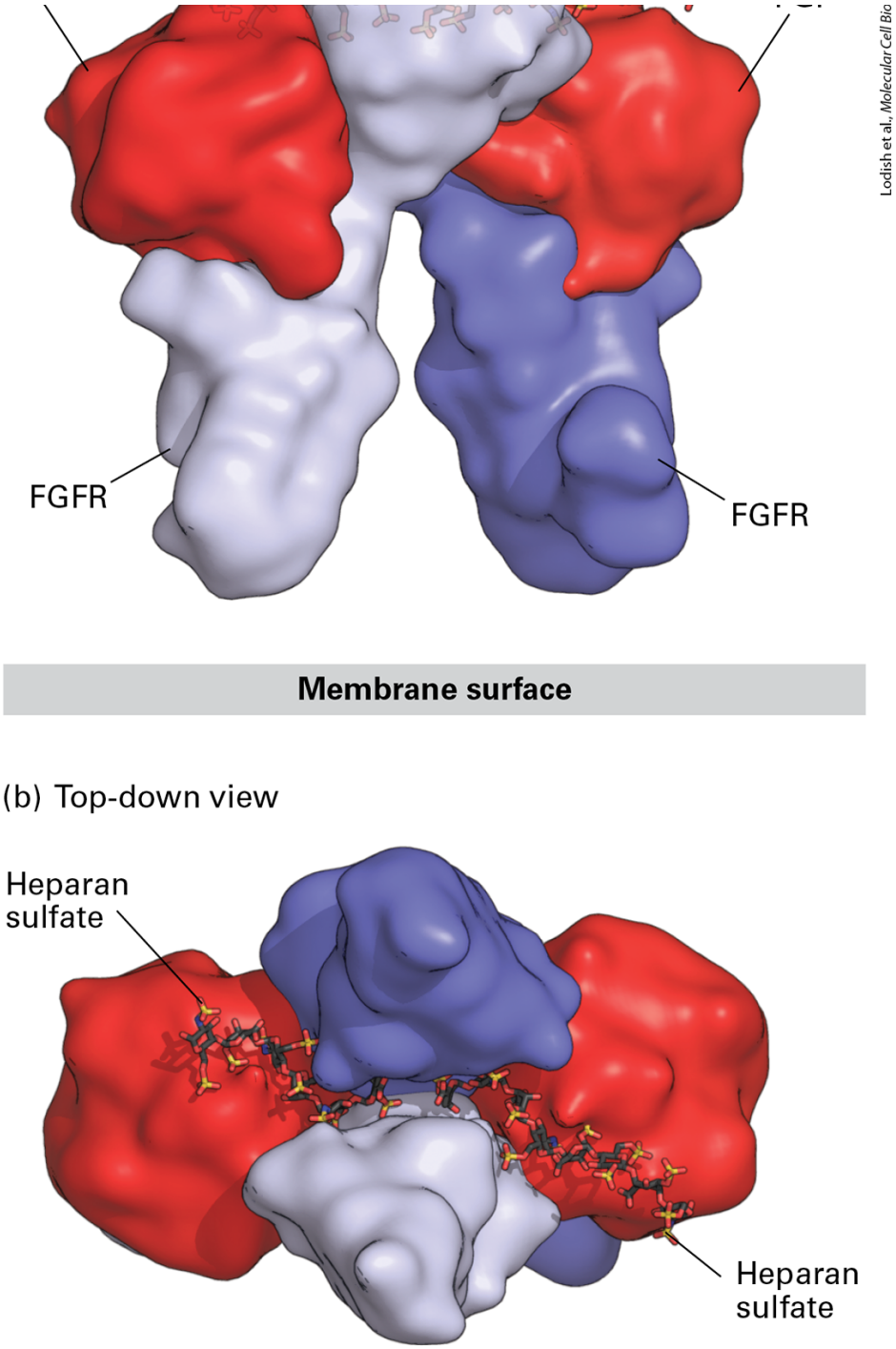


FIGURE 16-4 Structure of the extracellular domains of the active dimeric fibroblast growth factor (FGF) receptor, stabilized by binding of two ligands and by heparan sulfate. Shown here are side and top-down views of the complex comprising the extracellular domains of two FGF receptor (FGFR) monomers (purple and violet), two bound FGF molecules (red), and two short

heparan sulfate chains that are bound tightly to FGF. (a) In the side view, the upper domain of one receptor monomer (purple) is seen situated behind that of the other (violet); the plane of the plasma membrane is at the bottom. A small segment of the receptor's extracellular domain connects to the membrane-spanning α-helical segment of each of the two receptor monomers (not shown) that protrude downward into the membrane. (b) In the top view, the heparan sulfate chains are seen threading between and making numerous contacts with the upper domains of both receptor monomers. These interactions enhance binding of the FGF ligand to the receptor and receptor dimerization.

[Data from J. Schlessinger et al., 2000, *Mol. Cell* **6**:743, PDB ID 1fq9.]

Yet other RTKs, such as the insulin receptor, form disulfide-linked dimers even in the absence of hormone; binding of ligand to this type of *pre-dimerized* but inactive RTK alters the conformation of the dimer in such a way that the receptor's kinase becomes activated (Figure 16-5). In the absence of insulin, the structure of the dimeric receptor's extracellular domain forces the weakly active cytosolic kinase domains to be far apart, unable to phosphorylate and thus activate each other. Upon insulin binding, the receptor's extracellular domain converts from an inverted U-shape into a T-shaped conformation that brings the transmembrane domains, and thus the kinase domains, together, presumably facilitating phosphorylation of the activation loop in the tyrosine kinase domain of one receptor by the kinase in the other.

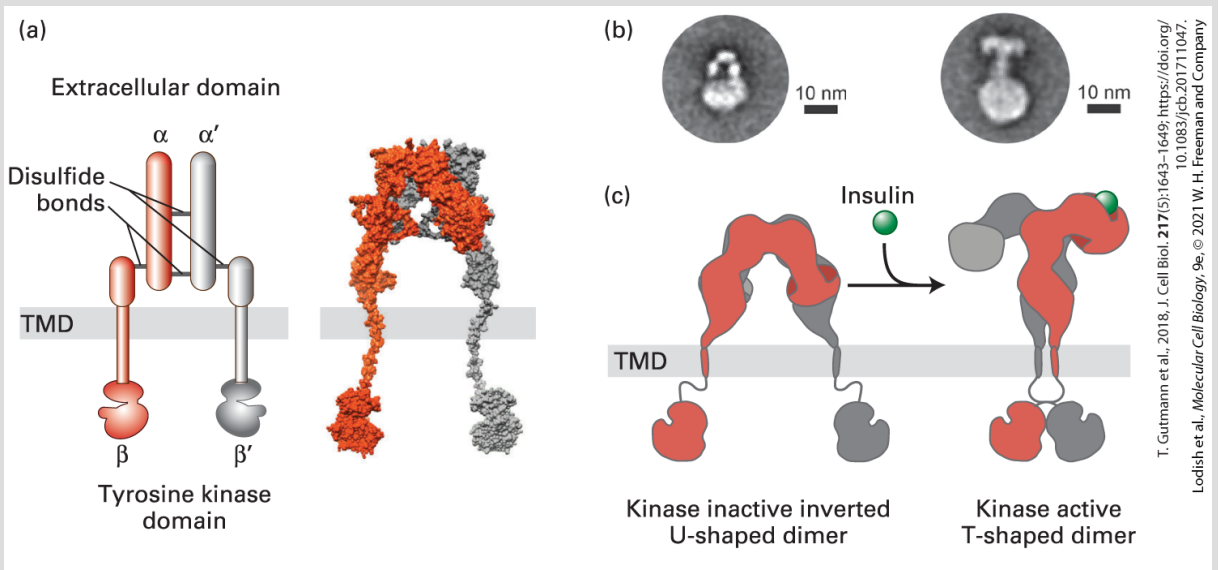


FIGURE 16-5 Insulin-induced activation of the insulin receptor. (a) Schematic cartoon showing the extracellular domain (ECD), transmembrane domain (TMD), and tyrosine kinase domain (TKD) as well as a structural model of the full-length human insulin receptor in the absence of insulin. The receptor is produced as a single polypeptide chain that undergoes cleavage in the Golgi complex; the two resulting subunits, α and β, are linked together by a disulfide bond, as are the two α subunits in a dimer. (b) Electron microscopic views of single insulin receptors reconstituted in tiny lipid discs, both in the absence and presence of insulin. (c) Cartoon illustrating the ligand (green) binding–induced conformational change in the extracellular domain of the insulin receptor and its coupling to the transmembrane domains (TMDs) with concomitant activation of the tyrosine kinase domains by phosphorylation. In the absence of insulin, the extracellular domain adopts a symmetric inverted U-shaped conformation; upon insulin binding, the extracellular domain is converted into a T-shaped conformation that brings the transmembrane domains together, presumably facilitating phosphorylation of the activation loop of one kinase by the poorly active kinase of the other (trans-phosphorylation), leading to activation of both tyrosine kinase domains.

This last example highlights the fact that simply having two identical receptor monomers in close contact is not sufficient for receptor activation — the proper conformational changes must accompany

receptor dimerization. Once an RTK converts into a functional state in a dimer, its associated tyrosine kinase becomes activated.

Homo- and Hetero-Oligomers of Epidermal Growth Factor Receptors Bind Members of the Epidermal Growth Factor Family

During evolution, the genes encoding some growth factors and their receptors underwent duplication and diverged to encode similar, but distinct, proteins with distinct functions. For example, in humans there are 22 FGF-related and 4 FGF receptor-related genes with many different functions. There are also multiple members of the *epidermal growth factor (EGF)* and EGF receptor families that have been studied intensively because of their involvement in many human diseases. Drugs targeting EGF receptors that are overproduced in tumors or that have undergone mutations that cause their kinase domains to be active even in the absence of a hormone signal are used to treat many cancers, as we learn in Chapter 25. The first known member of this family of hormones, epidermal growth factor, was originally identified by Stanley Cohen as a small protein that stimulated the early opening of eyelids in a baby mouse; it is now known to stimulate the proliferation of many types of epithelial cells. For this research and the parallel discovery

of nerve growth factor, Cohen and Rita Levi-Montalcini were awarded Nobel Prizes.

Four related EFG receptors participate in signaling by the many members of the epidermal growth factor family of signaling molecules. In the mouse these receptors are termed Erb-B1, 2, 3, and 4. In humans these receptors are called **HER** (*h*uman *e*pidermal growth factor *r*eceptor) 1, 2, 3, and 4, respectively.

EGF Receptor Homodimers

In the resting state, the majority of EGF receptor molecules are monomeric. We first consider the activation of the EGF receptor HER1 (Erb-B1). HER1 directly binds EGF as well as six other members of the EGF family; binding of any of these ligands leads to formation of a homodimer (homodimerization) of the HER1 extracellular domain (Figure 16-6). Binding of EGF binding triggers a dramatic conformational change in the extracellular domain of HER1, exposing a portion of the receptor that is inaccessible in the inactive monomer, termed the *dimerization arm*. The dimerization arms from two receptor monomers then bind tightly together, forming a receptor homodimer that is further stabilized by interactions between another extracellular domain (domain IV) in each of the two receptors. Several members of the EGF family — neuregulins 1, 2, 3, and 4 — bind to HER4 that, like HER1, can also

form homodimers. As detailed below, formation of a receptor dimer triggers activation of the kinase domain.

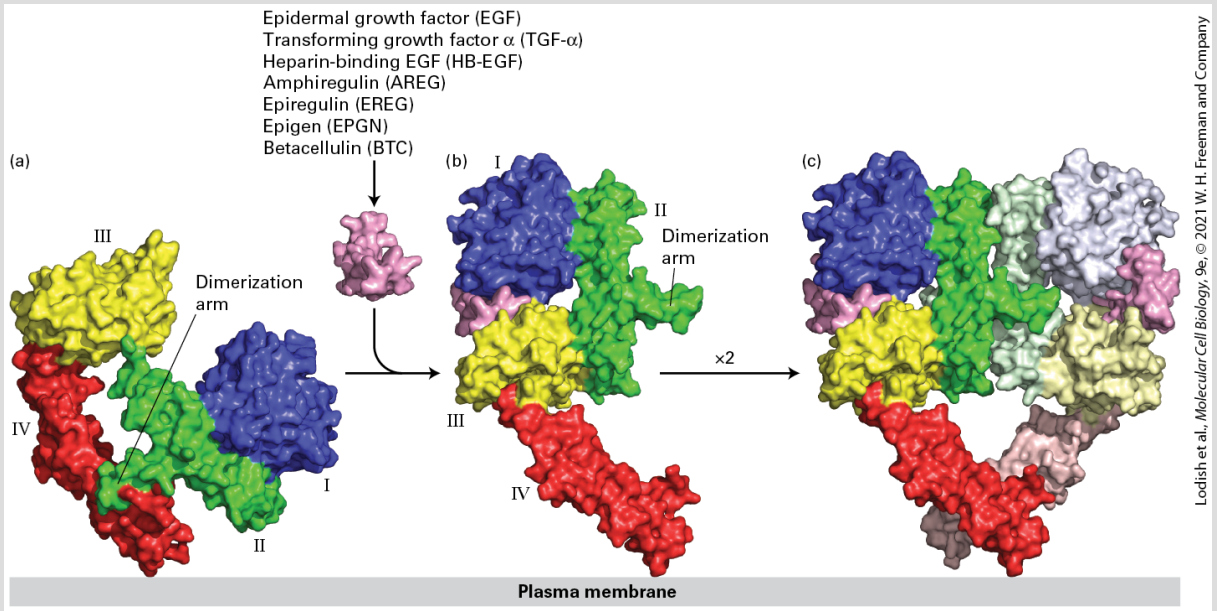
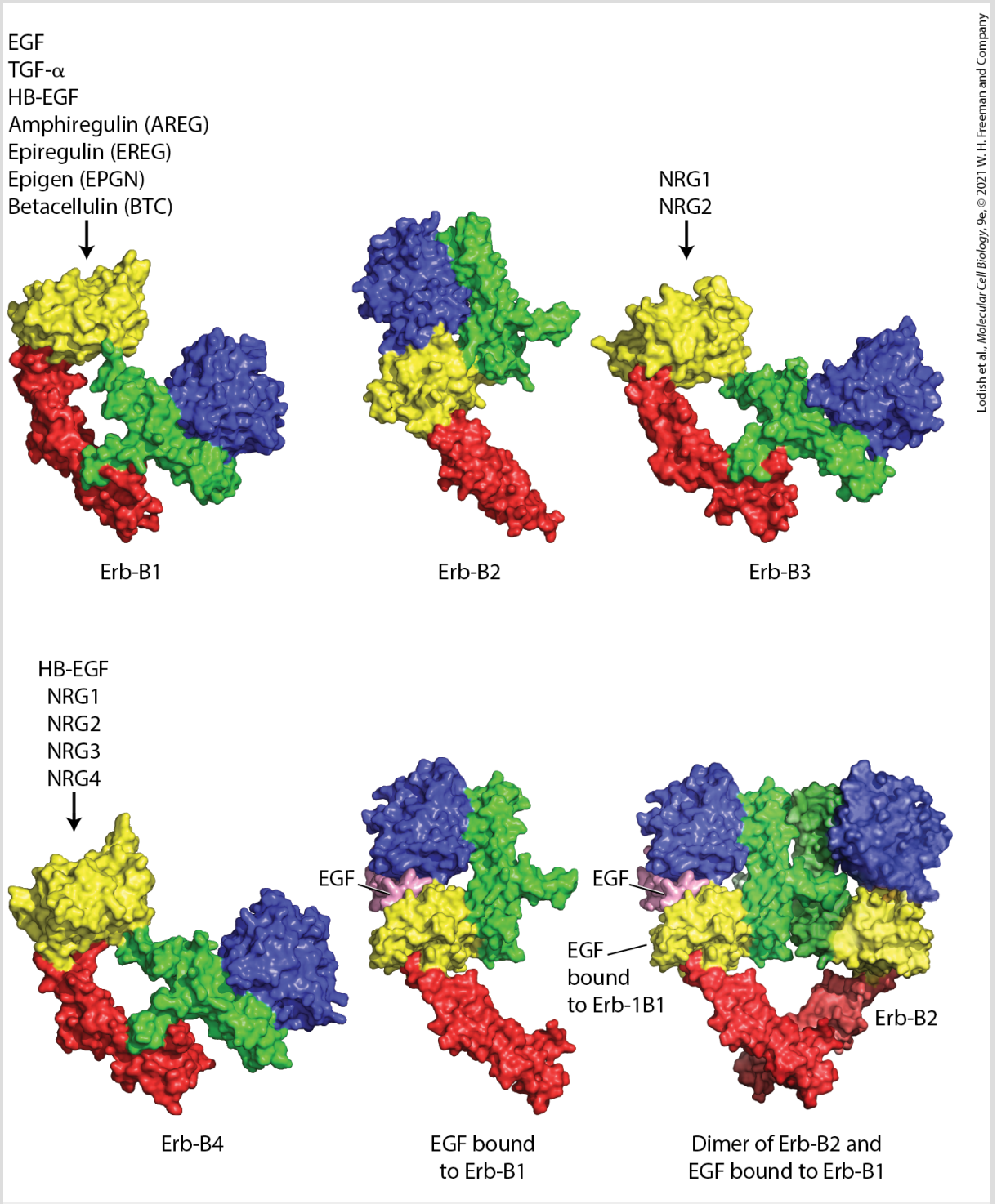


FIGURE 16-6 Ligand-induced dimerization of HER1, a human receptor for epidermal growth factor (EGF). (a) The extracellular region of all EGF receptors contains four domains: domains I (blue) and III (yellow) are closely related in sequence, as are domains II (green) and IV (red). In the absence of bound EGF, the receptor is mostly monomeric and the intracellular kinase is inactive. The extracellular region adopts a configuration in which the β-hairpin from domain II that forms the *dimerization arm* binds to domain IV of the same receptor molecule. (b) EGF, as well as the six other listed members of the EGF family, binds simultaneously to domains I and III; binding induces a major conformational change in the extracellular domain, exposing the dimerization arm of domain II. (c) Dimerization of two identical ligand-bound receptor monomers in the plane of the membrane occurs primarily through interactions between the dimerization arms of the two receptors.

[Data from H. Ogiso et al., 2002, *Cell* **110**:775, PDB ID 1ivo.]

EGF Receptor Heterodimers with HER2

Remarkably, although one of the EGF receptors, HER2 (mouse Erb-B2), cannot bind any ligand, it participates in signaling by forming heterodimers with HER1, HER3, or HER4 when they are bound to their respective ligands. HER2 can form heterodimers with these other EGF receptors because it adopts an active conformation with its dimerization arm always accessible for binding to the dimerization arms of ligand-bound receptors. Thus, by HER2 forming heterodimers, its kinase domain becomes activated and in so doing it facilitates signaling by all EGF family members that bind to one of the other three EGF receptors (Figure 16-7).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 16-7 The HER family of receptors and their ligands. Humans and mice express a family of four receptor tyrosine kinases, denoted HER1, 2, 3, and 4 in humans and Erb-B1, 2, 3, and 4 in mice and other animals. Only the extracellular domains of these receptors are depicted here. These receptors bind epidermal growth factor (EGF) and the other EGF family

members: heparin-binding EGF (HB-EGF), transforming growth factor α (TGF-α), amphiregulin (AREG), epiregulin (EREG), epigen (EPGN), betacellulin (BTC), and the four neuregulins (NRG1–4). Notice that Erb-B2 (HER2), which does not directly bind a ligand, exists in a conformation that is very similar to that of the activated Erb-B1 with a bound EGF. Erb-B2 can form a heterodimer with ligand-activated Erb-B1, 3, or 4; thus Erb-B2 facilitates signaling by all EGF family members. Erb-B3 (HER3) has a very poorly active kinase domain and can signal only when complexed with Erb-B2.

[Erb-B1 data from K. M. Ferguson et al., 2003, *Mol. Cell* **11**:507, PDB ID 1nql. Erb-B2 data from H.-S. Cho et al., 2003, *Nature* **421**:756, PDB ID 1n8z. Erb-B3 data from H. S. Cho and D. J. Leahy, 2002, *Science* **297**:1330, PDB ID 1m6b. Erb-B4 data from S. Bouyain et al., 2005, *Proc. Nat'l. Acad. Sci. USA* **102**:15024, PDB ID 1ahx. EFG and Erb-B1 data from H. Ogiso et al., 2002, *Cell* **110**:775, PDB ID 1ivo.]

HER3 is an EGF receptor that can bind ligands but lacks a functional kinase domain. HER3 can still participate in signaling; after binding a ligand, it heterodimerizes with HER2 and activates the HER2 kinase. Later we describe how the active kinases in the receptor cytoplasmic domains mediate downstream signal transduction pathways.

An increase in the numbers of HER2 molecules on the cell surface makes a cell more sensitive to signaling by many EGF family members, because the greater number of HER2 receptors enhances the rate and extent of formation of signaling heterodimers following binding of a member of the EGF family to HER1, HER3, or HER4. As we learn in Chapter 25, in about 25 percent of breast cancers the tumor cells have multiple copies of the *HER2* gene and thus elevated levels of HER2. Cells that overproduce HER2 are hypersensitive to ambient levels of all members of the EGF family of hormones, levels

that normally would not lead to activation of many receptors. As a consequence, at even low levels EGF family members can stimulate growth of these tumor cells inappropriately. Monoclonal antibodies that bind HER2 and thereby block signaling by EGF have proved useful in treatment of breast and other tumors that overexpress HER2. ■

Ligand Binding to the EGF Receptor and Receptor Dimerization Results in the Formation of an Active Asymmetric Kinase Domain Dimer

In the case of most receptor tyrosine kinases, the kinase domain is activated by phosphorylation of a tyrosine in the activation loop following formation of a receptor dimer. In contrast, loop phosphorylation is not necessary for initial activation of the kinase domains of the EGF receptors, though it is important for full activation. The mechanism of activation was uncovered through structural studies of the receptor cytosolic domain in both the active and inactive states.

The kinase domains are separated from the transmembrane segment by a so-called juxtamembrane segment (Figure 16-8). As we learned in Chapters 3 and 15, kinase domains typically contain two regions termed the *N* and *C lobes*. In the inactive, monomeric state

of the EGF receptor, the activation loop is situated in the active site of the kinase near the interface of the N and C lobes, blocking the site's activity; in this way the kinase is maintained in the inactive or "off" state (Figure 16-8, *left*). Receptor dimerization brings the kinase domains of the two subunits close together, so that the C lobe of one kinase domain (called the *donor*) binds to the N lobe of the other kinase domain (the *acceptor*), producing what is called an *asymmetric kinase dimer* (Figure 16-8, *right*). The binding of the two lobes changes the conformation of the N lobe of the acceptor kinase, displacing the activation loop and activating the kinase activity of the acceptor domain, thereby initiating intracellular signal transduction.

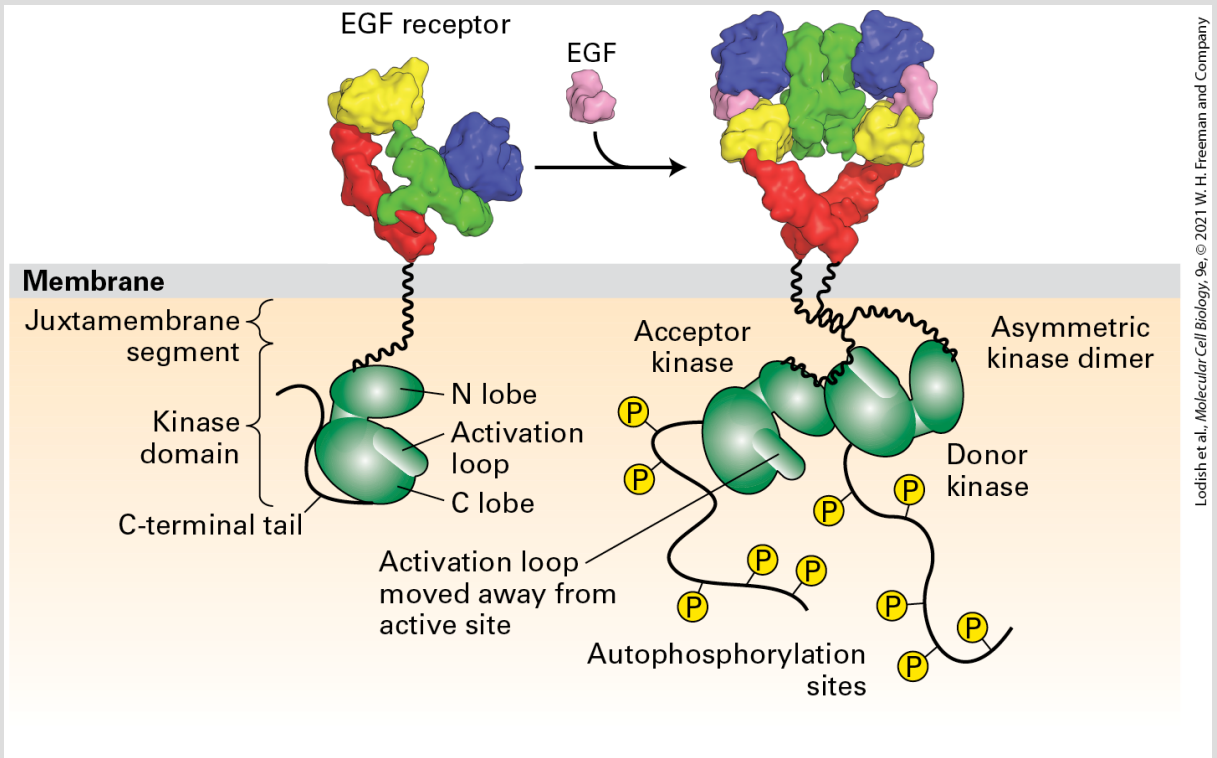


FIGURE 16-8 Activation of the EGF receptor by EGF results in the formation of an active asymmetric kinase domain dimer. In the inactive, monomeric state, the activation loop is

positioned in the kinase active site, blocking substrate binding and thus inhibiting kinase function. Receptor dimerization generates an asymmetric kinase dimer such that the C lobe of the donor kinase binds to the N lobe of the acceptor kinase in the opposite receptor; the dimer is stabilized by interactions between the juxtamembrane segments of the two receptors. These interactions cause a conformational change that pushes out the activation loop from the kinase site of the acceptor kinase, activating its kinase activity. The active acceptor kinase then phosphorylates tyrosine residues in the C-terminal segments of both receptor cytosolic domains.

[EGF receptor data from H. Ogiso et al., 2002, *Cell* **110**:775, PDB ID 1ivo; and K. M. Ferguson et al., 2003, *Mol. Cell.* **11**:507, PDB ID 1nql. Asymmetric kinase dimer data from E. Kovacs et al., 2015, *Annu. Rev. Biochem.* **84**:739.]

Thus evolution has produced many variations on the theme of ligand-induced RTK signaling: RTKs are activated by dimerization, but different receptors use different mechanisms to accomplish this. Similarly, kinases become activated by movement of the activation loop away from the kinase catalytic site, but, again, different receptors use different mechanisms to accomplish this.

Signal Transduction After Activation of RTKs: Phosphotyrosine Residues on the Receptor Are Binding Surfaces for Multiple Proteins with SH2 Domains

Once the RTK kinases become activated, they phosphorylate several tyrosine residues on their cytosolic domains (see Figures 16-3 and 16-8). Each of these phosphotyrosine residues, together with a few

adjacent residues, serve as binding sites for signal-transducing proteins that have conserved phosphotyrosine-binding domains (see Figure 16-1a). The recruitment of each signal-transducing protein to the activated receptor sets off a specific subsequent signal transduction pathway.

There are several classes of phosphotyrosine-binding domains; one is called the *SH2 domain*. An SH2 domain is able to bind to a short target sequence of polypeptide that contains a phosphotyrosine. The SH2 domains in different signal-transducing proteins have very similar three-dimensional structures, but each binds to a distinct, short target sequence of amino acids adjacent to a specific phosphotyrosine residue (often abbreviated "pY"). The unique amino acid sequence of each SH2 domain determines the specific target sequence to which it binds (Figure 16-9). Each SH2 domain contains a site for binding the phosphotyrosine and sites for binding the side chains of the adjacent residues in that SH2's target sequence.

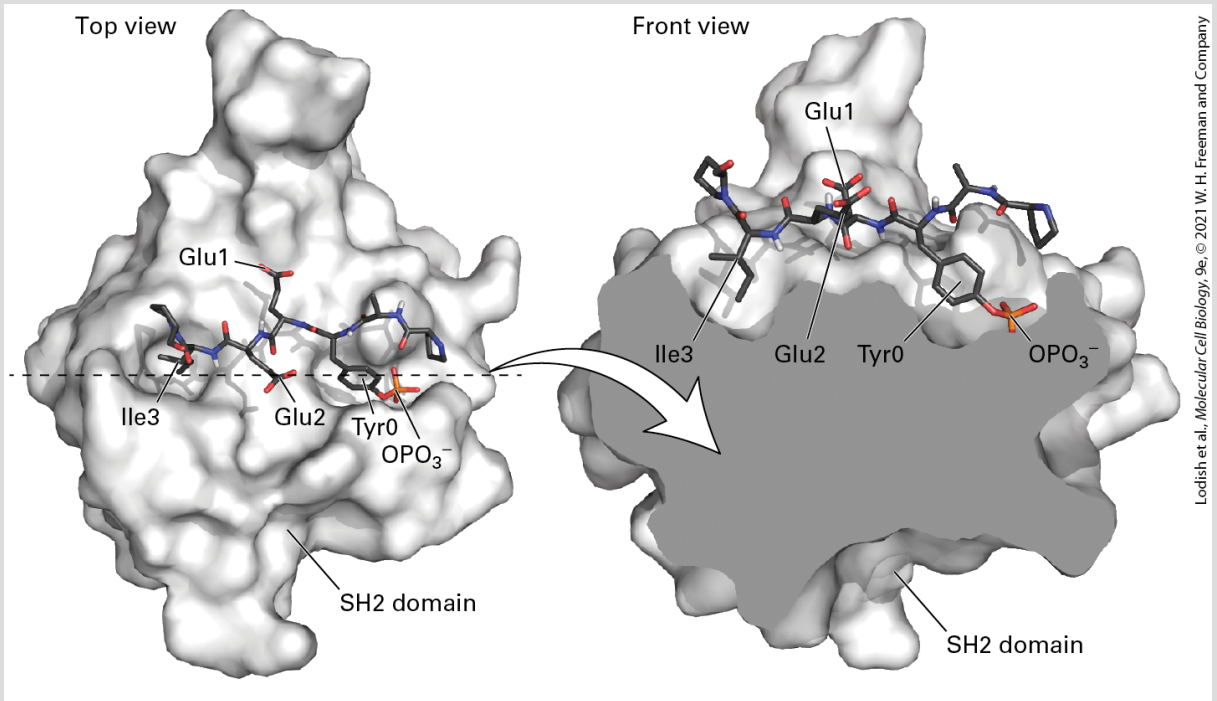


FIGURE 16-9 Surface model of an SH2 domain bound to a phosphotyrosine-containing peptide. The peptide bound by this SH2 domain, part of the Src tyrosine kinase, is shown in stick form (blue backbone with red oxygen atoms). This SH2 domain binds strongly to short target peptides containing a critical four-residue core sequence: phosphotyrosine (Tyr_0 and $\mathrm{OPO_3}^-$)–glutamic acid (Glu_1)–glutamic acid (Glu_2)–isoleucine (Ile_3). Binding resembles the insertion of a two-pronged plug — the phosphotyrosine and isoleucine side chains of the peptide — into a two-pronged socket in the SH2 domain. The two glutamate residues are bound to sites on the surface of the SH2 domain between the two sockets. Other SH2 domains bind phosphotyrosine residues surrounded by other sequences of amino acids.

[Data from G. Waksman et al., 1993, *Cell* **72**:779, PDB ID 1sps.]

As an example, consider the protein in which SH2 domains were first identified, the tyrosine kinase called *Src* (Src is an acronym for a *sarcoma* tumor). *A* mutant form of the *src* gene, which inappropriately activates a downstream signal transduction pathway, was originally found incorporated in the *Rous sarcoma virus* that in chickens causes sarcomas (a muscle tumor, as Chapter 25

details). The SH2 domain of the Src tyrosine kinase binds strongly to any peptide containing the four-residue sequence: phosphotyrosine–glutamic acid–glutamic acid–isoleucine (single letter code: pY-E-E-I, see Figure 16-9). These four amino acids make intimate contact with the peptide-binding site in the Src SH2 domain. Binding resembles the insertion of a two-pronged plug — the phosphotyrosine and isoleucine side chains of the peptide — into a two-pronged socket in the SH2 domain. The two glutamic acids fit snugly onto the surface of the SH2 domain between the phosphotyrosine binding socket and the hydrophobic socket that binds the isoleucine residue. The specificity of SH2 domains for their target sequence plays an important role in determining which SH2-domain-containing signal-transducing proteins bind to which receptors and thus which signal transduction pathways are activated.

Receptor-Mediated Endocytosis and Lysosomal Degradation Squelch Signaling from RTKs

In Chapter 15 we emphasized that, to prevent overstimulation, cells must be able to suppress or even turn off their signal transduction pathways. In addition, once the cell has adequately responded to the signal, or when the signal has been removed, it is important to prevent any further cellular response. Suppression of signaling from

RTKs is common, and different mechanisms have evolved to accomplish this task.

For example, phosphotyrosine phosphatases continuously hydrolyze the bonds linking phosphates to tyrosine residues on the receptor, reducing the ability of activated RTKs to bind proteins containing phosphotyrosine-binding domains and thus diminishing the activation of signal transduction pathways. A second mechanism — receptor-mediated endocytosis followed by receptor degradation in lysosomes — is another common method for dampening RTK signal transduction.

Treatment of cells with ligand often reduces the number of available cell-surface receptors, so that the cells have a less robust response to continuous exposure to a given concentration of ligand than they did before ligand addition. This *desensitization response* helps prevent inappropriately prolonged receptor activity. In the absence of EGF, for instance, cell-surface HER1 receptors are relatively long-lived, with an average half-life of 10–15 hours. Unbound receptors are internalized via clathrin-coated pits into endosomes at a relatively slow rate, on average once every 30 minutes, and are often recycled rapidly to the plasma membrane so that in the absence of ligands there is little reduction in the total number of surface receptors. Following binding of an EGF ligand, the rate of endocytosis of HER1 is increased about tenfold, and only a fraction of the internalized receptors return to the plasma membrane; the rest are degraded in

lysosomes. Each time a HER1–EGF complex is internalized through endocytosis (see Figure 14-29), the receptor has about a 20–80 percent chance of being degraded, depending on the cell type. In epithelial cells exposed to high levels of EGF, for example, almost all cell-surface EGF receptor molecules are internalized and degraded, markedly reducing the cell's sensitivity to EGF. In this way, prolonged treatment with EGF desensitizes the cell to that hormone. If the EGF is removed, newly synthesized receptors will eventually replenish the cell surface and allow the cell to respond to a second treatment of EGF.

Several processes influence whether surface receptors internalized in endosomes are recycled to the plasma membrane or transported to lysosomes for degradation. One is covalent modification by the small protein ubiquitin (see Chapter 3). The enzyme c-Cbl adds a single ubiquitin to a given lysine of a protein, a process called *monoubiquitinylation*. There is a strong correlation between the extent of monoubiquitinylation of the HER1 cytosolic domain and HER1 degradation; such monoubiquitinylation also occurs to other ligand-activated RTKs. The enzyme c-Cbl, which is an E3 ubiquitin ligase (see Figure 3-32), contains a domain that binds directly to phosphorylated EGF receptors, and a RING finger domain, which recruits ubiquitin-conjugating enzymes and mediates transfer of ubiquitin to the receptor. The ubiquitin functions as a tag on the receptor that stimulates its incorporation from endosomes into multivesicular bodies (see Figure 14-32) that are ultimately engulfed

by and degraded in lysosomes. A role for c-Cbl in EGF receptor trafficking emerged from genetic studies in *C. elegans*, which established that c-Cbl negatively regulates the function of the nematode EGF receptor (Let-23), probably by inducing its degradation. Similarly, knockout mice lacking c-Cbl show hyperproliferation of mammary gland epithelia, consistent with a role of c-Cbl as a negative regulator of EGF signaling.

Interestingly, the other EGF family receptors — HER2, HER3, and HER4 — do not undergo internalization after exposure to ligand, an observation that emphasizes how each receptor evolved to be regulated in its own appropriate manner.

KEY CONCEPTS OF SECTION 16.1

Growth Factors and Their Receptor Tyrosine Kinases

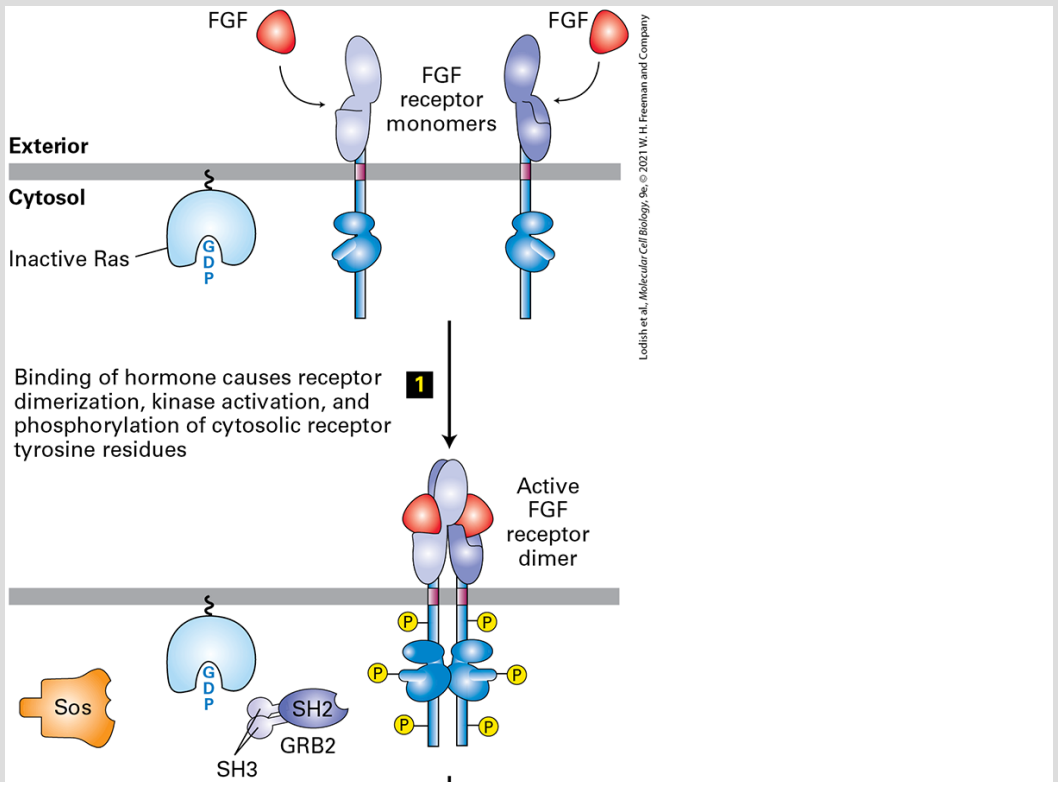
- Receptor tyrosine kinases, which bind to peptides and signaling proteins such as growth factors and insulin, may exist as preformed dimers or dimerize when ligands bind. Ligand binding triggers conformational changes leading to formation of functional dimeric receptors, a necessary step in activating the receptor's kinase domain. Different receptors accomplish this function in different ways (see Figures 16-3 to 16-7).
- Activation of many RTKs leads to phosphorylation of the activation loop in the protein tyrosine kinase that is an intrinsic part of the cytoplasmic domain, enhancing its catalytic activity (see Figure 16-3). The activated kinase then phosphorylates tyrosine residues in the receptor's cytosolic domain and in other protein substrates.
- Humans express many RTKs, four of which (HER1–HER4) define the family of epidermal growth factor receptors that mediates signaling from the many members of the epidermal growth factor family of signaling molecules (see Figure 16-7). One of these receptors, HER2, does not bind ligand; it forms active heterodimers with ligand-

bound monomers of the other three HER proteins. Overexpression of HER2 is implicated in many breast cancers.

- In RTKs, short amino acid sequences containing a phosphotyrosine residue are bound by signal-transducing proteins with conserved SH2 domains. The sequence of amino acids surrounding the phosphorylated tyrosine determines which SH2 domain will bind to it (see Figure 16-9). Such protein-protein interactions are important in determining specificity in many signaling pathways (see Figure 16-3).
- Endocytosis of receptor-ligand complexes and their degradation in lysosomes is a principal way of reducing the number of receptor tyrosine kinases on the cell surface and thus decreasing the sensitivity of cells to many peptide and protein hormones. Other receptors, such as cytokine receptors, are also regulated by endocytosis and degradation.

16.2 The Ras/MAP Kinase Signal Transduction Pathway

Signal transduction by RTKs begins with the binding of a protein with an SH2 domain to a phosphorylated tyrosine of the receptor's cytosolic domain. After one or two additional steps that we will see shortly, virtually all receptor tyrosine kinases activate the *Ras/MAP kinase signal transduction pathway* (Figure 16-10; also see Figure 16-1a, Pathway A). This pathway is also activated by most cytokine receptors. It is conserved in evolution, as versions are found in vertebrates, invertebrates, and even in yeasts, and it plays a key role in many developmental pathways and regulates the growth and differentiation of many types of cells.



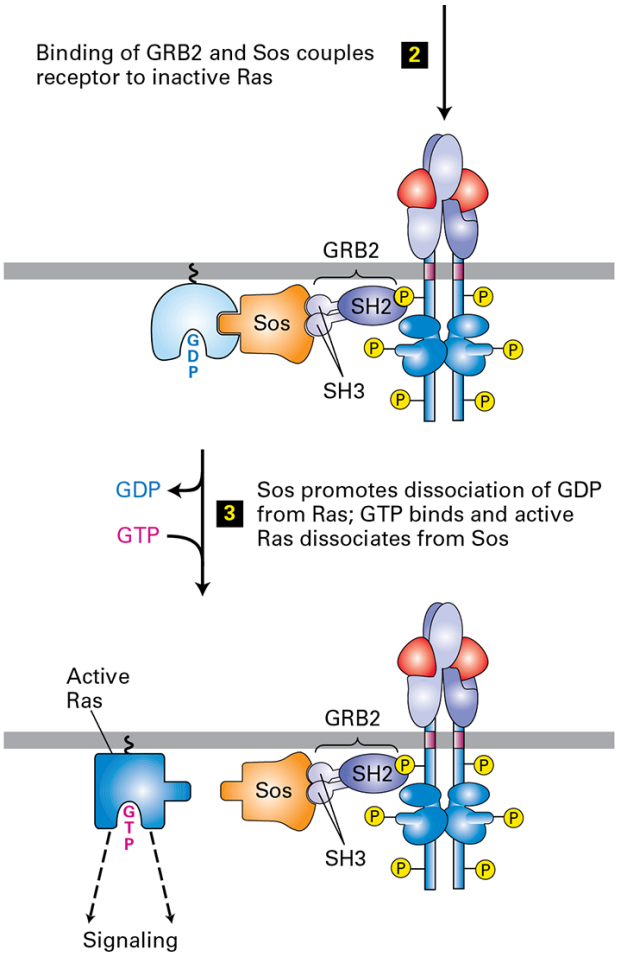


FIGURE 16-10 Activation of Ras following ligand binding to receptor tyrosine kinases (RTKs) or cytokine receptors. The receptor for fibroblast growth factor (FGF) exemplifies the activation of this signaling pathway. The SH2 domain of the cytosolic adapter protein GRB2 binds to a specific phosphotyrosine on an activated, ligand-bound receptor, and its SH3 domains bind to a cytosolic Sos protein, bringing it near the cytosolic surface of the plasma membrane and close to its substrate, the inactive Ras·GDP. The guanine nucleotide exchange factor (GEF) activity of Sos then promotes formation of active Ras·GTP from Ras·GDP. Note that Ras is tethered to the cytosolic surface of the plasma membrane by a hydrophobic farnesyl anchor (see Figure 10-19). See J. Schlessinger, 2000, *Cell* **103**:211; and M. A. Simon, 2000, *Cell* **103**:13.

The monomeric *Ras protein* belongs to the *GTPase superfamily* of intracellular switch proteins (see Figure 15-7). Activated Ras promotes the formation of a signal transduction complex, containing three sequentially acting protein kinases, at the cytosolic surface of the plasma membrane. This *kinase cascade* culminates in activation of certain members of the **MAP kinase** family, which can translocate into the nucleus and phosphorylate many different proteins. Among the target proteins for MAP kinase are transcription factors that regulate the expression of proteins with important roles in the cell cycle and in cell differentiation. Importantly, different hormones and their cell-surface receptors often differ in the strength or duration of the activation of a MAP kinase and thus in their effects on the cell. An activating mutation in an RTK, Ras, or a protein in the MAP kinase pathway that induces inappropriate cell division is found in almost all types of human tumors. Thus the RTK/Ras/MAP kinase signal transduction pathway has been studied extensively, and a great deal is known about its component proteins and their regulation (see Chapter 25).

We begin our discussion of the Ras/MAP kinase signal transduction pathway by reviewing how Ras cycles between the active GTP-bound and inactive GDP-bound states. We then describe how activated Ras passes a signal to the MAP kinase. Finally, we examine recent studies indicating that both yeasts and cells of higher eukaryotes contain multiple MAP kinase pathways, and we consider the ways in

which cells keep different MAP kinase pathways separate from one another through the use of scaffold proteins.

Ras, a GTPase Switch Protein, Operates Downstream of Most RTKs and Cytokine Receptors

Like the other G proteins, including the G_α subunits in trimeric G proteins that were discussed in Chapter 15, monomeric Ras proteins are G protein switches that alternate between an inactive "off" state with a bound GDP and a receptor-activated "on" state with a bound GTP (see Figure 15-7). Trimeric G proteins are activated upon binding directly to a ligand-bound GPCR that acts as a *guanine nucleotide exchange factor (GEF)* that triggers the release of the bound GDP. GTP then binds spontaneously to the nucleotide-free G protein. Activated RTKs and cytokine receptors are not GEFs; rather, the activated RTK or cytokine receptor recruits several adaptor proteins to the cytosolic surface of the plasma membrane, and one of these acts as the GEF that activates Ras.

Ras (~170 amino acids) is smaller than G_α proteins (~300 amino acids), but the GTP-binding domains of the two protein types have similar structures (see Figure 15-7 to review the structure of Ras). As with G_α proteins, hydrolysis of the bound GTP to GDP deactivates Ras. G_α proteins contain a GTPase-activating protein (GAP) domain

that increases the intrinsic rate of GTP hydrolysis, but this domain is not present in Ras. As a result, Ras has an intrinsically slower rate of GTP hydrolysis. Thus the average lifetime of a GTP bound to Ras is about 1 minute, much longer than the average lifetime of a G_α·GTP complex. Because of its low intrinsic GTPase activity, Ras·GTP requires the assistance of another protein, a GTPase-activating protein (GAP), to become deactivated. Binding of a GAP to Ras·GTP accelerates the intrinsic GTPase activity of Ras by more than a hundredfold; the actual hydrolysis of GTP to GDP and P_i is catalyzed by amino acids from both Ras and the GAP. The GAP for RAS (called a RAS-GAP) works by inserting one of its arginine side chains into the Ras active site, thus stabilizing an intermediate in the hydrolysis reaction.

As noted, mammalian Ras proteins have been studied in great detail because mutant Ras proteins are associated with many types of human cancer. These mutant proteins, which bind but cannot hydrolyze GTP, are permanently in the "on" state and contribute to the cellular transformation that leads to cancer (see Chapter 25). Determination of the three-dimensional structure of the Ras-GAP complex and tests of mutant forms of Ras explained the puzzling observation that most oncogenic, constitutively active Ras proteins (Ras^D) contain a mutation at position 12. Replacement of the normal glycine 12 with any other amino acid (except proline) blocks the functional binding of GAP proteins, vastly decreases the rate of

GTP hydrolysis, and in essence locks Ras in the active GTP-bound state. ■

The first indication that Ras functions downstream from RTKs in a common signaling pathway came from experiments on cultured fibroblast cells. These cells were induced to proliferate by treatment with a mixture of two protein hormones that activate RTKs: PDGF and EGF. Microinjection of anti-Ras antibodies into these cells blocked cell proliferation. Conversely, injection of Ras^{D}, a constitutively active mutant Ras protein, caused the cells to proliferate in the absence of the growth factors. These findings are consistent with studies using the pull-down assay method detailed in Figure 15-11 in which addition of FGF to fibroblasts led to a rapid increase in the proportion of Ras present in the GTP-bound active form.

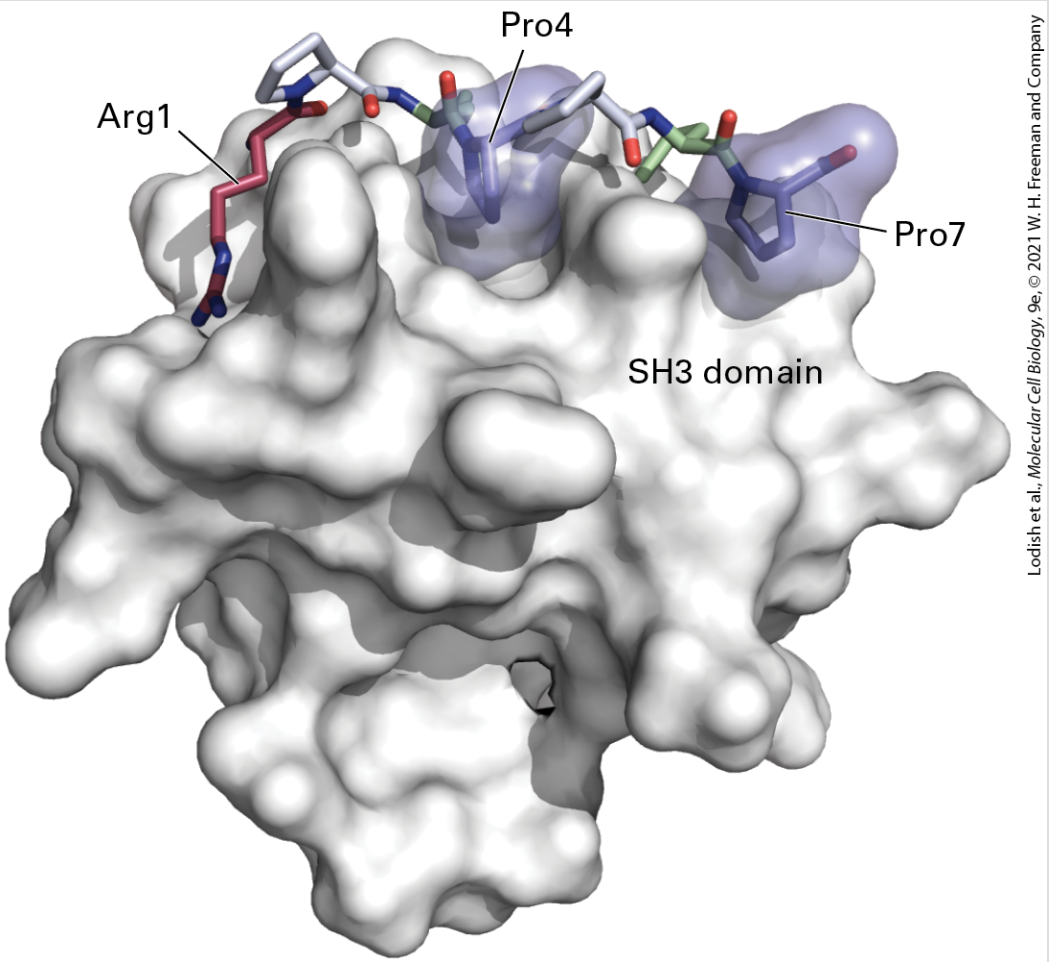
Receptor Tyrosine Kinases Are Linked to Ras by Adapter Proteins

In order for activated RTKs and cytokine receptors to activate Ras, two cytosolic proteins — GRB2 and Sos — must first be recruited to provide a link between the receptor and Ras (see Figure 16-10). GRB2 is an *adapter protein*, meaning that it has no enzyme activity and serves as a link, or scaffold, between two other proteins — in this case, between the activated receptor and Sos. Sos is a GEF that

catalyzes conversion of inactive GDP-bound Ras to the active GTP-bound form. (The name *Sos* is an acronym for *Son of Sevenless*, a gene defined by analysis of *Drosophila* mutants with defects in eye development, which is critically dependent on a Ras-mediated signaling pathway.)

GRB2 is able to serve as an adapter protein because it has an SH2 domain, which binds to specific phosphotyrosine residues in the cytosolic domains of many activated RTKs. In addition to its SH2 domain, the GRB2 adapter protein contains two *SH3 domains*, which bind to Sos (see Figure 16-10). Like phosphotyrosine-binding SH2 domains, SH3 domains are present in a large number of proteins involved in intracellular signaling. Although the three-dimensional structures of various SH3 domains are similar, their amino acid sequences differ. The SH3 domains in GRB2 selectively bind to proline-rich sequences in Sos; different SH3 domains in other proteins bind to proline-rich target sequences distinct from those in Sos. Proline residues play two roles in the interaction between an SH3 domain in an adapter protein (e.g., GRB2) and a proline-rich sequence in its corresponding target protein (e.g., Sos). First, the proline-rich sequence assumes an extended conformation that permits extensive contacts with the SH3 domain, thereby facilitating interactions. Second, a subset of the prolines fit into binding pockets on the surface of the SH3 domain (Figure 16-11). Several nonproline residues also interact with the SH3 domain and are responsible for determining which SH3 domain binds to which target protein

(binding specificity). Hence the binding of proteins to SH3 and to SH2 domains follows a similar strategy: certain residues provide the key structural motif necessary for binding (phosphotyrosine for binding to SH2 domains, prolines for binding to SH3 domains), and neighboring residues confer specificity to the binding.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 16-11 Surface model of an SH3 domain bound to a target peptide. The short, proline-rich target peptide is shown as a stick model and the SH3 domain as a space-filling model. In this target peptide, two prolines (Pro4 and Pro7, dark blue) fit into binding pockets on the surface of the SH3 domain. Interactions involving an arginine (Arg1, red), two other prolines (gray), and other residues in the target peptide (green) determine the specificity of binding of the target peptide to an SH3 domain.

[Data from S. Feng et al., 1995, *Proc. Nat'l. Acad. Sci. USA* **92**:12408, PDB ID 1qwf.]

Binding of Sos to Inactive Ras Causes a Conformational Change That Triggers an Exchange of GTP for GDP

Following activation and autophosphorylation of an RTK (e.g., the FGF receptor) and binding of GRB2, Sos binds to GRB2, forming a complex on the cytosolic face of the plasma membrane (see Figure 16-10). The formation of this complex depends on the ability of GRB2 to bind simultaneously to the activated receptor and to Sos. Thus on receptor activation, Sos relocates from the cytosol to the membrane, bringing Sos close to its substrate, Ras·GDP, which is already bound to the cytosolic surface of the plasma membrane by means of a covalently attached lipid.

Binding of Sos to Ras·GDP activates Ras·GDP by releasing its GDP so that GTP can replace it. The binding of Sos leads to conformational changes in the switch I and switch II segments of Ras, thereby permitting GDP to diffuse away (Figure 16-12). In other words, Sos functions as a GEF for Ras. Binding of GTP to Ras, driven by the high GTP concentration in the cytosol relative to that of GDP, induces a specific conformation of switch I and switch II that allows Ras·GTP to activate Raf, the next protein in the Ras/MAP kinase pathway.

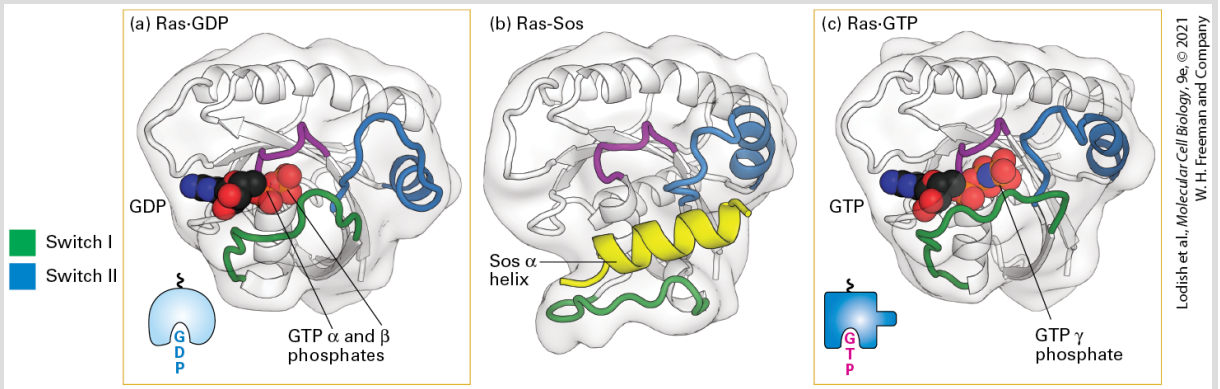


FIGURE 16-12 Structures of Ras bound to GDP, Sos protein, and GTP. (a) As with other G proteins bound to GDP, in Ras·GDP, the switch I (green) and switch II (blue) segments do not directly interact with GDP. (b) One α helix (yellow) in Sos binds to both switch segments of Ras·GDP, leading to a massive conformational change in Ras. In effect, Sos pries Ras open by displacing the switch I region, thereby allowing GDP to diffuse out. (c) GTP is thought to bind to the Ras-Sos complex first through its base (guanine); subsequent binding of the GTP phosphates completes the interaction. The resulting conformational change in the switch I and switch II segments of Ras, allowing both to bind to the GTP γ phosphate, displaces Sos and promotes interaction of Ras·GTP with its effectors (discussed later). The P loop (purple) is a sequence motif found in many ATP- and GTP-binding proteins that binds the β phosphate of the nucleotide. See Figure 15-7 for another depiction of Ras·GDP and Ras·GTP.

[Part (a) Data from M. V. Milburn et al., 1990, *Science* **247**:939, PDB ID 4q21. Part (b) Data from J. Sejbal et al., 1996, *J. Med. Chem.* **39**:1281, PDB ID 1bdk. Part (c) Data from M. E. Pacold et al., 2000, *Cell* **103**:931, PDB ID 1he8.]

Signals Pass from Activated Ras to a Cascade of Protein Kinases Ending with MAP Kinase

Biochemical and genetic studies in yeast, *C. elegans, Drosophila*, and mammals have revealed that downstream of Ras is a highly

conserved cascade of three protein kinases, culminating in MAP kinase. Although activation of the kinase cascade does not yield the same biological results in all cells, all Ras/MAP kinase pathways have a common set of three sequentially acting kinases, as outlined in Figure 16-13: Ras → Raf → MEK → MAP kinase.

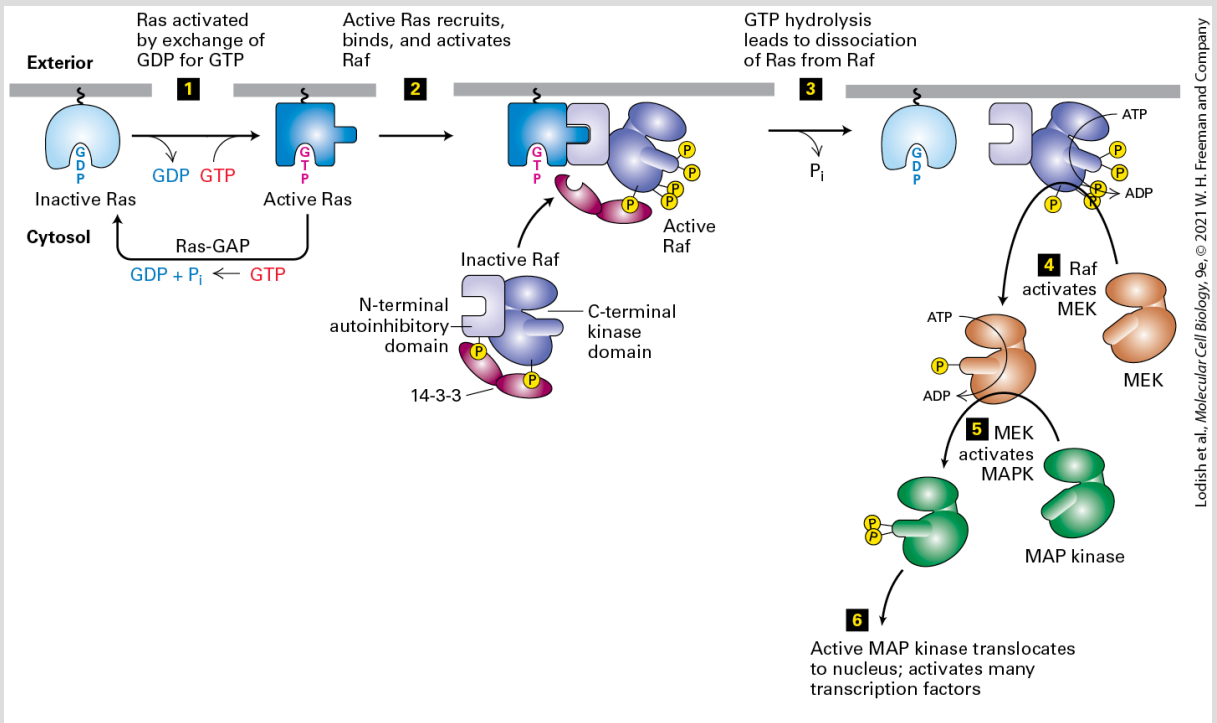


FIGURE 16-13 Ras/MAP kinase pathway. In unstimulated cells, most Ras is tethered to the cytosolic surface of the plasma membrane in the inactive form with bound GDP. Binding of a ligand to its RTK or cytokine receptor leads to formation of the active Ras·GTP complex (step 1; see also Figure 16-12). Activated Ras triggers the activation of the Raf kinase, activating the kinase cascade depicted in steps 2–6 and culminating in activation of MAP kinase (MAPK). In unstimulated cells, binding of a dimer of the 14-3-3 protein to Raf stabilizes its kinase domain in an inactive conformation. Each 14-3-3 monomer binds to a phosphoserine residue in Raf, one to phosphoserine-259 in the N-terminal domain and the other to phosphoserine-621 in the kinase domain; binding to 14-3-3 maintains the kinase in a closed, inactive state in the cytosol. Step 2: Binding of the Raf N-terminal regulatory domain to Ras·GTP results in dephosphorylation by a cytosolic enzyme of one of the serines that bind Raf to 14-3-3, leading to the consequent loss of 14-3-3 binding and partial activation of Raf

kinase activity. Maximum activation of Raf kinase activity occurs in part by formation of Raf dimers (not depicted), in which the kinase domain of one phosphorylates serine or threonine residues in the activation loop of the other. Step **3**: After the GTP bound to Ras is hydrolyzed to GDP, a process accelerated by a Ras-GAP, the inactive Ras·GDP dissociates from Raf. The Ras·GDP can be reactivated by signals from activated receptors, thereby recruiting additional Raf molecules to the membrane. As detailed in the text, in step **4** Raf phosphorylates and thus activates a MEK kinase, and in step **5** MEK phosphorylates and thus activates a MAP kinase. See E. Kerkhoff and U. Rapp, 2001, *Adv. Enzyme Regul.* **41**:261; J. Avruch et al., 2001, *Recent Prog. Horm. Res.* **56**:127; and D. Matallanas et al., 2011, *Genes Cancer* **2**:232.

In unstimulated cells, Raf is phosphorylated by several cytosolic kinases on two sites, one in its serine/threonine kinase domain and a second in its N-terminal autoinhibitory domain. These phosphate groups then bind to a dimer of the protein 14-3-3, which locks the autoinhibitory domain in a position where it blocks and inhibits the kinase active site — a *molecular handcuff*. Binding to 14-3-3 also retains the inactive Raf protein in the cytosol.

As noted, Ras is activated by the exchange of GDP for GTP (Figure 16-13, step **1**). Active Ras·GTP (but not Ras·GDP) binds to the N-terminal autoinhibitory domain of Raf, altering its conformation and allowing the phosphate attached to this domain to be removed by cytosolic phosphatases, thereby causing loss of 14-3-3 binding. This unblocks the kinase active site and partially activates its kinase activity (step **2**). Hydrolysis of Ras·GTP to Ras·GDP releases Raf from Ras (step **3**), and the released Raf then forms dimers in which the kinase activity of one phosphorylates serine or threonine

residues on the activation loop of the other, further increasing Raf kinase activity. In summary, binding of Ras·GTP to Raf, followed by hydrolysis of Ras·GTP to Ras·GDP, releases active Raf from its inhibitory complex with 14-3-3 (step **3**) and leads to activation of Raf kinase activity.

The now active Raf subsequently phosphorylates *MEK* on one or two serine residues in its activation loop, thereby activating MEK kinase activity (step **4**). MEK is a kinase that phosphorylates its target proteins on both tyrosine and serine/threonine residues; hence it is called a *dual-specificity* protein kinase. Active MEK mainly phosphorylates MAP kinase in its activation loop (step **5**). Active MAP kinase phosphorylates many different proteins, including nuclear transcription factors that mediate cellular responses that are discussed later (step **6**).

Several types of experiments have established the $\mathrm{Ras} \rightarrow \mathrm{Raf} \rightarrow \mathrm{MEK} \rightarrow \mathrm{MAP}$ kinase pathway. For example, cultured mammalian cells that express a mutant, nonfunctional Raf protein cannot be stimulated to proliferate uncontrollably by a constitutively active $\mathrm{Ras}^{\mathrm{D}}$ protein, one that has a permanently bound GTP. This finding established a link between the Raf and Ras proteins and showed that Raf lies downstream of Ras in the signaling pathway. In vitro binding studies further showed that the purified Ras·GTP protein binds directly to the N-terminal regulatory domain of Raf and activates its catalytic activity.

That MAP kinase is activated in response to Ras activation was demonstrated in quiescent (nondividing) cultured cells expressing a constitutively active Ras^D protein. In these cells, activated MAP kinase is generated in the absence of stimulation by growth-promoting hormones. Biochemical studies showed, however, that Raf cannot directly phosphorylate MAP kinase or otherwise stimulate its activity. The final link in the kinase cascade activated by Ras·GTP emerged from studies in which scientists fractionated extracts of cultured cells to search for a kinase activity that could phosphorylate MAP kinase and that was active only in cells stimulated with growth factors, not in unstimulated cells. This work led to the identification of MEK. Phosphorylation promotes not only the catalytic activity of MAP kinase, but also its dimerization. The dimeric form of MAP kinase is translocated to the nucleus, where it regulates the activity of many nuclear transcription factors. Later studies showed that MEK binds to the C-terminal catalytic domain of Raf and is phosphorylated on one or two serine residues in its activation loop by the Raf serine/threonine kinase (see Figure 16-13, step **4**); this phosphorylation activates the catalytic activity of MEK.

Given that the Ras/MAP kinase pathway is used in many cells and activated by many different receptors, it is not surprising that there are multiple isoforms of each of its components. In humans, there are three RAS, three Raf, two MEK, and two MAP proteins, and the isoforms have overlapping but also nonredundant functions.

Because Raf phosphorylates MEK, Raf and analogous MEK phosphorylating kinases in cells are often called *MEK kinases*, or *MEKK*. All eukaryotes including mammals possess multiple highly conserved Ras/MAP kinase pathways that are activated by different extracellular signals and that activate different MAP kinase proteins that phosphorylate different transcription factors; these transcription factors, in turn, trigger different changes in cell division, differentiation, or function. Mammalian MAP kinases include *Jun N-terminal kinases (JNKs)* and *p38 kinases*, which are activated by signaling pathways in response to various types of stresses and which phosphorylate various transcription factors and other types of signaling proteins that affect cell division.

Mutations in the *Raf* gene that lock its kinase domain in a permanently activated conformation occur in many types of tumors, in particular in about 50 percent of melanomas, skin cancers that are often caused by exposure to the ultraviolet radiation in sunlight. One such mutation, a glutamic acid substitution for the valine at position 600, occurs in over 90 percent of melanomas caused by exposure to UV radiation. This mutant constitutively active *Raf* kinase stimulates MAP-kinase signaling in cells in the absence of growth factors, promoting cell proliferation and preventing apoptosis (programmed cell death; see Chapter 22). Selective inhibitors of this mutant form of Raf have recently entered the clinic and generally produce short-term remission of the cancer. A combination of drugs that inhibit Raf and MEK appear more

promising and have recently entered clinical trials in patients with melanomas triggered by this mutant Raf. ■

MAP Kinase Regulates the Activity of Many Transcription Factors Controlling Early Response Genes

Addition of a growth factor such as EGF or PDGF to quiescent, nondividing cultured mammalian cells causes a rapid increase in the expression of as many as a hundred different genes. These genes are called *early response genes* because they are induced well before cells, in the G_0 or G_1 phase of the cell cycle, enter the S phase and replicate their DNA (see Chapter 19); many are turned on downstream of activated (phosphorylated and dimeric) MAP kinase (Figure 16-14). Consider the master transcription factor c-Fos. Together with other transcription factors, c-Fos induces the expression of many genes that encode proteins necessary for cells to progress through the cell cycle. The enhancer that regulates c-*fos* gene expression contains a *serum response element (SRE)*, so named because it is activated by many growth factors in serum. Activated MAP kinase induces transcription of the c-*fos* gene by directly phosphorylating and thus activating one transcription factor, *ternary complex factor (TCF)*. MAP kinase also phosphorylates and activates yet another kinase, termed $p90^{RSK}$, that in turn phosphorylates another transcription factor termed *serum response factor (SRF)*.

Association of phosphorylated TCF with two molecules of phosphorylated SRF forms an active transcription factor complex that activates c-*fos* gene transcription.

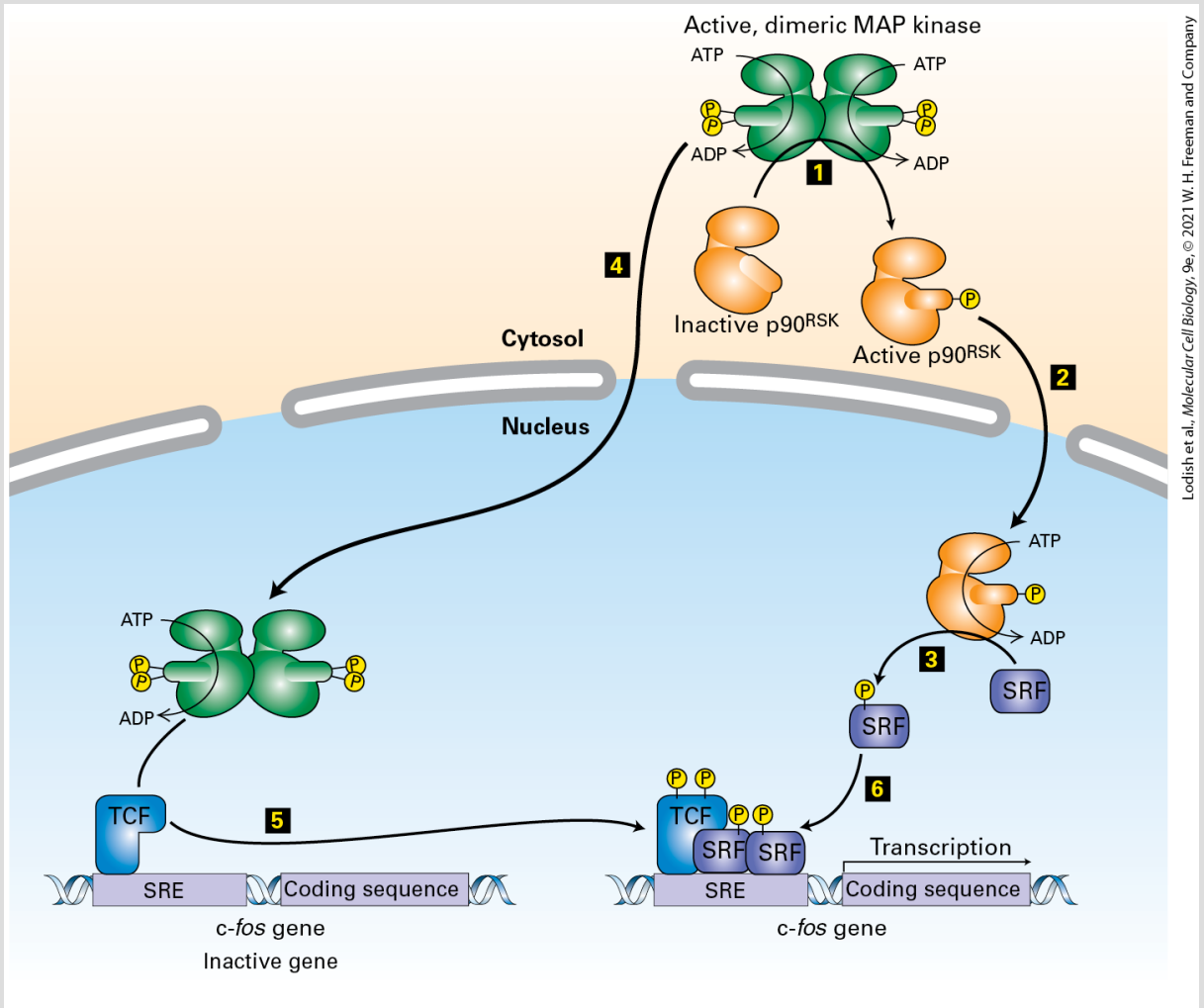


FIGURE 16-14 Induction of gene transcription by MAP kinase. Steps 1–3: In the cytosol, MAP kinase phosphorylates and activates the kinase $p90^{RSK}$, which then moves into the nucleus and phosphorylates a specific serine in the SRF transcription factor. Steps 4 and 5: After translocating into the nucleus, MAP kinase directly phosphorylates specific serines in the transcription factor TCF, which is already bound to the promoter of the c-*fos* gene. Step 6: Phosphorylated TCF and SRF act together to stimulate transcription of c-*fos* and other genes that contain an SRE sequence in their promoter. See the text for details. See R. Marais et al., 1993, *Cell* **73**:381; and V. M. Rivera et al., 1993, *Mol. Cell Biol.* **13**:6260.

Most RTKs that bind growth factors use the Ras/MAP kinase pathway to activate genes encoding proteins such as c-Fos, which in turn propel the cell through the cell cycle. Unlike the Raf and MEK kinases, which phosphorylate one or only a few target kinases in the Ras/MAP kinase cascade, MAP kinases are known to phosphorylate more than 200 proteins in the nucleus and cytosol. Many MAP kinase targets are regulators of gene expression, and different MAP kinase targets are expressed in different mammalian cell types. As with other transcription factors activated directly or indirectly by cell-surface receptors, the precise proteins induced by MAP kinase depend on the particular target proteins expressed in the cell, on epigenetic markers on DNA and chromatin proteins, and on the presence of other transcription factors (see Figure 16-2).

Multiple Feedback Mechanisms Restrain MAP Kinase Activation

Prolonged activation of the RTK–Ras/MAP kinase cascade may lead to inappropriate cell proliferation that would be deleterious to the organism. Accordingly, to maintain proper normal cell regulation, MAP kinase activation is tightly regulated through several feedback mechanisms. As with feedback control of other cellular pathways, these fall into two main categories: post-translational modification of pathway components and induction of synthesis of proteins that inhibit specific steps in the pathway. Inhibition of the Ras/MAP

kinase pathway by direct post-translational modification is nearly instantaneous, while many minutes are required to produce inhibitory protein(s) by de novo gene expression and protein synthesis and for these inhibitory proteins to influence their targets.

At the post-translational level, activated MAP kinase phosphorylates multiple proteins that act earlier in the pathway, reducing their ability to activate the MAP kinase pathway. For example, activated MAP kinase phosphorylates the HER1 EGF receptor on a threonine residue near the juxtamembrane domain (see Figure 16-8). Phosphorylation of this threonine residue decreases the ability of the receptor kinase to become fully active. Active MAP kinase also phosphorylates Raf on multiple sites, reducing its association with Ras and thus its ability to activate Ras. Such regulation effectively limits both the intensity and duration of MAP kinase activation.

Activation of the Ras/MAP kinase pathway also promotes increased transcription of regulators that lower the intensity and duration of MAP kinase signaling, a classic example of regulation by negative feedback. For example, activated MAP kinase triggers an increase in the expression of *du*al-*s*pecificity *p*hosphatases (DUSPs), particularly DUSP6. Their name is derived from their ability to dephosphorylate both the threonine and tyrosine phosphorylation sites in the MAP kinase activation loop, causing the loop to block the MAP kinase active site. DUSP6, along with several other negative regulators of the RTK–Ras/MAP kinase signaling cascade, are among the early

response genes whose transcription is increased rapidly — within the first 15 minutes — following MAP kinase activation. Increased expression of these phosphatases typically leads to transient inhibition of MAP kinase and can also suppress the response to a subsequent stimulation.

Scaffold Proteins Isolate Distinct MAP Kinase Pathways in the Same Cell from One Another

Two distinct signal transduction pathways within the same cell, initiated by two distinct receptors and having two distinct results, may share the identical protein kinase at some step. Yet this kinase will phosphorylate a different substrate depending on the pathway, even though both potential substrates are present in the cell. The principal mechanism that ensures that the correct substrate is phosphorylated is the formation of a complex by a scaffold protein. An example from yeast illustrates how the action of a scaffold protein guides the kinase to the correct substrate.

Yeast cells, like mammalian cells, utilize several MAP kinase pathways (Figure 16-15), but they are not activated by Ras-like proteins (which do not exist in yeasts). One pathway is activated by placing the cells in medium of high osmotic strength. The MAP kinase in this pathway, Hog1, phosphorylates transcription factors

that in turn activate genes whose proteins protect the cell against this osmotic challenge (Figure 16-15b). A second and distinct pathway is activated during mating of yeast cells of opposite mating type. Haploid yeast cells are either of the **a** or α mating type. An **a** haploid cell secretes the **a** mating factor and has cell-surface GPCRs that bind the α factor; an α cell secretes the α factor and has GPCRs that bind the **a** factor (see Figure 1-24). GPCRs that bind these factors initiate a MAP kinase pathway; the MAP kinase in this pathway, Fus3, translocates into the nucleus. There it phosphorylates and activates several transcription factors that turn on genes whose encoded proteins inhibit progression of the cell cycle and other genes that enable cells of opposite mating type to fuse together and ultimately form a diploid cell (Figure 16-15a).

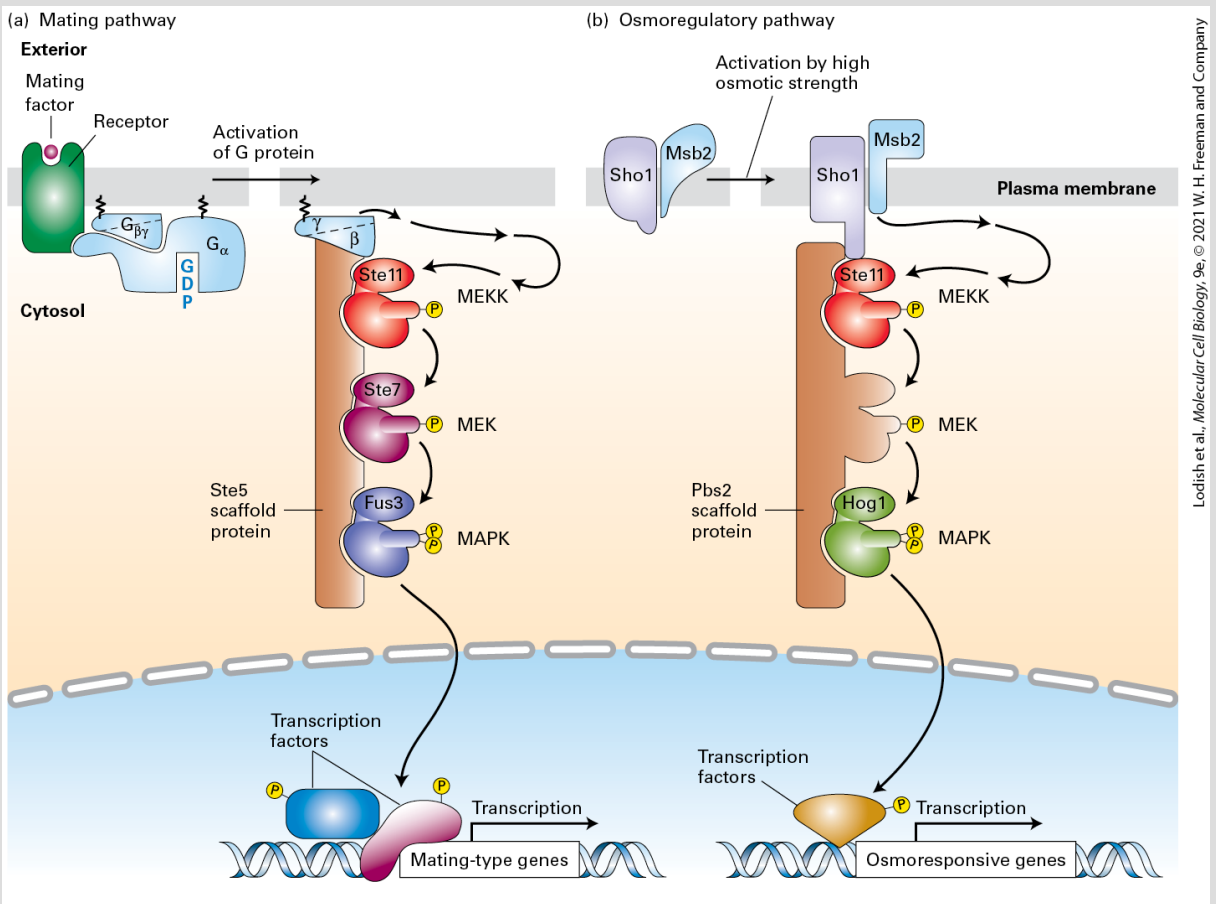


FIGURE 16-15 Scaffold proteins separate yeast MAP kinase cascades in the mating and osmoregulatory pathways. In the budding yeast *Saccharomyces cerevisiae*, different cell-surface receptors activate different MAP kinase pathways, two of which are outlined here. (a) Mating pathway: The GPCR receptors for yeast α and **a** mating factors are coupled to the same trimeric G_α protein. Following ligand binding and dissociation of the G protein subunits, the released membrane-tethered $G_{\beta\gamma}$ subunit binds the Ste5 scaffold to the plasma membrane. $G_{\beta\gamma}$ then activates a signal transduction pathway (arrows) that culminates with the phosphorylation and thus activation of Ste11, which is analogous to Raf and other mammalian MEK kinase (MEKK) proteins. Ste11 initiates a MAP kinase cascade in which the final component, Fus3, is functionally equivalent to the MAP kinases (MAPK) in higher eukaryotes. Like other MAP kinases, activated Fus3 then translocates into the nucleus where it leads to phosphorylation and activation of several transcription factors — in this case ones that initiate transcription of genes that enable cells of opposite mating type to fuse together and ultimately form a diploid cell. (b) Osmoregulatory pathway: Two plasma membrane proteins, Sho1 and Msb1, are activated by exposure of yeast cells to media of high osmotic strength. Activated Sho1 recruits the Pbs2 scaffold protein, which contains a

MEK domain, to the plasma membrane, and activates a signal transduction pathway (arrows) that also results in phosphorylation and activation of Ste11. This initiates a different MAP kinase cascade that activates Hog1, a MAP kinase. After translocating to the nucleus, Hog1 phosphorylates several transcription factors and chromatin-modifying enzymes that induce transcription of genes that support survival in high-osmotic-strength media. Note that the MEK in this pathway is an integral domain of the Ste5 scaffold. See N. Dard and M. Peter, 2006, *BioEssays* **28**:146; and R. Chen and J. Thorner, 2007, *Biochim. Biophys. Acta* **1773**:1311.

There is a potential source of signaling confusion because, as Figure 16-15 shows, the same MEKK protein kinase, Ste11, is found in both pathways but phosphorylates different MEK kinases that are unique to each pathway. **Scaffold proteins** prevent Ste11 from phosphorylating the wrong MEK kinase. The scaffold protein Ste5, employed in the mating pathway, has no obvious catalytic function but stabilizes a large complex that includes the MEKK kinase Ste11, the MEK kinase Ste7, and the MAP kinase Fus3 in the mating signal transduction pathway. By binding to the $G_{\beta\gamma}$ protein that is released following activation of the mating factor GPCR, Ste5 ensures that the Ste11 bound to it is activated by the mating pathway. In contrast, the receptors that sense changes in extracellular osmotic strength, Sho1 and Msb2, cannot activate the MEKK Ste11 in the mating pathway complex.

Similarly, the Pbs2 scaffold protein is employed in the kinase cascade in the osmoregulatory pathway (see Figure 16-15) and binds to Sho1, an activator of the osmolarity pathway. This ensures that the Ste11 bound to it is activated only by the osmotic pathway; the

$G_{\beta\gamma}$ protein released following activation of the mating factor GPCR cannot activate the MEKK Ste11 in this osmo-sensing pathway complex.

Thus signaling downstream from Ste11 is restricted to the complex in which it is localized. As a result, exposure of yeast cells to mating factors induces activation of a single MAP kinase, Fus3, whereas exposure to a high osmolarity induces activation of a different MAP kinase, Hog1.

As emphasized in Figure 3-31, scaffolds also increase the rate of enzyme–catalyzed reactions by bringing the enzyme (e.g., a MEKK) close together to its substrate (a MEK), reducing the time needed for an enzyme to diffuse through an aqueous solution and find a substrate molecule.

Scaffolds for MAP kinase pathways are well documented in yeast, fly, worm, and mammalian cells. In addition to their ability to function as insulators for specific pathways, as described above for yeast Fus3 and Hog1, scaffold proteins can also regulate pathway activity. IQGAP1 is a mammalian scaffold protein that binds to RAF, MEK, and ERK kinases, as well as to a number of other proteins. Expression of IQGAP1 is increased in multiple different cancer types, and when its expression increases, so does the activity of the Ras–MAP kinase signaling cascade. Presumably IQGAP1 acts by bringing Raf, MEK, and ERK together into a single complex and thus facilitating their activation. Intriguingly, expressing in cells just the

domain of IQGAP that binds ERK is sufficient to inhibit the growth of melanoma cells that express a mutant, activated Raf protein. This small protein blocks the binding of ERK to the full-length IQGAP1 scaffold, thus preventing ERK's incorporation into the functional scaffold complex. This function of IQGAP suggests an alternate strategy for reducing MAP kinase activity in cancers. The signal specificity of the different MAP kinases in mammalian cells may arise from their association with different scaffold-like proteins, but much additional research is needed to test this possibility.

KEY CONCEPTS OF SECTION 16.2

The Ras/MAP Kinase Signal Transduction Pathway

- Ras is an intracellular GTPase switch protein that acts downstream from most RTKs and cytokine receptors. Like G_α, Ras cycles between an inactive GDP-bound form and an active GTP-bound form. Ras cycling requires the assistance of two proteins: a guanine nucleotide exchange factor (GEF) and a GTPase-activating protein (GAP).
- RTKs are linked indirectly to Ras via two proteins: GRB2, an adapter protein, and Sos, which has GEF activity (see Figure 16-10).
- The SH2 domain in GRB2 binds to a phosphotyrosine in activated RTKs, while its two SH3 domains bind Sos, thereby bringing Sos, a GEF, close to membrane-bound Ras·GDP and activating it.
- Binding of Sos to inactive Ras causes a large conformational change that permits release of GDP and binding of GTP, forming active Ras (see Figure 16-12).
- Activated Ras triggers a kinase cascade in which Raf (a MEK kinase or MEKK), MEK, and MAP kinase are sequentially phosphorylated and thus activated. Activated MAP kinase then translocates to the nucleus (see Figure 16-13).
- Activation of MAP kinase following stimulation of a growth-factor receptor leads to phosphorylation and activation of two transcription factors that promote transcription of various early response genes (see Figure 16-14).
- Different extracellular signals induce activation of different MAP kinase pathways, which regulate diverse cellular processes by phosphorylating different sets of

transcription factors.

- Multiple negative feedback mechanisms restrain MAP kinase activation — some depending on proteins whose gene transcription is induced by activated MEK.
- The kinase components of MAP kinase cascades can assemble into a large pathway-specific complex often stabilized by a scaffold protein (see Figure 16-15). This ensures that activation of one MAP kinase pathway by a particular extracellular signal does not lead to inappropriate activation of other pathways containing shared components.

16.3 Phosphoinositide Signal Transduction Pathways

In previous sections, we have seen how activated receptor tyrosine kinases (RTKs) initiate the Ras/MAP kinase pathway (see Figures 16-10 and 16-13). Here we discuss how these same receptors initiate some very different signaling pathways. These pathways have as intermediates certain phosphorylated phospholipids derived from phosphatidylinositol. As discussed in Chapter 15, these membrane-bound lipids are collectively referred to as **phosphoinositides**.

These signaling pathways include several enzymes that synthesize different phosphoinositides on the cytosolic face of the plasma membrane as well as cytosolic proteins with domains that can bind to these molecules. These proteins can thus be recruited to the cytosolic surface of the plasma membrane when the corresponding phosphoinositide is generated.

Several of these phosphoinositide pathways have long-term effects on patterns of gene expression as well as short-term effects. We will see that different phosphoinositide pathways end with activation of specific kinases, including protein kinase C (PKC) and protein kinase B (PKB), that play key roles in cell growth and metabolism. As an example, we will see in Chapter 21 how insulin activation of PKB plays a key role in rapidly stimulating glucose import into muscle and fat cells.

Phospholipase C_γ Is Activated by Many RTKs and Cytokine Receptors

In Chapter 15 we learned that hormonal stimulation of certain G protein–coupled receptors leads to the activation of phospholipase C (PLC). This membrane-associated enzyme then cleaves phosphatidylinositol 4,5-bisphosphate [$PI(4,5)P_2$] to generate two important second messengers: 1,2-diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3). One of the resulting products, IP_3, triggers an increase in the concentration of Ca^{2+} in the cytosol with its many and diverse effects in different types of cells. A rise in the cytosolic Ca^{2+} level causes protein kinase C to translocate to the cytosolic surface of the plasma membrane, where it binds to DAG and becomes activated (see Figure 15-28); activation of PKC in different cells results in a varied array of cellular responses, indicating that it plays a key role in many aspects of cellular growth and metabolism.

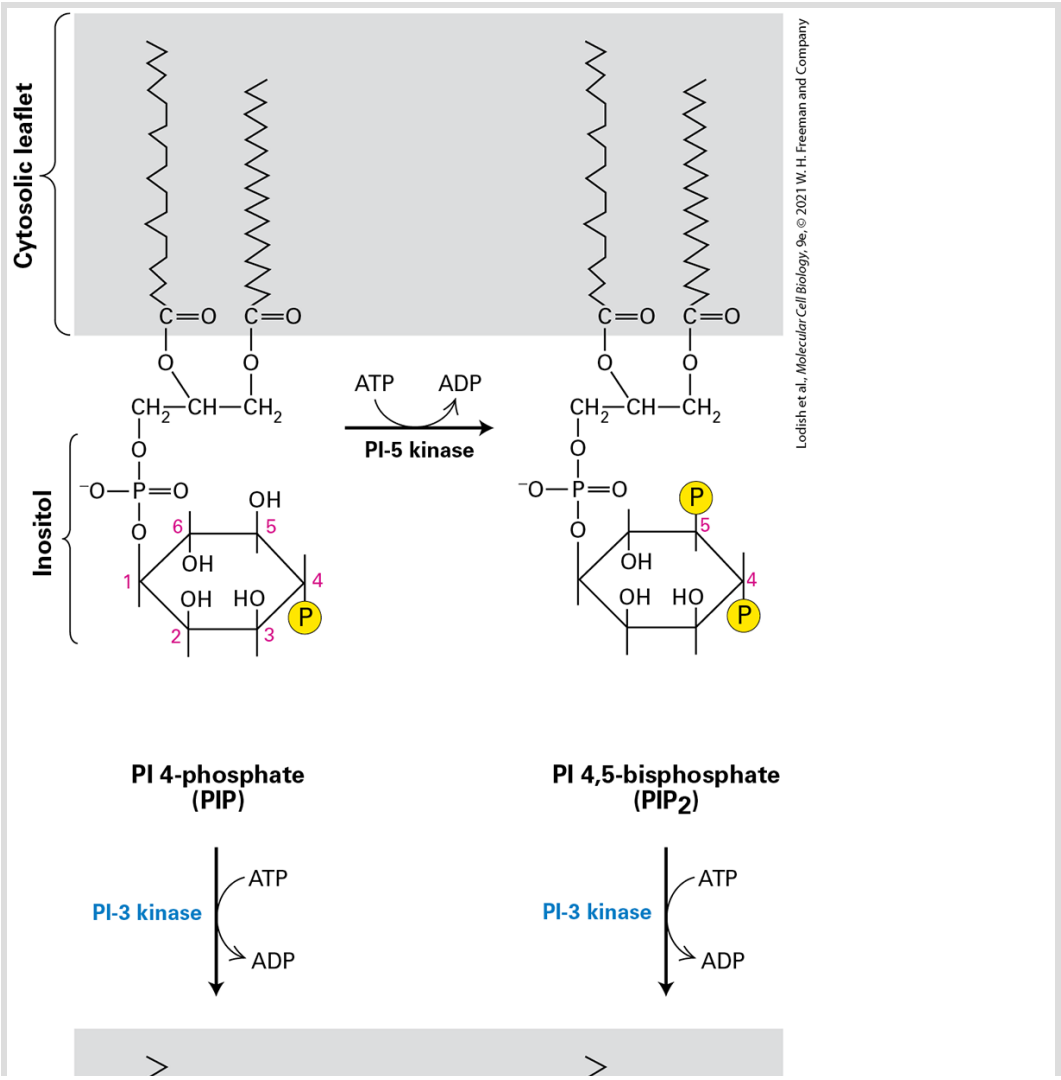
As in the case for many signal transduction proteins, there are multiple isoforms of PLC. Via trimeric G proteins, GPCRs specifically activate the β isoform of this enzyme (PLC_β). Many RTKs (and the cytokine receptors discussed in the following section) activate a different isoform of phospholipase C, the γ isoform (PLC_γ), which contains an SH2 domain. This SH2 domain binds to specific phosphotyrosines on the cytosolic domains of these activated receptors, thus recruiting the PLC_γ enzyme from the cytosol to the cytosolic face of the plasma membrane, close to its

substrate, $PI(4,5)P_2$. In addition, the receptor phosphorylates tyrosine residues on the bound PLC_γ, enhancing its hydrolase activity. Thus activated RTKs and cytokine receptors promote PLC_γ activity in two ways: by localizing the enzyme to the membrane and by phosphorylating it. As detailed in Chapter 15, the IP_3/DAG pathway initiated by PLC has multiple physiological effects.

Recruitment of PI-3 Kinase to Activated Receptors Leads to Accumulation of Phosphatidylinositol-3-Phosphates in the Plasma Membrane and to Activation of Several Downstream Kinases

In addition to the IP_3/DAG pathway, many activated RTKs and cytokine receptors initiate another phosphoinositide pathway by recruiting the enzyme *phosphatidylinositol-3 kinase (PI-3 kinase)* to the cytosolic surface of the plasma membrane. PI-3 kinase is recruited to the plasma membrane by binding of its SH2 domain to phosphotyrosines on the cytosolic domain of many activated RTKs and cytokine receptors. This recruitment positions the catalytic domain of PI-3 kinase near its phosphoinositide substrates on the cytosolic face of the plasma membrane. Unlike kinases we have encountered earlier, which phosphorylate proteins, PI-3 kinase adds

a phosphate to the hydroxyl group on the 3′ carbon in one of two separate phosphatidylinositol substrates, leading to formation of either phosphatidylinositol 3,4-bisphosphate [$PI(3,4)P_2$] or phosphatidylinositol 3,4,5-trisphosphate [$PI(3,4,5)P_3$] — collectively called *PI 3-phosphates* (Figure 16-16). These membrane-bound PI 3-phosphate products of the PI-3 kinase act as docking sites for multiple signal-transducing proteins. Once bound to these phosphoinositides, these proteins transduce signals in several important pathways.



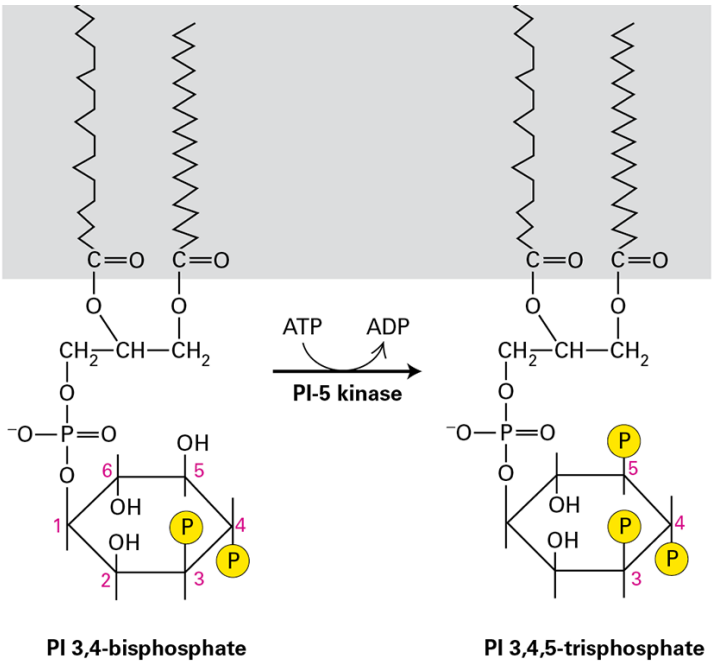


FIGURE 16-16 Generation of phosphatidylinositol 3-phosphates. The enzyme phosphatidylinositol-3 kinase (PI-3 kinase) is recruited to the cytosolic face of the plasma membrane by many activated receptor tyrosine kinases (RTKs) and cytokine receptors. The 3-phosphate added by this enzyme, to yield $\mathrm{PI(3,4)P_2}$ or $\mathrm{PI(3,4,5)P_3}$, is a binding site for various signal-transducing proteins, such as the PH domain of protein kinase B. See L. Rameh and L. C. Cantley, 1999, *J. Biol. Chem.* **274**:8347.

In some cells, this *PI-3 kinase pathway* triggers cell division and prevents apoptosis, thus ensuring cell survival. In other cells, this pathway induces specific changes in cell metabolism.

An important signal transduction protein that binds to PI 3-phosphates is **protein kinase B (PKB)**, a serine/threonine kinase that is also called **Akt**. In addition to its kinase domain, PKB contains a *PH domain*, a conserved protein domain present in a wide variety of signaling proteins that binds with high affinity to the 3-

phosphate in both $PI(3,4)P_2$ and $PI(3,4,5)P_3$. In unstimulated, resting cells, the concentration of these phosphoinositides is low; PKB is present in the cytosol in an inactive form because its PH domain physically blocks the kinase active site (Figure 16-17). Following ligand-induced receptor activation, the concentration of PI 3-phosphates rises. The PH domain of PKB binds to these PI 3-phosphates, so it no longer blocks the PKB kinase active site. PKB is thus recruited to the cytosolic face of the plasma membrane, and its unblocked kinase domain becomes partially activated. Maximal activation of PKB depends on recruitment to the cytosolic surface of the plasma membrane of two other kinases, named *PDK1* and *PDK2*.

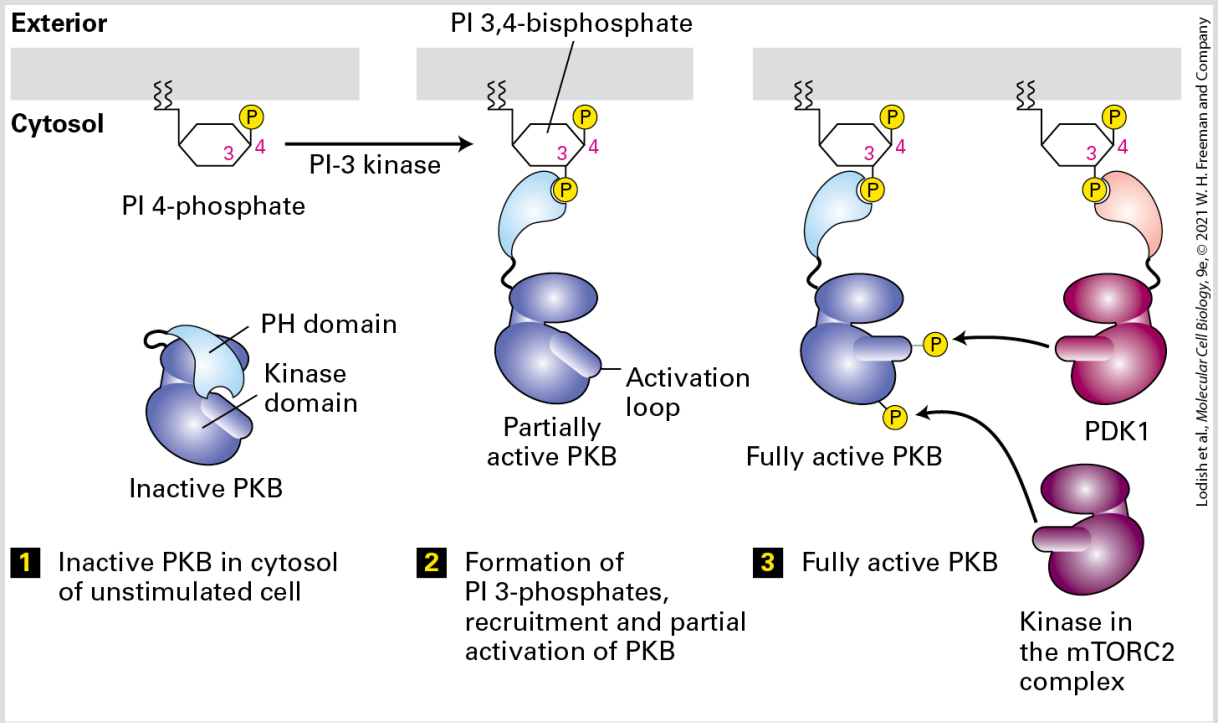


FIGURE 16-17 Recruitment and activation of protein kinase B (PKB) in PI-3 kinase pathways. In unstimulated cells (step 1), PKB is in the cytosol with its PH domain bound to its catalytic kinase domain, inhibiting its activity. Hormone stimulation leads to activation of PI-3 kinase and subsequent formation of $PI(3,4)P_2$ or $PI(3,4,5)P_3$ (see Figure 16-16).

The 3-phosphate group on both serves as a docking site on the plasma membrane for the PH domain of PKB (step **2**) and another kinase, PDK1. Full activation of PKB requires phosphorylation both in the activation loop by PDK1 and at the C-terminus by a second kinase, PDK2, that is part of the mTORC2 complex described in Chapter 21 (step **3**). See A. Toker and A. Newton, 2000, *Cell* **103**:185; and S. Sarbassov et al., 2005, *Curr. Opin. Cell Biol.* **17**:596.

PDK1 is recruited to the plasma membrane via binding of its own PH domain to PI 3-phosphates. Anchored to distinct PI 3-phosphates, both PKB and PDK1 diffuse randomly in the plane of the plasma membrane, eventually coming close enough together so that PDK1 can phosphorylate PKB on a critical threonine residue in its activation loop — yet another example of kinase activation by phosphorylation. Phosphorylation by the mTORC2 kinase complex, described in Chapter 21, of a second serine on PKB that is not in the activation loop is necessary for maximal PKB activity (see Figure 16-17). As in the regulation of Raf activity (see Figure 16-13), release of an inhibitory domain and phosphorylation by other kinases regulate the activity of PKB.

Activated Protein Kinase B Induces Many Cellular Responses

Once fully activated, PKB can dissociate from the plasma membrane and phosphorylate its many target proteins throughout the cell. These targets have a wide range of effects on cell behavior. In many

cells, activated PKB directly phosphorylates and inactivates a protein termed *Bad*, which as its name suggests is not good for a cell's survival — it will induce programmed cell death (apoptosis; see Figure 22-42). Phosphorylation of Bad by PKB protects against apoptosis.

Activated PKB also promotes survival of many cultured cells by phosphorylating the Forkhead transcription factor FOXO3a on multiple serine and threonine residues. When FOXO3a is not phosphorylated, it mainly localizes to the nucleus, where it activates transcription of several genes encoding proteins that induce apoptosis. When growth factors are added to the cells, PI-3 kinase is activated, and consequently so is PKB. Activated PKB phosphorylates FOXO3a. The cytosolic phosphoserine-binding protein 14-3-3 then binds to phosphorylated FOXO3a and sequesters it in the cytosol. (Recall that many other phosphorylated proteins, including Raf, that bind to 14-3-3 are similarly retained in an inactive state in the cytosol; see Figure 16-13.) A FOXO3a mutant in which the three serine residues that are targets for PKB are mutated to alanines is constitutively active as a transcription factor and initiates apoptosis even in the presence of activated PKB. This finding demonstrates the importance of phosphorylation of FOXO3a by PKB in preventing induction of apoptosis.

Misregulation of PKB is implicated in the pathogenesis of both cancer and diabetes, and in Chapter 21 we will see how PKB,

activated downstream of the insulin RTK, promotes glucose uptake and storage in muscle and liver. This is another example of one signaling pathway controlling different cellular functions in different cells.

The PI-3 Kinase Pathway Is Negatively Regulated by PTEN Phosphatase

Like most intracellular signaling events, phosphorylation by PI-3 kinase is reversible. The phosphatase that removes the phosphates added by PI-3-kinase, termed *PTEN*, has an unusually broad specificity. Although PTEN can remove phosphate groups attached to serine, threonine, and tyrosine residues in proteins, its major function in cells is thought to be its ability to remove the 3-phosphate from $PI(3,4)P_2$ and $PI(3,4,5)P_3$ and thus suppress the PI-3 kinase signaling pathway. Overexpression of PTEN in cultured mammalian cells promotes apoptosis by reducing the level of $PI(3,4)P_2$ and $PI(3,4,5)P_3$, and hence reducing the activation of PKB and limiting its anti-apoptotic effect.

The *PTEN* gene is deleted in multiple types of advanced human cancers, and its loss contributes to the uncontrolled growth of cells. Indeed, cells lacking PTEN have elevated levels $PI(3,4,5)P_3$ and PKB activity. Because PKB inactivates a protein that induces apoptosis, loss of PTEN reduces the programmed cell death that is the normal

fate of many cells including many tumors. In certain cells, such as neuronal stem cells, absence of PTEN not only prevents apoptosis but also stimulates cell cycle progression and enhances the rate of cell proliferation. Knockout mice lacking PTEN have big brains with an excess number of neurons, attesting to PTEN's importance in the control of normal brain development. ■

KEY CONCEPTS OF SECTION 16.3

Phosphoinositide Signal Transduction Pathways

- Many RTKs and cytokine receptors initiate the $\mathrm{IP_3/DAG}$ signaling pathway by activating phospholipase $\mathrm{C_\gamma}$ $(\mathrm{PLC_\gamma})$, a different PLC isoform than the one activated by G protein–coupled receptors.
- Activated RTKs and cytokine receptors also can initiate another phosphoinositide pathway by binding a PI-3 kinase, thereby allowing the enzyme access to its membrane-bound phosphoinositide substrates, which then become phosphorylated at the 3 position (PI 3-phosphates $\mathrm{PI(3,4)P_2}$ and $\mathrm{PI(3,4,5)P_3}$; see Figure 16-16).
- The PH domains in various proteins bind to PI 3-phosphates, forming signaling complexes associated with the cytosolic face of the plasma membrane.
- Protein kinase B (PKB) becomes partially activated by binding with its PH domain to PI 3-phosphates. Full activation of PKB requires phosphorylation by the kinase PDK1, which is also recruited to the membrane by binding to PI 3-phosphates, and by a second kinase, PDK2 (see Figure 16-17).
- Activated PKB promotes survival of many cells by directly phosphorylating and inactivating several pro-apoptotic proteins and by phosphorylating and inactivating the FOXO3a transcription factor, which otherwise induces synthesis of pro-apoptotic proteins.
- Signaling via the PI-3 kinase pathway is terminated by the PTEN phosphatase, which hydrolyzes the 3-phosphate in PI 3-phosphates. Loss of PTEN, a common occurrence in human tumors, promotes cell survival and proliferation.

16.4 Cytokines, Cytokine Receptors, and the JAK/STAT Signaling Pathway

The **cytokines** form a family of relatively small, secreted signaling molecules, generally containing about 160–200 amino acids, that control growth and differentiation of many types of cells. All cytokines have a similar three-dimensional structure — four long, conserved α helices folded into a conserved three-dimensional structure. All cytokine receptors also have similar three-dimensional structures; they function similarly to RTKs and activate similar signal transduction pathways. Whereas in RTKs the protein tyrosine kinase is part of the receptor cytosolic domain, **cytokine receptors** and the kinase enzyme are separate polypeptides, encoded by different genes, yet are bound tightly together. The tightly bound kinase is known as a *JAK kinase*. (JAK is an acronym for "*J*ust *A*nother *K*inase" because when the first member of the family was cloned researchers knew it was a kinase from its sequence, but had no idea of its function.) In this section, we focus here on a pathway — the JAK/STAT pathway — mainly employed by cytokine receptors: the JAK kinase phosphorylates and activates a STAT transcription factor that moves directly to the nucleus. More specifically, the SH2 domain of a STAT transcription factor binds to a phosphotyrosine in the cytosolic domain of an activated cytokine receptor; the STAT transcription factor then becomes

phosphorylated by the JAK kinase, moves to the nucleus, and directly activates gene expression.

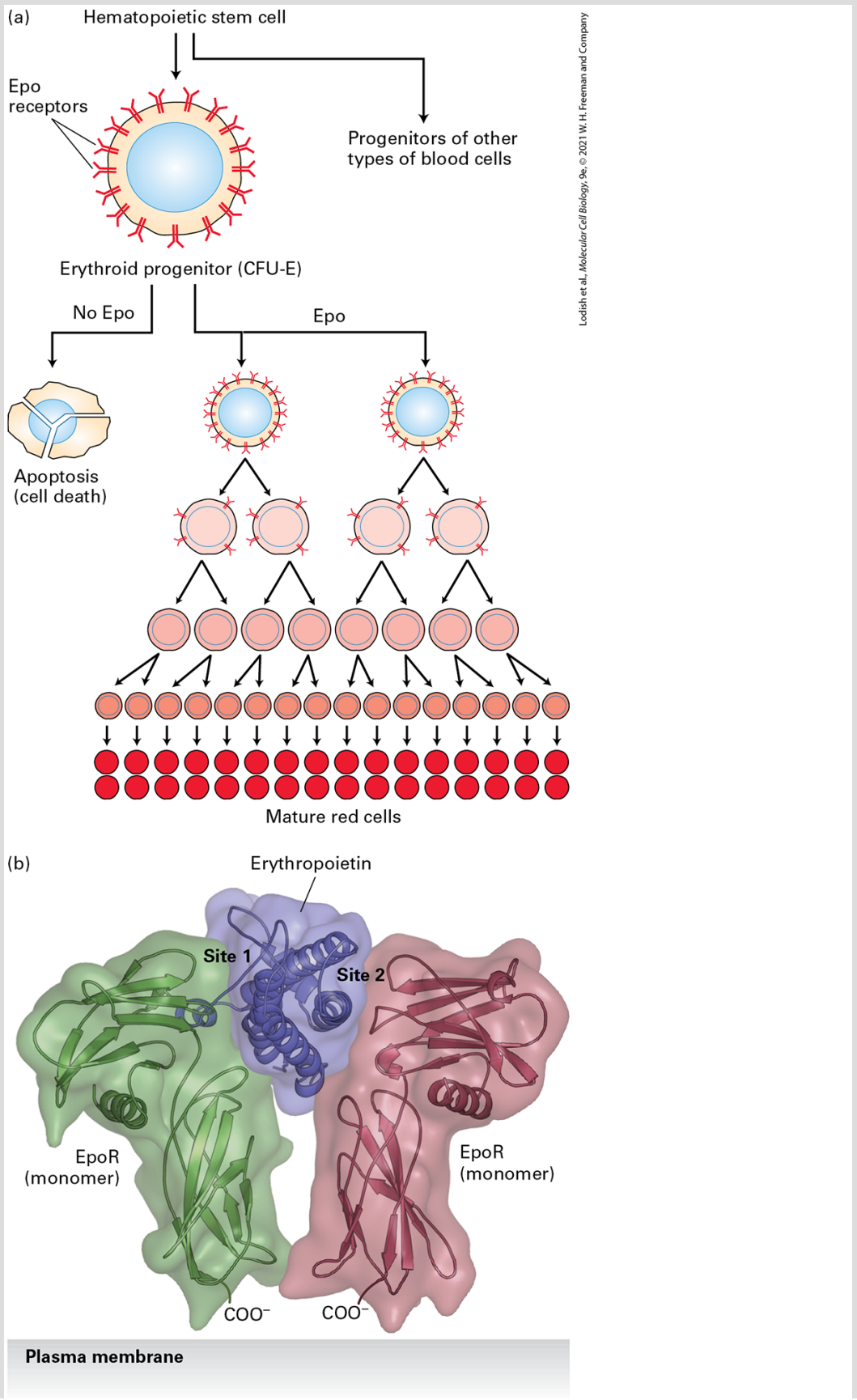
Cytokines Regulate the Development and Function of Many Cell Types

Growth hormone (GH), as its name implies, is a 191-amino-acid protein that stimulates proliferation of many types of body cells; it is made and secreted by cells in the anterior pituitary gland in response to another hormone, growth hormone–releasing hormone, that is made by the part of the brain termed the *hypothalamus*. GH was one of the first protein drugs to be made by recombinant DNA; it is used clinically to treat short stature caused by GH deficiency in children and adults. The bovine version is used to increase milk production in dairy cows. During pregnancy, a related hormone, the cytokine **prolactin**, induces epithelial cells lining the immature ductules of the mammary gland to differentiate into the acinar cells that produce milk proteins and secrete them into the ducts. GH and prolactin have three-dimensional structures very similar to those of the cytokines that induce the formation of important types of blood cells.

All blood cells are derived from hematopoietic stem cells, which form a series of progenitor cells that then differentiate into the mature blood cells (see Figure 22-18). For instance, the cytokine

granulocyte colony–stimulating factor (G-CSF) induces a granulocyte progenitor cell in the bone marrow to divide several times and then differentiate into granulocytes, a type of white blood cell that inactivates invading bacteria and other pathogens. A related cytokine, **thrombopoietin**, stimulates a different progenitor cell to divide and differentiate into megakaryocytes, huge cells that fragment into the platelets that are essential for blood clotting.

A structurally related cytokine, **erythropoietin (Epo)**, triggers production of erythrocytes (red blood cells) by inducing the proliferation and differentiation of erythroid progenitor cells in the bone marrow (Figure 16-18a). Erythropoietin is synthesized by certain kidney cells. The major function of erythrocytes is to transport oxygen complexed to hemoglobin. Thus a drop in blood oxygen, such as that caused by loss of blood from a large wound, signifies a lower than optimal level of red blood cells. As we learn in Chapter 21, the oxygen-sensing transcription factor HIF-1α becomes activated and induces the transcription of the erythropoietin gene; the kidney cells synthesize more of this cytokine and secrete it into the blood. As the level of erythropoietin rises, more and more erythroid progenitors are induced to divide and differentiate; each progenitor produces 30 to 50 erythrocytes in only a few days. In this way, the body can respond to the loss of blood or reduction in available atmospheric oxygen at high elevations by accelerating the production of erythrocytes. GH, prolactin, G-CSF, thrombopoietin, and Epo undoubtedly evolved from a common ancestral protein, since all of these cytokines have a similar tertiary structure.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 16-18 Erythropoietin and formation of red blood cells (erythrocytes). (a) Differentiation of erythroid progenitors to erythrocytes. Erythroid progenitor cells, called *colony-forming units–erythroid (CFU-E),* are derived from hematopoietic stem cells, which also give rise to progenitors of other blood cell types (see Figure 22-18). In the absence of erythropoietin, CFU-E cells undergo apoptosis (programmed cell death). Binding of Epo to its receptors on a CFU-E cell induces transcription of several genes whose encoded proteins prevent apoptosis, allowing the cell to survive. Other Epo-induced proteins trigger a developmental program of three to six terminal cell divisions, induction of hemoglobin and many other erythroid-important genes, reduction in cell and nuclear size, and finally, loss of the cell nucleus. If CFU-E cells are cultured with Epo in a semisolid medium (e.g., containing methylcellulose), daughter cells cannot move away, and thus each CFU-E cell produces a colony of 30–100 erythroid cells; hence its name. (b) Structure of erythropoietin bound to an erythropoietin receptor. Like other cytokines, erythropoietin (Epo) contains four conserved, long α helices that are folded into a conserved structural arrangement. The activated erythropoietin receptor (EpoR) is a dimer of identical subunits. Like the extracellular domains of all cytokine receptors, the extracellular domain of each monomer of the EpoR is constructed of two subdomains, each containing seven conserved β strands folded together in a characteristic and conserved structure. Side chains of residues on two of the α helices in Epo, termed *site 1*, contact loops on one EpoR monomer, while residues on the two other Epo α helices, termed *site 2*, bind to the same loop segments in a second receptor monomer, thereby stabilizing the dimeric receptor in a specific active conformation.

[Part (a) Data from M. Socolovsky et al., 2001, *Blood* **98**:3261. Part (b) Data from R. S. Syed et al., 1998, *Nature* **395**:511, PDB ID 1eer.]

Both Epo and G-CSF are produced commercially by recombinant expression in cultured mammalian cells. Patients with kidney disease, especially those undergoing dialysis, frequently are anemic (have a low red blood cell count) and therefore are treated

with recombinant Epo to boost red cell levels. Epo and G-CSF are used as adjuncts to certain cancer therapies because many cancer treatments reduce the production of red cells and granulocytes in the bone marrow. Epo can also be misused to increase red blood cell levels in endurance athletes (blood "doping"). ■

Cytokines include the *interferons*, which are produced and secreted by several cell types following viral infection. Interferons act on nearby cells to induce enzymes that render those cells more resistant to viral infection. Other cytokines, exemplified by Interleukin 2 (IL-2), mainly regulate cells of the immune system. IL-2 promotes the proliferation, differentiation, and survival of key cells of the immune system, including antibody-forming B cells and also several types of T cells (Chapter 24).

Binding of a Cytokine to Its Receptor Activates One or More Tightly Bound JAK Protein Tyrosine Kinases

GH, prolactin, G-CSF, thrombopoietin, and Epo all have similar structures, and they activate receptors of similar structure by forming dimers of two identical cytokine receptor proteins, a process termed *receptor homodimerization*. The extracellular domains of these cytokine receptors are constructed of two subdomains, each of which contains seven conserved β strands

folded together in a characteristic fashion (Figure 16-18b). Cytokines bind their receptors very tightly, as a consequence of molecular complementarity and multiple weak, noncovalent forces — that is, ionic, van der Waals, and hydrophobic interactions [see Figure 15-4; dissociation constants (K_d) of about 10^{-10} M]. Thus very low concentrations of GH and most other cytokines — about 1 microgram per liter (roughly one part in a billion) — are sufficient to activate cytokine receptors.

Other cytokines, including the interleukins and the interferons, bind simultaneously to two or more different cytokine receptors, a process called *receptor hetero-dimerization* or *hetero-oligomerization* (Figure 16-19b). IL-2 can bind first to the alpha chain (IL-2Rα), which does not directly interact with downstream signal transduction molecules, but rather enhances the binding of IL-2 to the two signaling subunits, the beta and gamma chains IL-2Rβ and IL-2Rγ. The same gamma chain is also an essential subunit of the receptors for several other interleukins, including IL-4, IL-7, IL-9, IL-15, and IL-21, cytokines that are essential for formation of B cells and other types of immune-system cells.

Severe combined immunodeficiency (SCID) is a genetic disease in which neither T nor B cells are produced. People with SCID cannot cope with any bacterial or viral infection, so they must be kept in a sterile environment (like the famous “bubble boy” who was kept in a plastic bubble and eventually succumbed to an infection).

Many cases of SCID are due to a deficiency in the IL-2 receptor gamma chain, whose gene is located on the X chromosome. These children can now be cured by gene therapy: a lentiviral vector (see Figure 6-37) is used to introduce a functional gamma chain gene into the hematopoietic stem cells that generate all immune-system cells (see Chapter 23), and the children can live normal lives. ■

The JAK family of kinases contain four members. An identical JAK kinase of one of these types is bound to the cytosolic domain of each of the two polypeptides in homodimeric receptors (Figure 16-19a, step **1**). Two distinct JAK kinase isoforms bind to heterodimeric receptors (Figure 16-19b). Despite this difference, the signaling pathways activated by all cytokine receptors are broadly similar; here we focus on the simpler case of receptors that homodimerize. Our example is the interaction of one erythropoietin molecule with two identical erythropoietin receptor (EpoR) proteins (Figures 16-19a and 16-20a).

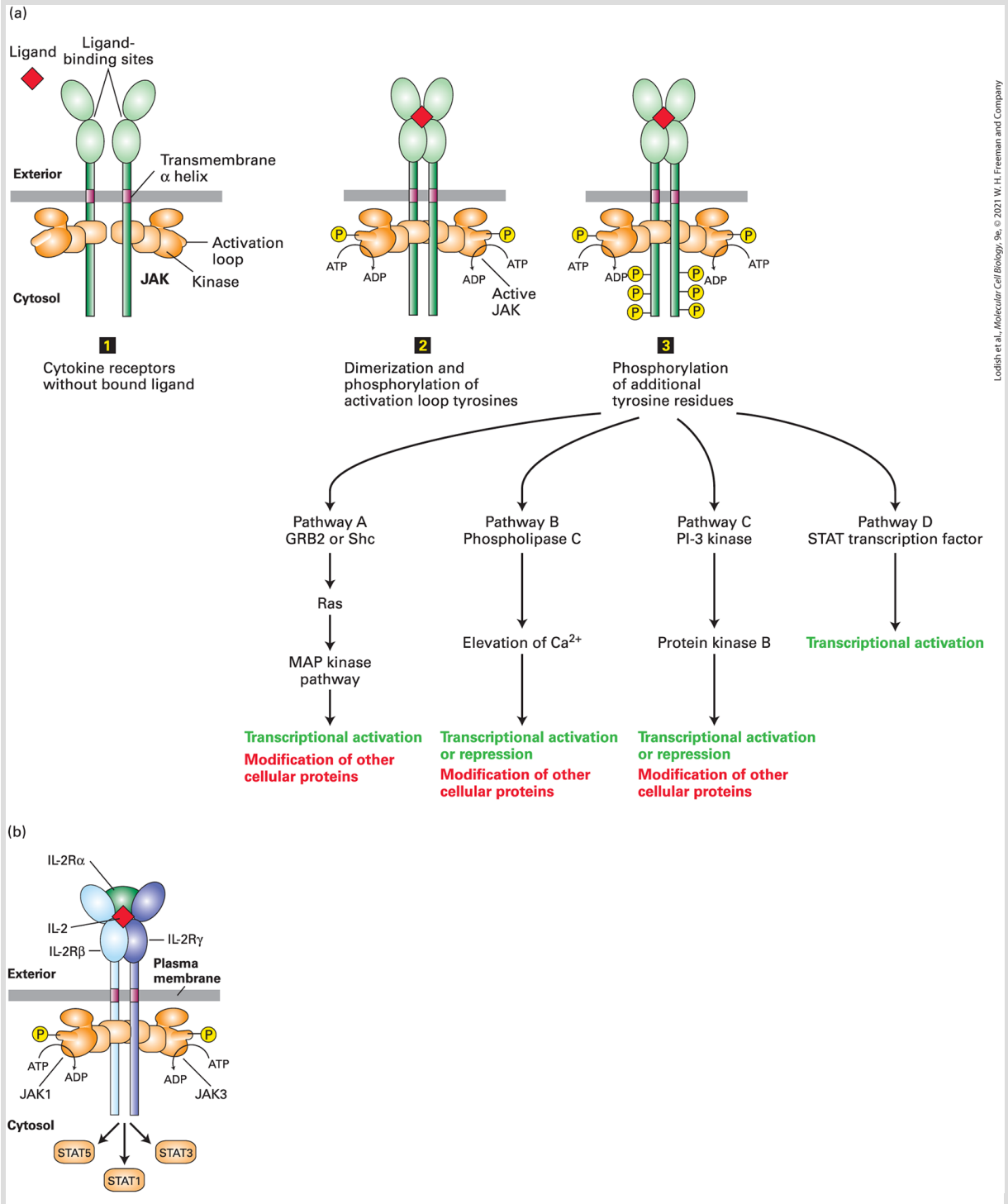


FIGURE 16-19 Structure and activation of cytokine receptors. (a) Homodimeric receptors. The cytosolic domains of both cytokine receptor subunits bind tightly and irreversibly to the same JAK protein tyrosine kinase. In the absence of ligand (step 1), the receptors are largely monomers and the JAK kinases are poorly active. Ligand binding causes a conformational change in the receptors that brings together the JAK kinase domains, which

then phosphorylate each other on a tyrosine residue in the activation loop, activating kinase activity (step **2**). The active JAK kinases then phosphorylate multiple tyrosine residues in the receptor cytosolic domain (step **3**). As in RTKs (see Figure 16-1a), the resulting phosphotyrosines function as docking sites for SH2 domains of multiple signal-transducing proteins, including STAT proteins. (b) Heteromeric receptors, as exemplified by the binding of Interleukin 2 (IL-2) to the extracellular domains of the three subunits of the IL-2 receptor. Both the beta and gamma subunits of the IL-2 receptor (IL-2Rβ and IL-2Rγ) have structures similar to that of the Epo receptor. The alpha subunit (IL-2Rα) is expressed in some but not all responding cells and does not activate a signal transduction pathway. Its function is to bind IL-2 and facilitate the binding of IL-2 to the signaling IL-2Rβ and IL-2Rγ subunits. Each of the two signaling receptors, IL-2Rβ and IL-2Rγ, is bound to a different JAK kinase and activates different STAT signal-transducing proteins but otherwise the receptor complex activates similar same downstream signaling pathways as depicted in (a).

Each of the four members of the JAK family of kinases contains an N-terminal domain that binds to the receptor, a C-terminal kinase domain that has low catalytic activity when no ligand is bound, and a middle *pseudokinase* domain that regulates kinase activity by an as yet unknown mechanism. These kinases become activated after ligand binding and receptor dimerization (Figure 16-19a, step **1**); as with many RTKs, ligand binding triggers a conformational change in the JAKs such that they phosphorylate each other on a critical tyrosine in the activation loop (Figure 16-19a, step **2**; see Figure 15-6), which greatly enhances kinase activity. Also as with RTKs, the active JAK kinase phosphorylates several tyrosine residues in the receptor's cytosolic domain (Figure 16-19a, step **3**), creating binding sites for SH2 domains of many signal transduction proteins, including those that activate the Ras/MAP kinase and other signaling pathways discussed in Sections 16.2 and 16.3. Here we

focus on a pathway — the JAK/STAT pathway — mainly employed by cytokine receptors: the JAK kinase phosphorylates and activates a STAT transcription factor that moves directly to the nucleus (Figure 16-20a).

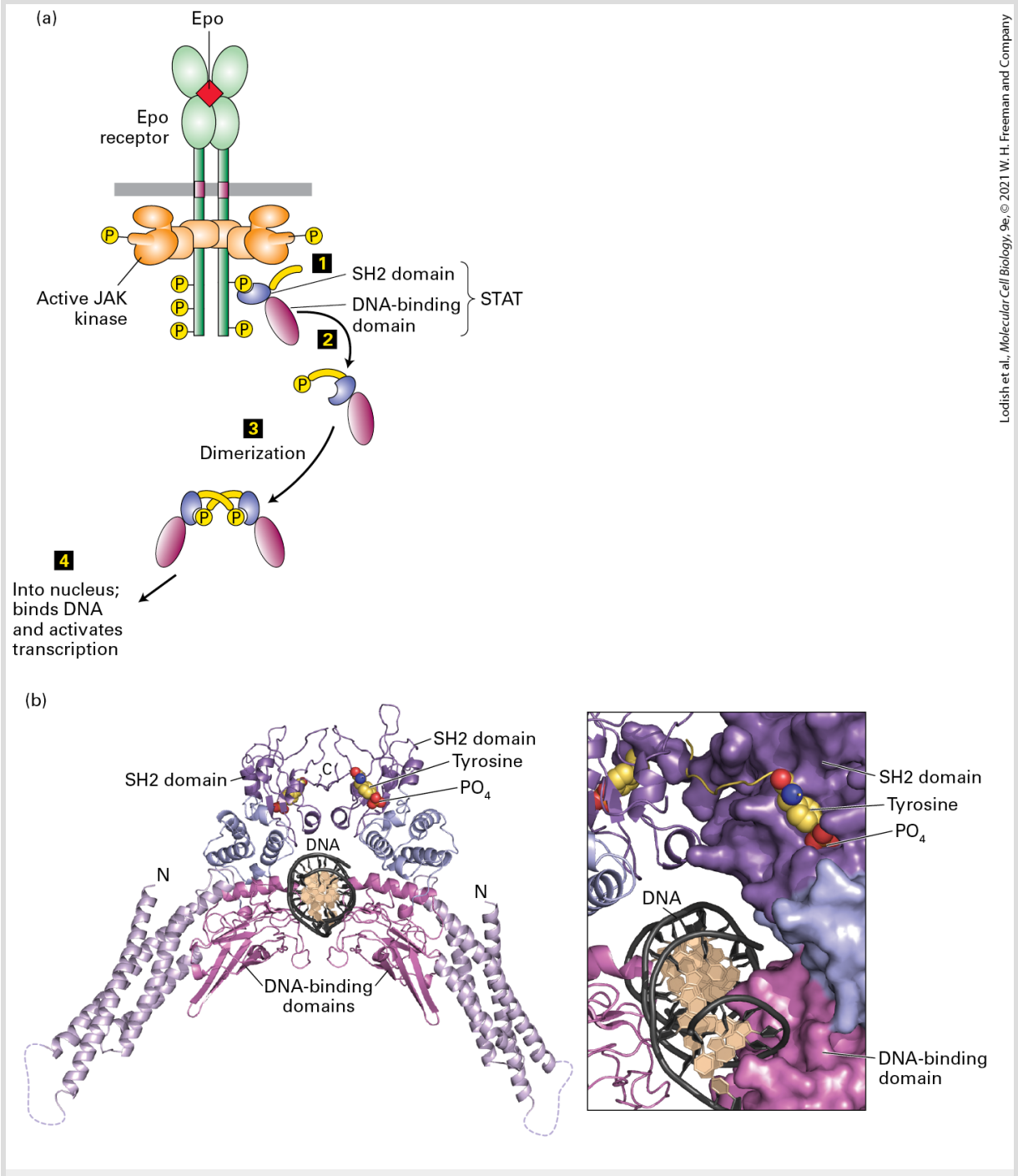


FIGURE 16-20 Activation and structure of STAT proteins. (a) Phosphorylation and dimerization of STAT proteins. Step **1**: Following dimerization and activation of a JAK kinase bound to a cytokine receptor (see Figure 16-19a), the SH2 domain of an inactive monomeric STAT transcription factor binds to a phosphotyrosine in the receptor cytosolic domain, bringing the STAT close to the active JAK associated with the receptor. The JAK then phosphorylates a specific C-terminal tyrosine in the STAT. Steps **2** and **3**: Phosphorylated STATs spontaneously dissociate from the receptor and spontaneously homodimerize. Because the STAT homodimer has two phosphotyrosine-SH2 domain interactions, whereas the receptor-STAT complex is stabilized by only one such interaction, phosphorylated dimeric STATs tend not to dissociate and rebind to the receptor. Step **4**: The STAT dimer moves into the nucleus, where it binds to promoter sequences and activates transcription of target genes. (b) Ribbon diagram of the STAT1 dimer bound to DNA (black). The STAT1 dimer forms a C-shaped clamp around DNA that is stabilized by reciprocal and highly specific interactions between the SH2 domain (purple) of one monomer and the phosphorylated tyrosine residue (yellow with red oxygens) on the C-terminal segment of the other. The phosphotyrosine-binding site of the SH2 domain in each monomer is coupled structurally to the DNA-binding domain (magenta), suggesting a potential role for the SH2-phosphotyrosine interaction in the stabilization of DNA interacting elements.

[Part (b) Data from X. Chen et al., 1998, *Cell* **93**:827, PDB ID 1bf5.]

JAK Kinases Phosphorylate and Activate STAT Transcription Factors

All STAT proteins contain an N-terminal DNA-binding domain, an SH2 domain, and a C-terminal domain with a critical tyrosine residue. Through its SH2 domain, a monomeric STAT binds to a specific phosphotyrosine in the cytosolic domain of an activated receptor (Figure 16-20a, step **1**). The C-terminal tyrosine then becomes phosphorylated by the adjacent JAK kinase (step **2**). This

arrangement ensures that in a particular cell, only those STAT proteins with an SH2 domain that can bind to a particular receptor will be activated when that receptor is activated by cytokine binding. For example, the erythropoietin receptor, as well as the receptors for GH, prolactin, G-CSF, and several other cytokines, activates STAT5, but not STATs 1, 2, 3, or 4. In contrast the IL-2 receptor complex activates STATs 1, 3, and 5 but not STATs 2 or 4 (see Figure 16-19b).

A phosphorylated STAT dissociates spontaneously from the receptor, and two phosphorylated STAT proteins form a homodimer in which the SH2 domain on each molecule in the dimer binds to the phosphotyrosine in the other molecule (Figure 16-20b). Conformational changes take place on dimerization that expose the nuclear-localization signal (NLS), which had been inaccessible in the unphosphorylated STAT. NLSs are present in virtually all transcription factors found in the cytosol and are required for their transport into the nucleus (see Chapter 13). Targeted by the NLS, the STAT dimers move into the nucleus, where they bind to specific **enhancers** or **promoters** (DNA regulatory sequences) controlling target genes and thus alter gene expression.

As noted above (see Figure 16-2), because different cell types have unique complements of transcription factors and unique epigenetic modifications of their chromatin, the genes that are available to be activated by any STAT are different in different cell types. For example, in mammary gland cells, STAT5 becomes activated following prolactin binding to the prolactin receptor and induces

transcription of genes encoding milk proteins. In contrast, when STAT5 becomes activated in erythroid progenitor cells following binding of Epo to the EpoR, it induces expression of the protein Bcl-x_L · Bcl-x_L prevents the apoptosis of these progenitors (see Chapter 22), allowing them to proliferate and differentiate into red blood cells. More generally, it is thought that in each type of cell, activated STAT proteins, like the other transcription factors discussed earlier, bind only to DNA sites in open chromatin and mainly to sites that have master transcription factors or other cell-specific gene regulatory proteins bound at adjacent sites. This strategy of activating sets of genes specific to specific cell types by a relatively limited set of receptors, JAK kinases, and STAT proteins, allows the control, via combinatorial diversity, of a vast array of cellular activities.

Multiple Mechanisms Suppress Signaling from Cytokine Receptors

In Chapter 15 and in this one, we saw several ways in which signaling from G protein–coupled receptors and RTKs is terminated to prevent cells from overresponding to external stimuli. Here we discuss two mechanisms by which cytokine receptor signaling is regulated.

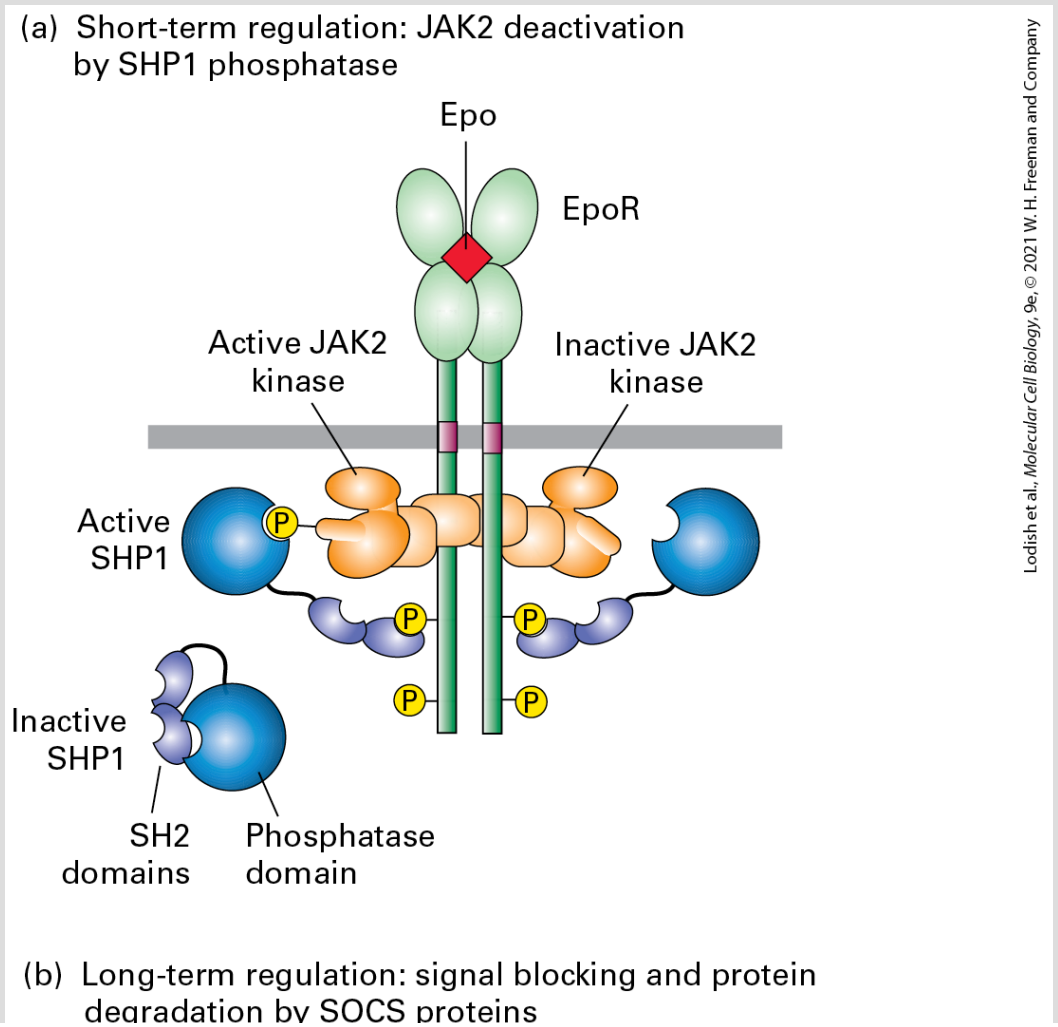
Phosphotyrosine Phosphatases

Phosphotyrosine phosphatases are enzymes that remove (hydrolyze) phosphates from specific phosphotyrosines on specific target proteins. An excellent example of how phosphotyrosine phosphatase enzymes function to suppress the activity of protein tyrosine kinases is provided by SHP1, a phosphatase that negatively regulates signaling by several types of cytokine receptors. Its role was first identified by analysis of mutant mice (termed *motheaten*) that lacked this protein; they died because of excess production of several types of blood cells, including erythrocytes.

SHP1 dampens cytokine signaling by binding to a cytokine receptor and inactivating the associated JAK protein, as depicted in Figure 16-21a. In addition to a phosphatase catalytic domain, SHP1 has two SH2 domains. When cells are in the resting state, unstimulated by a cytokine, SHP1 is free in the cytosol and one of its SH2 domains physically binds to and masks the catalytic site in the enzyme's phosphatase domain. In the stimulated state, however, this blocking SH2 domain preferentially binds to a specific phosphotyrosine residue in the activated receptor and unblocks the catalytic site. The accompanying conformational change brings the SHP1 catalytic site adjacent to the phosphotyrosine residue in the activation loop of the JAK associated with the receptor. By removing this phosphate, SHP1 inactivates the JAK, so that it can no longer phosphorylate the receptor or other substrates including the STATs.

SOCS Proteins

In another example of negative feedback, among the genes whose transcription is induced by STAT proteins are those encoding a class of small proteins termed *s*uppressors *o*f *c*ytokine *s*ignaling (SOCS), which terminate signaling by cytokine receptors. All SOCS proteins contain an SH2 domain that binds to specific phosphotyrosines on an activated receptor (Figure 16-21b). Binding of SOCS to an activated receptor results in the degradation of both the receptor and its associated JAK kinase.



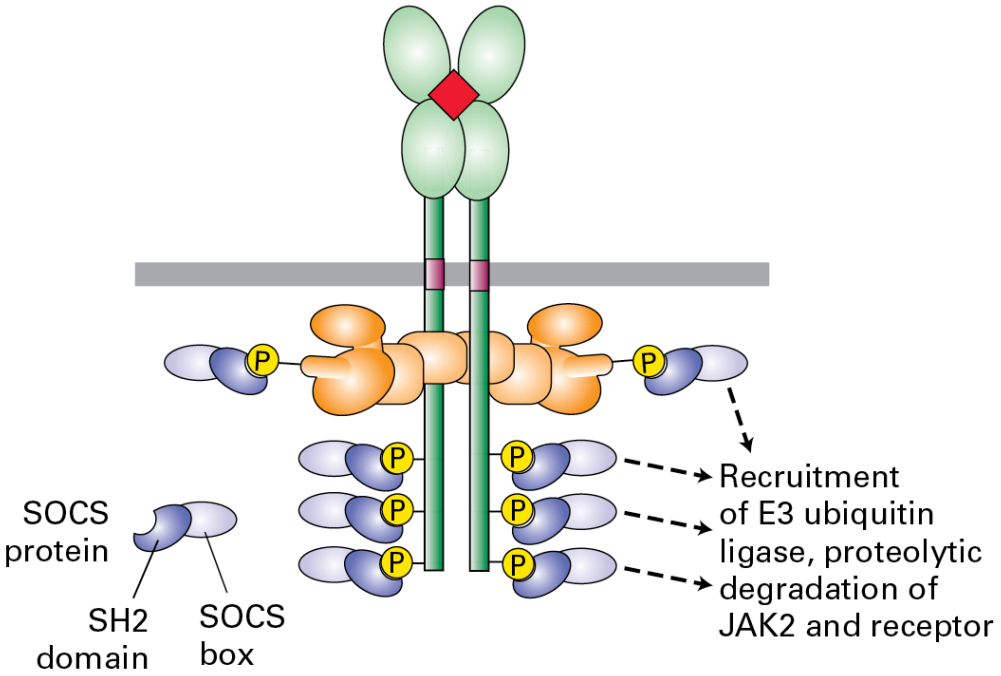


FIGURE 16-21 Two mechanisms for terminating cytokine signal transduction, as exemplified by the erythropoietin receptor (EpoR). (a) Short-term regulation: SHP1, a phosphotyrosine phosphatase, is present in an inactive form in the cytosol of unstimulated cells. Binding of an SH2 domain in SHP1 to a particular phosphotyrosine in the activated receptor unmasks its phosphatase catalytic site and positions it near the phosphorylated tyrosine in the activation loop region of JAK2. Removal of the phosphate from this tyrosine inactivates the JAK kinase. See S. Constantinescu et al., 1999, *Trends Endocrin. Met.* **10**:18. (b) Long-term regulation: SOCS proteins, whose expression is induced by the STAT5 protein in erythropoietin-stimulated erythroid progenitor cells, inhibit or permanently terminate signaling over longer periods. Binding of SOCS to specific phosphotyrosine residues on EpoR or JAK2 blocks binding of other signaling proteins (*left*). The SOCS box also targets the receptor as well as JAK2 for degradation by the ubiquitin-proteasome pathway (*right*). Similar mechanisms regulate signaling from other cytokine receptors. See B. T. Kile and W. S. Alexander, 2001, *Cell. Mol. Life Sci.* **58**:1627.

SOCS acts by targeting these two proteins to be degraded by a proteasome. SOCS proteins contain a domain, called the *SOCS box*, that recruits components of E3 ubiquitin ligases (see Figure 3-32) that polyubiquitinylate the receptor and the JAK kinase (a polymer of ubiquitins is covalently attached). The *polyubiquitin tail* marks the proteins for destruction by a proteasome (see Chapter 3), thereby permanently turning off JAK2-mediated signaling pathways until new receptors and JAK2 proteins can be made. The observation that proteasome inhibitors prolong JAK2 signal transduction supports this mechanism. One SOCS protein, SOCS-1, also binds to the critical phosphotyrosine in the activation loop of an activated JAK2 kinase, thereby inhibiting its catalytic activity.

Studies with cultured mammalian cells have shown that the cytokine receptor for growth hormone is suppressed by another SOCS protein, SOCS-2. Strikingly, mice deficient in SOCS-2 grow significantly larger than their wild-type counterparts; these mice have long bones and proportionate enlargement of most organs. Thus SOCS proteins play an essential, negative role in regulating intracellular signaling from the receptors for erythropoietin, growth hormone, and other cytokines.

KEY CONCEPTS OF SECTION 16.4

Cytokines, Cytokine Receptors, and the JAK/STAT Signaling Pathway

- Cytokines play numerous roles in development. Erythropoietin, a cytokine secreted by kidney cells, promotes proliferation and differentiation of erythroid progenitor cells in the bone marrow to increase the number of mature red cells in the blood (see Figure 16-18a).
- Cytokines such as GH, prolactin, Epo, and G-CSF have very similar tertiary structures, as do their receptors. These and related cytokines form a cytokine: receptor homodimer complex on the cell surface (see Figures 16-18b and 16-19a).
- Other cytokines, exemplified by Interleukin 2, interact with two or more different receptor subunits, forming cytokine: receptor heteromultimeric complexes (see Figure 16-19b).
- The cytosolic domains of cytokine receptors are tightly bound to a JAK protein tyrosine kinase, which becomes activated after cytokine binding and receptor dimerization or heteromerization and phosphorylates tyrosine residues on the cytosolic domain of the receptor (see Figure 16-19a).
- The JAK/STAT pathway operates downstream from all cytokine receptors and some RTKs. STAT monomers bound to phosphotyrosines on receptors are phosphorylated by receptor-associated JAKs, then dissociate from the receptor, dimerize and move to the nucleus, where they activate transcription (see Figure 16-20).
- Signaling from cytokine receptors is terminated by the phosphotyrosine phosphatase SHP1 and several SOCS proteins (see Figure 16-21).

16.5 The TGF-β Family of Growth Factors, Their Receptor Serine Kinases, and the Smad Transcription Factors They Activate

In this section, we discuss an evolutionarily conserved large family of signaling molecules, called the **transforming growth factor β (TGF-β)** family, as well as a conserved family of cell-surface receptors, called the *TGF-β receptor family*. Members of these two families are present in invertebrate organisms in phyla such as Porifera (sponges), and Cnidaria (coral and hydra) as well as in all vertebrates.

Kinase domains in TGF-β receptors phosphorylate and thus trigger the activation of a conserved class of transcription factors (the **Smads**) that regulate many growth and differentiation pathways. Like the STATs discussed in the previous section, Smads are located in the cytosol in unstimulated cells, but when phosphorylated they move into the nucleus to regulate transcription. As with cytokine signaling, the TGF-β pathways have widely diverse effects in different types of cells because different members of the TGF-β family activate different members of the TGF-β receptor family, which activate different members of the Smad group of transcription factors. In addition, as with other receptor-activated transcription factors, we will see that the same activated Smad

protein partners with different transcription factors and thus activates different sets of genes in different types of cells.

The mammalian TGF-β family contains 33 members and includes a number of related extracellular signaling molecules that play widespread roles in regulating development. The founding member of the TGF-β family, TGF-β1, was identified on the basis of its ability to induce a malignant phenotype in several cultured early-stage mammalian cancer cell lines (hence the name "transforming growth factor"); in this case, TGF-β1 promoted metastases, the spreading and invasion of primary tumors, as discussed in Chapter 25. Ironically, the principal function of all three human TGF-β isoforms, TGF-β1, 2, and 3, in most normal (non-cancerous) mammalian cells is to prevent their proliferation by inducing the synthesis of proteins, including $p15^{INK4B}$, which inhibit the cyclin-dependent kinases (CDKs) that are essential for progression into the DNA synthesis (S) phase of the cell cycle (see Figures 1-22 and 19-12).

TGF-β is secreted by many cells in the body and inhibits the growth of both the secreting cell (autocrine signaling) and neighboring cells (paracrine signaling). Loss of TGF-β receptors, or of any of several proteins in the TGF-β signal transduction pathway, releases cells from this growth inhibition and is seen frequently in the early development of human tumors. TGF-β proteins also promote expression of cell-adhesion molecules and extracellular-matrix

molecules, which play important roles in tissue organization (see Chapter 20).

Other members of the TGF-β family, *bone morphogenetic proteins (BMP)*, were initially identified by their ability to induce bone formation when implanted in mice. One of these, BMP2, is used clinically to strengthen bone after severe fractures. Of the numerous BMP proteins subsequently recognized, many induce key steps in development, including formation of mesoderm and the earliest blood-forming cells; some are important for maintaining the undifferentiated state of cultured embryonic and adult stem cells (see Chapter 22). Many have nothing to do with bones; for example, the BMP4/BMP7 heterodimer is important in dorsal-ventral (back-to-front) patterning in vertebrate embryos. Similarly, a *Drosophila* homolog of TGF-β, called *Dpp*, participates in dorsal-ventral patterning in fly embryos. In the early human male embryo another member of the human TGF-β family, Müllerian-inhibiting hormone (MIH), inhibits the development of the female reproductive tract, and thus is critical to sex differentiation at a specific time during the fetal development of human males.

TGF-β Proteins Are Stored in an Inactive Form in the Extracellular Matrix

In mammals, the monomeric active forms of all members of the TGF-β family contain a conserved cysteine residue that forms a disulfide bond linking two monomers into a dimer (Figure 16-22a), and the precursor proteins of all TGF-β family members dimerize in the endoplasmic reticulum immediately after synthesis. The genes for all members of the family encode a precursor polypeptide comprising a secretion signal peptide (Chapter 13), a ~250-residue prodomain, and the C-terminal ~110-residue TGF-β domain. The amino-terminal *prodomain* is required for proper folding and dimerization of the C-terminal TGF-β domain. As these dimeric precursor proteins pass through the Golgi complex, a protease severs the covalent connection between the prodomain and the TGF-β domain, but the prodomain and the mature TGF-β protein remain noncovalently attached.

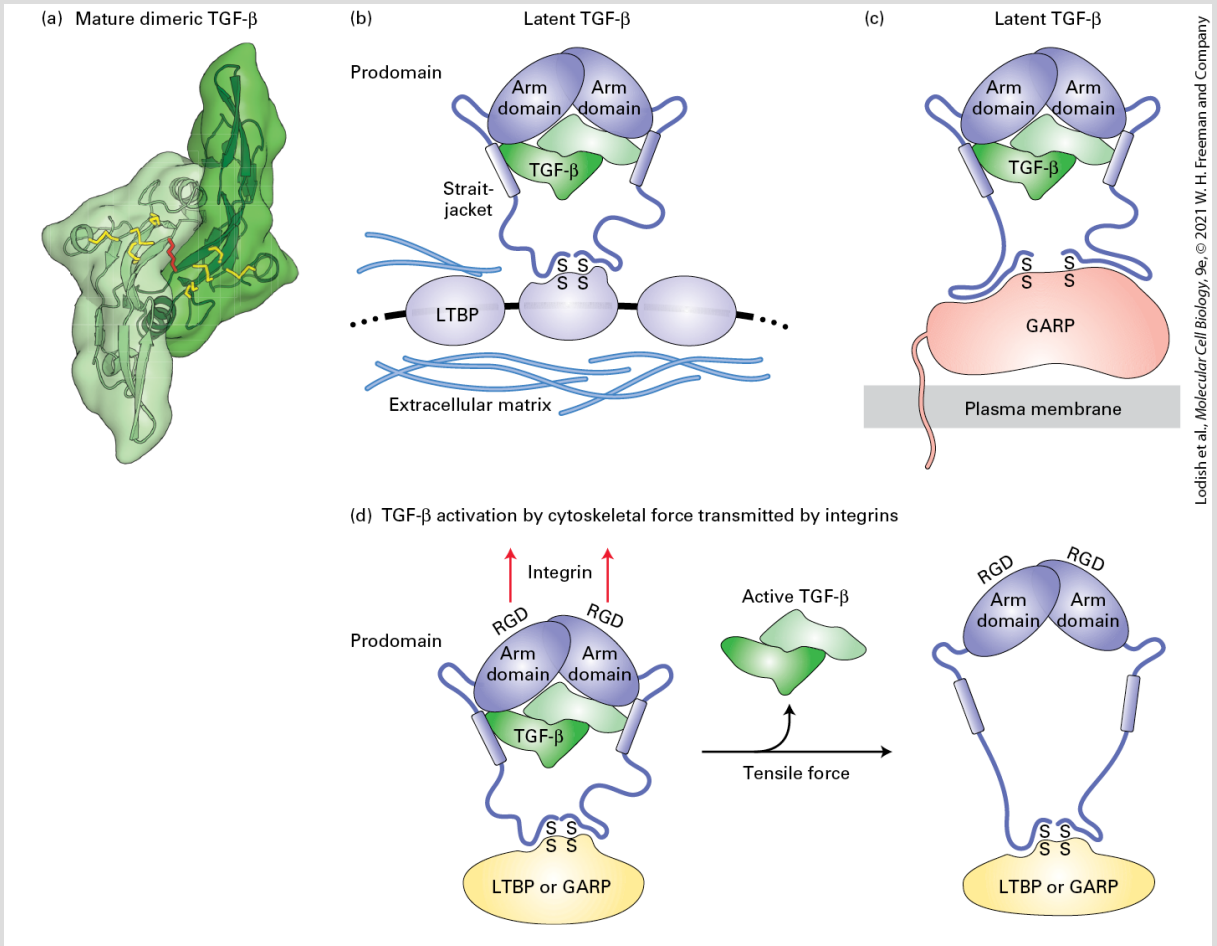


FIGURE 16-22 Model for TGF-β latency and activation. (a) Ribbon diagram structure of a mature TGF-β homodimer. Three intrachain disulfide linkages (yellow) in each monomer stabilize a cystine-knot domain; another disulfide bond (red) links the two monomers. (b, c) TGF-β latency. TGF-β is synthesized with a 250-residue amino-terminal prodomain that in the Golgi is proteolytically cleaved from the mature growth factor but remains noncovalently attached even after secretion. Both the mature growth factor and the prodomain form homodimers. The prodomain, its "arm" and "straitjacket" segments remaining tightly wrapped around the TGF-β growth factor, prevents the latent TGF-β from binding to a cell-surface receptor. Disulfide bonds on the prodomain link it covalently to LTBP or to other proteins in the extracellular matrix (b), or to the GARP protein on the surface of the secreting cell (c), localizing the latent TGF-β near the cell that produced it. (d) TGF-β activation. Certain integrins, perhaps on the surface of a nearby cell, bind to the RGD motifs on the arm segment of the prodomain. Because these integrins are also linked to the rigid actin cytoskeleton on the cytosolic side of the plasma membrane (Chapter 20), movements of the cell exert a traction force against the RGD motif, pulling and elongating

the straitjacket segment of the prodomain. Consequently, the TGF-β dimer is released from the complex and is now able to bind to a cell-surface receptor and initiate signaling.

[Part (a) Data from S. Daopin et al., 1992, *Science* **257**:369, PDBID 2tgi. Parts (c, d) From A. Hinck et al., 2016, *Cold Spring Harb. Perspect. Biol.* **8**:a022103.]

After TGF-β is secreted from the cell, the presence of the prodomain prevents the mature TGF-β protein from binding to its cell-surface receptors; the TGF-β growth factor is thus said to be *latent*. In addition, the prodomain forms linkages that localize the latent TGF-β to the surface of the secreting cell, to a nearby cell, or to the extracellular matrix. Specifically, the side chain of a cysteine residue near the N-terminus of the prodomain becomes disulfide linked to one of several types of proteins in the extracellular matrix, such as *latent TGF-β binding protein* (*LTBP*; Figure 16-22b) or to a protein called *GARP* on the surface of the secreting cell (Figure 16-22c).

Several mechanisms can release the mature TGF-β dimer from the inhibitory prodomain, enabling the active TGF-β dimer to rapidly bind to nearby TGF-β receptors and initiate highly localized autocrine or paracrine signaling (Figure 16-22d). These mechanisms include cleavage of the prodomain by extracellular proteases and mechanical stress that pulls the prodomain away.

Mechanical stress may be generated by the binding of the prodomains to a protein on the surface of a moving cell. The prodomains of some TGF-β family members (e.g., TGF-β1 and TGF-β3) contain a three-amino-acid sequence motif, RGD (Arg-Gly-Asp),

that allows them to bind to certain members of the class of membrane proteins termed *integrins* on the surface of a nearby cell (discussed in Chapter 20); integrins bind to RGD motifs in many extracellular-matrix proteins. During cell migration, for example, a tensile force is generated due to movement of the cell containing the integrin relative to the extracellular matrix containing LTBP and the prodomain/TGF-β dimer. The force physically distorts the prodomain and frees the active TGF-β dimer, initiating TGF-β signaling.

Three Separate TGF-β Receptor Proteins Participate in Binding TGF-β and Activating Signal Transduction

The approach taken to identify the receptors for TGF-β1 is representative of typical biochemical approaches to identifying receptors (see Section 15.2). Investigators first labeled the purified TGF-β1 protein with the radioisotope iodine-125 (^{125}I; see Chapter 3) using conditions that covalently link ^{125}I to exposed tyrosine residues. The ^{125}I-labeled TGF-β1 (^{125}I-TGF-β1) was then incubated at 4 °C with cultured fibroblast cells that express the receptor. At 4 °C ligands bind to receptors, but endocytosis is inhibited. At the end of the incubation, unbound ^{125}I-TGF-β1 in the extracellular fluid was washed away and the cells with receptor-bound ^{125}I-TGF-β1 were treated with a chemical that covalently cross-links proteins that are

physically close to one another, including ligands bound to their receptors. The cells were then disrupted and the investigators purified any proteins (and cross-linked protein complexes) that were radioactively labeled. Analysis of the purified proteins showed that ^{125}I-TGF-β1 was cross-linked primarily to three different polypeptides with molecular weights of 55, 85, and 280 kDa, referred to as *TGF-β receptors RI*, *RII*, and *RIII*, respectively.

Many additional experiments established that RI and RII are each homodimeric proteins; each polypeptide chain has both a single transmembrane segment and a serine/threonine kinase as part of its cytosolic domain. RI and RII together are sufficient to serve as a functional heterotetrameric TGF-β receptor. RIII is a transmembrane cell-surface protein expressed in some but not all cell types. Because RIII contains large, covalently bound oligosaccharide chains, it is a member of the class of glyco (sugar containing) proteins called *proteoglycans* (see Figure 20-33). The function of RIII for TGF-β receptors is similar to that of the α subunit of the IL-2 receptor described above (Figure 16-19b); RIII binds and concentrates TGF-β molecules near the cell surface, facilitating their subsequent binding to the signaling receptors RII and RI (Figure 16-23, step **1a**). As most cells lack RIII, TGF-β binds directly to the RII dimer (Figure 16-23, step **1b**). This step takes place before the RII dimer and RI bind each other (Figure 16-23, step **2**). The complex is inactive until RI binds.

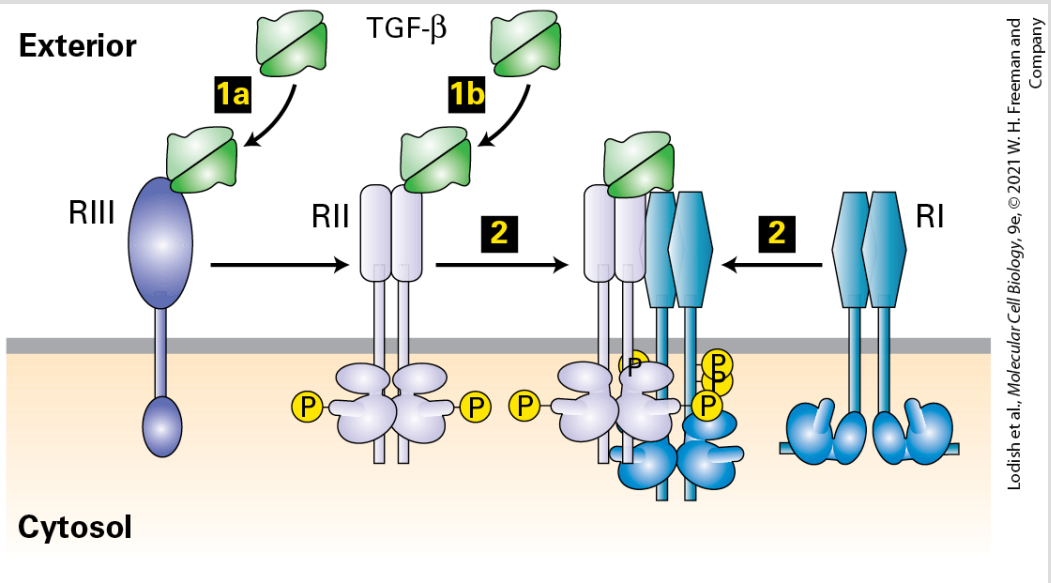


FIGURE 16-23 Activation of the TGF-β receptor kinases. Step **1a**: In cells that express the type III TGF-β receptor (RIII), TGF-β binds first to this receptor, which concentrates TGF-β near the cell surface and thus facilitates the binding of TGF-β to one of the two type II TGF-β signaling receptors (RII) in a homodimeric complex. Step **1b**: In most cells, RIII is not expressed and TGF-β binds directly to one molecule of RII, a constitutively active kinase. Step **2**: Ligand-bound RII recruits a type I TGF-β receptor (RI), such that one monomer of the TGF-β dimer binds simultaneously to one RII and one RI protein. The constitutively active RII protein kinase then phosphorylates serine and threonine residues in the juxtamembrane segment of the RI receptor, releasing the inhibition of RI kinase activity. The now active type I receptor kinase then phosphorylates a Smad transcription factor.

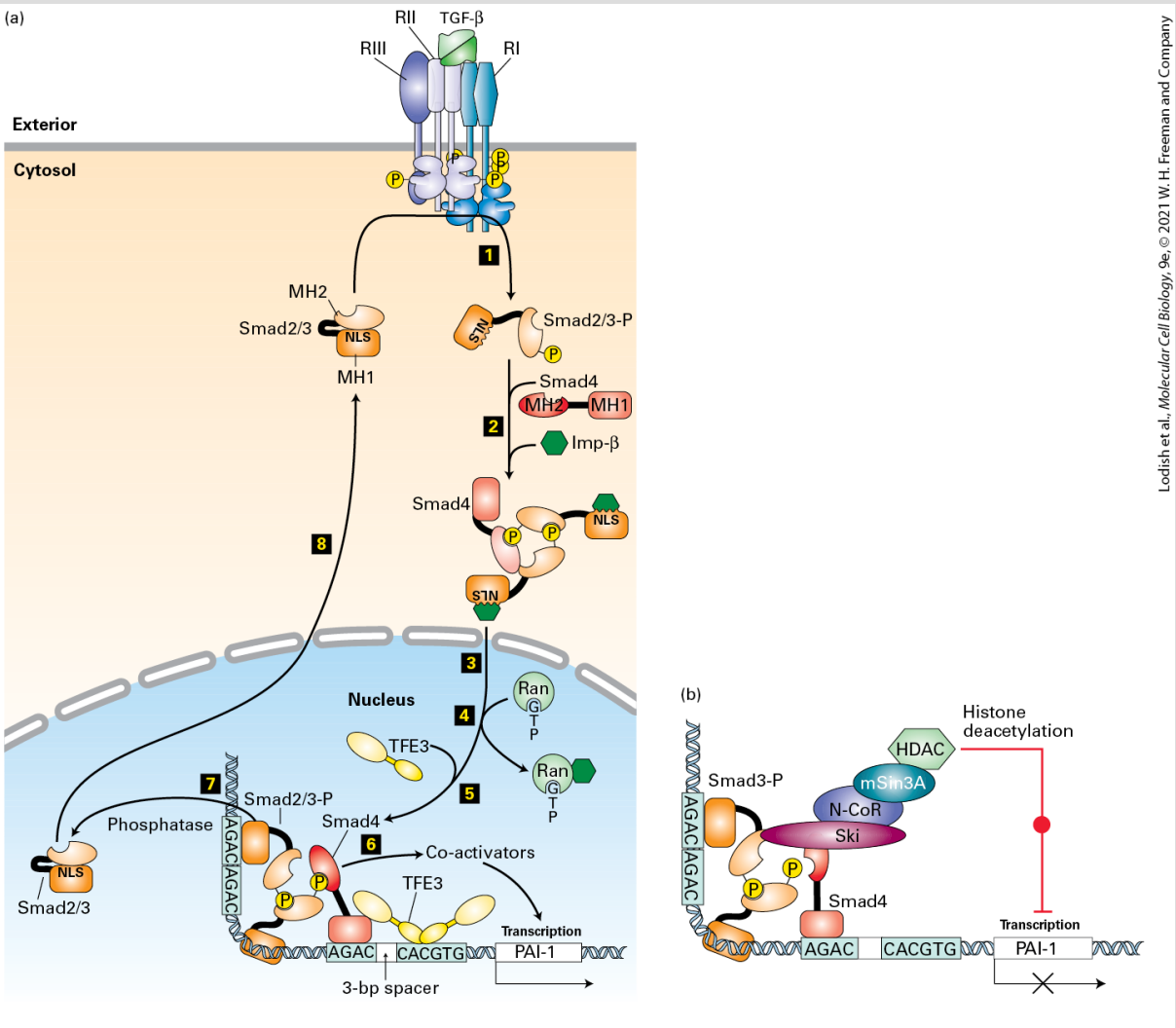
In the tyrosine kinase and cytokine receptors we have already discussed, ligand binding induces a substantial increase in kinase activity, which is required for activating intracellular signal transduction pathways. In contrast, the dimeric RII subunits of the TGF-β receptor exhibit *constitutive* kinase activity; that is, their kinase domains are enzymatically active even when there is no TGF-

β bound. Active RII kinase activity is not sufficient to generate an intracellular signal. However, the binding of TGF-β to RII generates a new molecular surface at the TGF-β–RII interface that binds to a dimeric RI receptor (Figure 16-23, step **2**), generating a complex that contains one dimeric TGF-β molecule and one each of the RI and RII receptor homodimers — another example of ligand-induced receptor hetero-oligomerization. RI will not bind to RII unless RII is bound to TGF-β. An RII subunit in the complex then phosphorylates serine and threonine residues in a highly conserved sequence of the attached RI subunit adjacent to the cytosolic face of the plasma membrane, thereby activating RI kinase activity. The RI kinase then sets in motion the downstream intracellular signaling pathway.

Activated RI TGF-β Receptors Phosphorylate Smad Transcription Factors

The transcription factors downstream of the TGF-β receptors are called *Smads*. Three types of Smad proteins function in the TGF-β signaling pathway: *R-Smads* (receptor-regulated Smads; Smad2 and Smad3), *co-Smads* (Smad4), and *I-Smads* (inhibitory Smads; Smad7). Smad2 and/or Smad3 become phosphorylated by active RI receptors and together with Smad4 move into the nucleus to activate gene transcription; in contrast, I-Smads block the TGF-β signaling pathway, as discussed at the end of this section.

As illustrated in Figure 16-24a, all R-Smads share a similar domain architecture comprising an N-terminal MH1 domain, a central proline-rich linker, and a C-terminal MH2 domain. The MH1 domain contains a DNA-binding segment as well as a *nuclear-localization signal (NLS)*. However, when R-Smads are in their inactive, nonphosphorylated state, the NLS is masked so that it cannot direct the Smad into the nucleus (see Figure 13-35), and the MH1 and MH2 domains associate in such a way that they could not bind to DNA or to a co-Smad. Phosphorylation of two serine residues in a common Ser-X-Ser motif at the extreme C-terminus of the MH2 domain by the activated RI subunits of the TGF-β receptor separates the two domains, exposing the NLS for binding to importins and nuclear import (see Figure 13-36). Simultaneously, the two serines in each Smad2 or Smad3 that were phosphorylated by the TGF-β RI receptor kinase bind to phosphoserine-binding sites in the MH2 domains of a Smad4 or Smad2/3 protein, forming a stable complex containing two molecules of phosphorylated Smad3 (or Smad2) and one molecule of the co-Smad (Smad4). (Smad4 does not get phosphorylated but is essential to form a functional heterotrimeric Smad complex.) The bound importin then mediates translocation of the R-Smad/co-Smad complex into the nucleus. After importin dissociates inside the nucleus, the Smad3/Smad4 (or Smad2/Smad4) complex binds to other transcription factors to activate transcription of specific target genes, for example, the protease inhibitor gene PAI-1 in Figure 16-24a.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 16-24 Activation of the TGF-β/Smad signaling pathway and its inhibition. (a) Activation. Step **1**: The activated type RI receptor kinase (see Figure 16-23) phosphorylates Smad2 or Smad3 (shown here as Smad2/3), causing a conformational change that unmasks its nuclear-localization signal (NLS). Step **2**: Two phosphorylated molecules of Smad2/3 bind to one co-Smad (Smad4) molecule, which is not phosphorylated, and to an importin, forming a multiprotein cytosolic complex. Steps **3** and **4**: After the entire complex translocates into the nucleus, Ran·GTP causes dissociation of the importin, as discussed in Chapter 13. Step **5**: One or more nuclear transcription factors (e.g., TFE3) then associates with the Smad2/3/Smad4 complex, forming an activation complex that cooperatively binds to a regulatory sequence of a target gene. Step **6**: This complex then recruits transcriptional co-activators and induces gene transcription (see Chapter 8). Smad2/3 is eventually dephosphorylated by a nuclear phosphatase (step **7**) and escorted by an exportin (not shown) returns to the cytosol through a nuclear pore (step **8**), where it can

be reactivated by another TGF-β receptor complex. Shown is the activation complex for the gene encoding plasminogen activator inhibitor (PAI-1); similar transcription complexes activate expression of genes encoding other extracellular-matrix proteins such as fibronectin. (b) Ski-mediated inhibition of transcription mediated by the Smad complex. Ski, whose synthesis is induced by the Smad2/3/Smad4 complex formed after TGF-β addition, represses Smad function by binding directly to Smad4. Since the Ski-binding domain on Smad4 significantly overlaps with the Smad4 MH2 domain required for binding the phosphorylated C-terminus of Smad3, binding of Ski disrupts the normal interactions between Smad3 and Smad4 necessary for transcriptional activation. In addition, Ski recruits the protein N-CoR, which binds directly to mSin3A; in turn, mSin3A interacts with a histone deacetylase (HDAC), an enzyme that promotes histone deacetylation on nearby promoters and enhancers, repressing gene expression (see Chapter 8). As a result of both processes, transcription activation induced by TGF-β and mediated by Smad complexes is shut down. The related protein SnoN functions similarly to Ski in repressing TGF-β signaling.

[Part (a) Data from S. Daopin et al., 1992, *Science* **257**:369, PDB ID 2tgi; A. Moustakas and C.-H. Heldin, 2009, *Development* **136**:3699; and D. Clarke and X. Liu, 2008, *Trends Cell Biol*. **18**:430. Part (b) Data from J. Deheuninck and K. Luo, 2009, *Cell Res.* **19**:47.]

An R-Smad's time in the nucleus is limited to prevent overstimulation of the cell. Within the nucleus, R-Smads are further modified by phosphorylation of their linker domains, acetylation of their MH1 domains, and dephosphorylation of the C-terminal serines by nuclear phosphatases. Collectively, these many modifications result in diminution of transcriptional activity and ultimately in dissociation of the R-Smad/co-Smad complex and export of the Smads from the nucleus *via* exportin proteins. Thus the concentration of active Smads within the nucleus closely reflects the levels of activated TGF-β receptors on the cell surface, allowing the regulation of gene transcription to closely follow the level of active TGF-β in the environment.

BMP proteins, which also belong to the TGF-β family, bind to and activate a different set of receptors that are similar to the TGF-β RI and RII proteins but phosphorylate other R-Smads, Smad1, Smad5, and Smad8. Two of these phosphorylated Smads then form a trimeric complex with Smad4, and this Smad complex activates different transcriptional responses from those induced by the TGF-β receptors.

The R-Smad/co-Smad Complex Activates Expression of Different Genes in Different Cell Types

Virtually all mammalian cells secrete at least one TGF-β isoform, and most have TGF-β receptors on their surface. However, the cellular responses induced by TGF-β vary among cell types, depending on the transcription factors that can partner with the R-Smad/co-Smad complex. In epithelial cells and fibroblasts, for example, TGF-β induces expression of extracellular-matrix proteins (e.g., fibronectins and collagens; see Chapter 20); it also induces expression of proteins that inhibit serum proteases, which otherwise would degrade these extracellular-matrix proteins. The inhibition of serum proteases stabilizes the matrix, allowing cells to form stable tissues. In many cells the R-Smad/co-Smad complex promotes expression of genes encoding other proteins such as $p15^{INK4B}$, which arrests the cell cycle at the G_1 stage and thus blocks

cell proliferation (see Chapter 19). More generally, binding of the R-Smad/co-Smad complex to DNA requires other transcription factors to bind at adjacent sites in the DNA; often these transcription factors are master transcription factors that determine the identity of a cell during its development. Equally important, in order for the R-Smad/co-Smad complex to bind to a DNA regulatory region and activate a given gene, the DNA binding segment must be in an active "open" chromatin conformation (see Figure 16-2).

Loss of TGF-β signaling plays a key role in the early development of many cancers. Many human tumors contain inactivating mutations in either TGF-β receptors or Smad proteins and thus are resistant to growth inhibition by TGF-β (see Figure 25-18). Most human pancreatic cancers, for instance, contain a deletion in the gene encoding Smad4 and thus cannot induce cell-cycle inhibitors such as $p15^{INK4B}$ in response to TGF-β. In fact, Smad4 was originally called *DPC* (*d*eleted in *p*ancreatic *c*ancer). Retinoblastoma, colon and gastric cancer, hepatomas, and some T- and B-cell malignancies are also unresponsive to TGF-β growth inhibition. This loss of responsiveness correlates with loss of RI or RII; responsiveness to TGF-β can be experimentally restored by recombinant expression of the "missing" protein. Loss-of-function mutations in Smad2 also commonly occur in several types of human tumors. ■

Negative Feedback Loops Restrain TGF-β/Smad Signaling

As discussed previously in the context of other signaling pathways, the response to most growth factors and other signaling molecules decreases with time, a phenomenon called *desensitization*. This response is adaptive because it prevents overreaction and makes fine-tuned control of cellular responses possible. Two cytosolic proteins called *SnoN* and *Ski* (*Ski* stands for "*S*loan-*K*ettering Cancer *I*nstitute") are induced by TGF-β signaling in virtually all body cells and serve to suppress the TGF-β/Smad signaling pathway. These proteins were originally identified as cancer-causing *oncoproteins* because their expression is elevated in many cancers, including melanomas and certain breast cancers; consequently, growth-inhibitory proteins normally induced by the TGF-β signal transduction pathway are not produced. Indeed, when overexpressed in cultured primary fibroblast cells, Ski or SnoN can cause abnormal cell proliferation, and suppression of Ski in pancreatic cancers reduces tumor growth.

SnoN and Ski trigger abnormal cell proliferation by binding to both the co-Smad (Smad4) and phosphorylated R-Smads (Smad3) after stimulation by TGF-β. SnoN and Ski do not prevent an R-Smad/co-Smad complex from forming or a Smad complex from binding to DNA regulatory regions. Rather, they block a Smad complex bound to DNA from activating transcription, in part by inducing

deacetylation of histones in adjacent chromatin segments. With Smad blocked from activating transcription, the growth-inhibitory effects of TGF-β are diminished (Figure 16-24b). The increased levels of SnoN and Ski proteins induced by TGF-β are thought to dampen the long-term effects of signaling due to continued exposure to TGF-β; this is another example of negative feedback, in which a gene induced by TGF-β signaling, in this case *SnoN*, inhibits further signaling by TGF-β.

Among other proteins induced after TGF-β stimulation are the I-Smads, especially Smad7. Smad7 binds to the activated RI subunit of TGF-β receptors and blocks its ability to phosphorylate Smad2 or Smad3. Like SOCS proteins in the cytokine signaling pathway (see Figure 16-21), Smad7 also recruits an E3 ubiquitin ligase that targets TGF-β receptors for degradation. In these ways, Smad7, like Ski and SnoN, participates in a negative feedback loop: its induction inhibits intracellular signaling stimulated by long-term exposure to the stimulating TGF-β hormone.

KEY CONCEPTS OF SECTION 16.5

The TGF-β Family of Growth Factors, Their Receptor Serine Kinases, and the Smad Transcription Factors They Activate

- The transforming growth factor β (TGF-β) family includes a number of related extracellular signaling molecules that play widespread roles in regulating development.
- TGF-β dimers are stored in an inactive form bound to the inhibitory prodomain on a cell surface or in the extracellular matrix; release of active dimers — by mechanical

stretching or protease digestion of the prodomain — initiates TGF-β signaling (see Figure 16-22).

- TGF-β receptors consist of three types of subunits (RI, RII, RIII). Binding of members of the TGF-β family to the RII subunit, a constitutively active kinase, allows recruitment of the RI subunit to the RII subunit. RII phosphorylates the cytosolic domain of the RI subunit, activating RI's serine/threonine kinase domain. RI then phosphorylates an R-Smad, exposing a nuclear-localization signal (see Figures 16-23 and 16-24).
- After phosphorylated R-Smads bind a co-Smad, the resulting complex translocates into the nucleus, where it interacts with various transcription factors to induce expression of target genes (see Figure 16-24).
- The Smad3/Smad4 complex often induces different genes in different cells by binding to regulatory DNA sequences at sites adjacent to those occupied by cell-specific master transcription factors (see Figure 16-24a).
- TGF-β signaling generally inhibits cell proliferation. Loss of any of several components of this signaling pathway can contribute to abnormal cell proliferation and malignancy.
- Oncoproteins (e.g., Ski and SnoN) and I-Smads (e.g., Smad7) suppress the TGF-β signaling pathway by inhibiting transcription mediated by the Smad2/3/Smad4 complex (see Figure 16-24b).

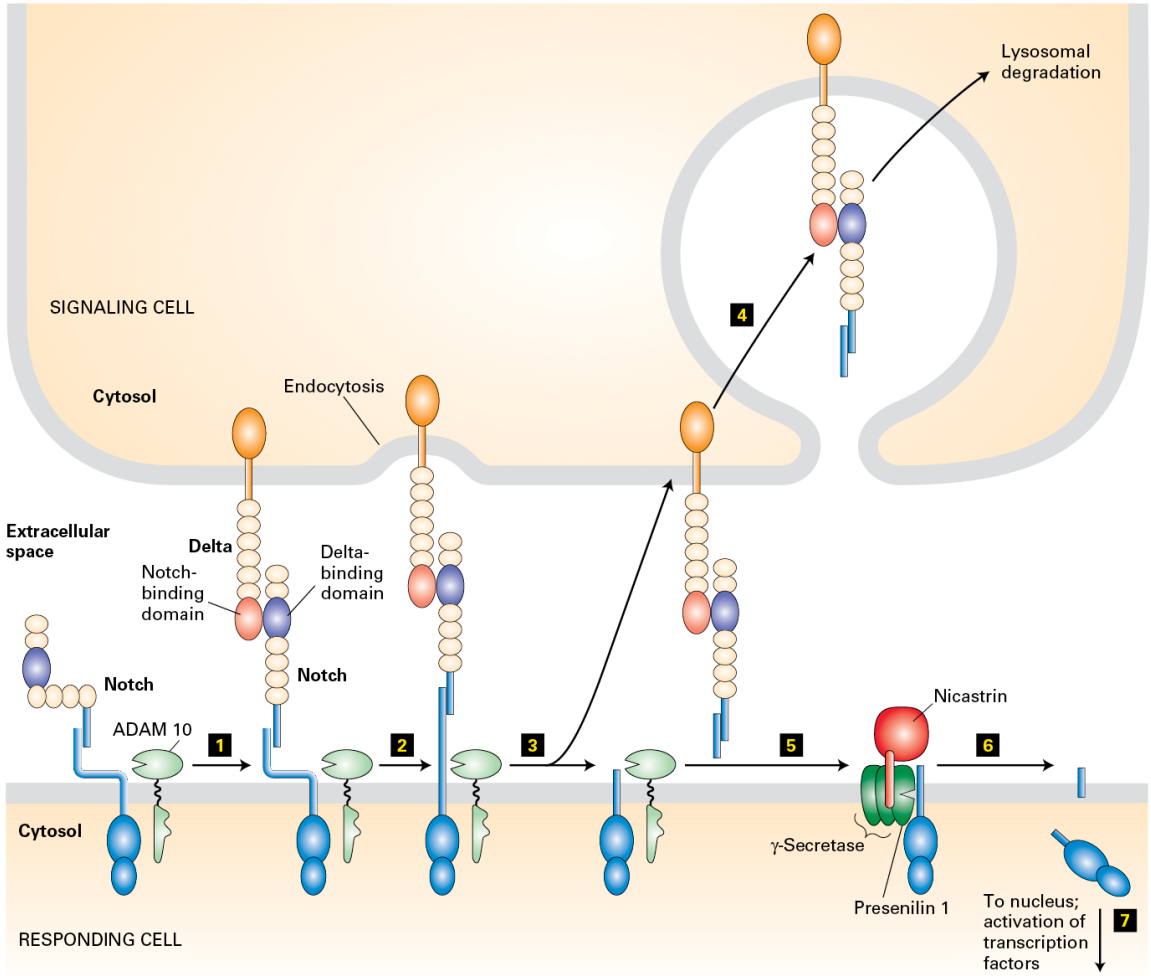
16.6 Signal Transduction Pathways That Utilize Regulated, Site-Specific Protein Cleavage: Notch/Delta and EGF Precursors

All of the signaling pathways we have discussed so far are reversible and so can be suppressed or turned off relatively quickly if the extracellular signal is removed. In this section and the next, we discuss several pathways that are irreversible or only slowly reversible. Many of these pathways trigger key steps in metazoan differentiation, and the irreversibility of a signal transduction pathway can lock the cell into undergoing a specific series of developmental events. In these pathways, a critical protein — a receptor, a transcription factor, or a protein in a signaling cascade that regulates a transcription factor — undergoes regulated cleavage.

On Binding Delta, the Notch Receptor Is Cleaved, Releasing a Component Transcription Factor

We begin by looking at a pathway in which cleavage of a protein releases a domain that affects the transcription of multiple target genes. In the Notch/Delta pathway (Figure 16-25a), the signaling molecule is Delta, an integral membrane protein bound to the

surface of the signaling cell. Delta binds to the extracellular domain of the receptor, Notch, on a neighboring responding cell. The binding of Delta triggers two sequential proteolytic cleavages of Notch, releasing the Notch intracellular domain. That cleavage product then moves into the nucleus and functions as a transcriptional activator.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

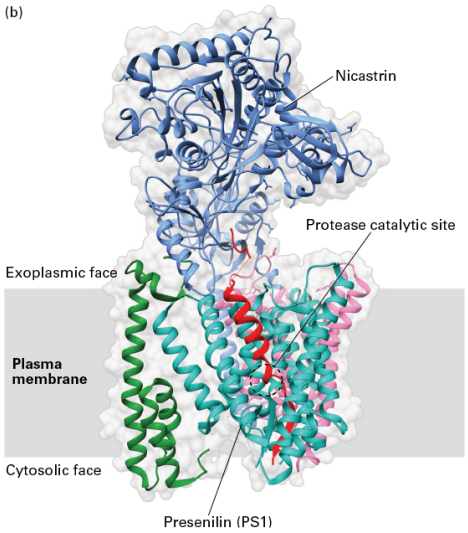


FIGURE 16-25 Notch/Delta signaling pathway. (a) In the absence of Delta, the transmembrane subunit of Notch on the surface of a responding cell is noncovalently

associated with its extracellular subunit; the extracellular domain is folded so that it cannot be cleaved by the cell-surface matrix metalloprotease ADAM 10. Binding of Notch to its ligand Delta on an adjacent signaling cell (step **1**) is followed by endocytosis of Delta by the signaling cell (step **2**), stretching the Notch extracellular domain so that ADAM 10 is able to cleave it (step **3**). The released fragment of the Notch extracellular domain remains bound to Delta and is endocytosed by the signaling cell and degraded in lysosomes (step **4**). Next the four-protein γ-secretase complex binds to the stump generated by ADAM 10 (step **5**), and then the protease in the complex, presenilin 1, catalyzes cleavage of the Notch transmembrane alpha helix near the cytosolic face of the membrane, releasing the cytosolic segment of Notch (step **6**). Following translocation to the nucleus, this Notch segment does not directly bind to DNA but rather interacts with several transcription factors to affect expression of genes that in turn influence the determination of cell fate during development (step **7**). See R. Kovall et al., 2017, *Dev. Cell* **41**:228; and S. Bray, 2016, *Nat. Rev. Mol. Cell Biol.* **17**:772. (b) High resolution structure of the four-protein γ-secretase complex determined by cryoelectron microscopy. A bound Notch transmembrane peptide that contains the intermembrane cleavage site is colored red and the presenilin (PS1) subunit is colored cyan. The protease catalytic site in PS1 is located near the cytosolic surface.

[Part (b) Data from G. Yang et al., 2019, *Nature* **565**:192.]

Both Notch and Delta are cell-surface, single membrane-spanning proteins. Notch also has other ligands, but the molecular mechanisms of Notch activation are the same with each. Notch is synthesized as a monomeric membrane protein in the endoplasmic reticulum. In the Golgi complex, it undergoes a proteolytic cleavage that generates an extracellular subunit and a transmembrane subunit that remain noncovalently associated with each other on the cell surface.

The first of two cleavages of the Notch extracellular domain is performed by ADAM 10, a *metalloprotease* and member of a class of metal-containing enzymes localized to the plasma membrane that cleave the extracellular segments of target proteins at a site near the plasma membrane. ADAM is an abbreviation for *a d*isintegrin *a*nd *m*etalloprotease; a *disintegrin* is a conserved protein domain that binds integrins and disrupts cell-matrix interactions (see Chapter 20). ADAM 10 cannot cleave Notch in the absence of Delta on an adjacent cell, because then the Notch extracellular domain is folded such that ADAM 10 cannot access the protease cleavage site (see Figure 16-25a).

After Delta binds to Notch (Figure 16-25a, step **1**), the Delta in the signaling cell undergoes endocytosis (step **2**). The force accompanying the movement of Delta into the signaling cell stretches the Notch protein on the responding cell, changing its conformation and allowing access by ADAM 10, which cleaves the Notch extracellular domain (step **3**). The Notch extracellular domain remains bound to Delta, is internalized by the signaling cell, and is probably degraded in lysosomes (step **4**).

A second cleavage of Notch follows quickly but this time it occurs within the hydrophobic membrane-spanning region of Notch; it is catalyzed by a transmembrane complex termed *γ-secretase* that contains *presenilin 1* (PS1), the protease enzyme, and three other essential subunits (Figure 16-25a, step **5**, and Figure 16-25b).

Cleavage by the *γ-secretase* complex releases the Notch cytosolic segment, which immediately translocates to the nucleus (Figure 16-25a, step **6**). This Notch intracellular domain does not bind directly to DNA. Rather it interacts with a single DNA-binding transcription factor, CSL (RBPJ in vertebrates) whereby it affects transcription of different target genes depending on the cell type. Its effect, like those of other transcription factors activated downstream of other cell-surface receptors, depends on the constellation of epigenetic chromatin marks and the presence of cell-specific transcription factors; it can both increase and decrease gene expression.

Metalloproteases Catalyze Cleavage of Many Signaling Proteins from the Cell Surface

We next look at a pathway in which the cleavage of a cell surface protein releases a signaling molecule rather than a transcription factor. Many growth factors and other signaling molecules are synthesized as transmembrane proteins whose signal domain extends into the extracellular space. Some such proteins, like Delta described above, are often biologically active but can signal only by binding to receptors on adjacent cells. In other cases, proteolytic cleavage of such proteins can release the soluble, active signaling molecule into the extracellular space; from there the signaling molecule can bind to receptors on nearby cells. The cleavage

producing these signaling molecules is often carried out by ADAMs. The human genome encodes 21 metalloproteases in the ADAM family, but only 12 are known to be catalytically active; many are involved in cleaving the transmembrane precursors of signaling proteins just outside their transmembrane segment, releasing the soluble signaling protein into the extracellular space.

Medically important examples of signaling proteins released from the cleavage of precursors are members of the EGF family, including EGF, HB-EGF, TGF-α, NRG1, and NRG2 discussed earlier (see Figures 16-6 and 16-7), which affect the proliferation and differentiation of many types of body cells. The activity of one or more ADAMs is increased in many cancers, leading to high levels of extracellular EGF-family growth factors. These stimulate the secreting cells (autocrine signaling) or adjacent cells (paracrine signaling) to proliferate inappropriately.

ADAM proteases also are an important factor in heart disease. As we learned in the last chapter, stimulation of β-adrenergic receptors in heart muscle by epinephrine (adrenaline) causes glycogenolysis and an increase in the rate of muscle contraction. Prolonged treatment of heart muscle cells with epinephrine, however, leads to activation of an ADAM by an unknown mechanism. This ADAM protease cleaves the transmembrane precursor of HB-EGF, and the released HB-EGF then binds to EGF receptors on heart muscle cells and stimulates their inappropriate growth. The excessive proliferation of

heart muscle cells can lead to an enlarged but weakened heart — a condition known as *cardiac hypertrophy*, which may cause early death. How cleavage of these proteins by ADAM metalloproteases is regulated and how specificity of substrate cleavage on the cell surface is achieved are still not completely understood. ■

KEY CONCEPTS OF SECTION 16.6

Signal Transduction Pathways That Utilize Regulated, Site-Specific Protein Cleavage: Notch/Delta and EGF Precursors

- On binding to its ligand, Delta, on the surface of an adjacent cell, the receptor Notch protein undergoes two proteolytic cleavages (see Figure 16-25), first by an ADAM protease and then within the plane of the plasma membrane by γ-secretase. The released Notch cytosolic segment then translocates into the nucleus and modulates transcription of target genes critical in determining cell fate during development.
- Many important growth factors and other signaling proteins such as EGFs are synthesized as transmembrane proteins; regulated cleavage of the precursor near the plasma membrane by members of the ADAM protease family releases the active molecule into the extracellular space to signal distant cells.
- Inappropriate cleavage of EGF-family precursors can result in abnormal cell proliferation, potentially leading to cancer, cardiac hypertrophy, and other diseases.

16.7 Signal Transduction Pathways That Utilize Proteasomal Degradation of Signaling Components: Wnt, Hedgehog, and the Many Hormones That Activate NF-κB

Next we discuss signal transduction pathways in which either a transcription factor or an inhibitor of a transcription factor becomes ubiquitinated and then is either cut or totally degraded by the proteasome. We begin with signal transduction pathways induced by the **Wnt** and **Hedgehog (Hh)** proteins, two evolutionarily conserved families of signaling proteins that play key roles in many developmental pathways and often induce expression of genes required for a cell to acquire a new identity or fate. Although Wnt and Hedgehog signaling pathways use different sets of receptors and signaling proteins, they do share similarities, which is why we group them together:

- In the unstimulated state, key transcription factors in both pathways are initially localized in large cytosolic multiprotein complexes; enzymes in the complex add ubiquitin chains to the transcription factors. In the Wnt pathway the transcription factor becomes totally degraded by the proteasome, whereas in

the Hh pathway the transcription factor is cleaved by the proteasome, generating a repressor of transcription.
- Activation of both pathways involves disassembly of these protein complexes, inhibition of cleavage or degradation of the transcription factor, and movement of the full-length, active transcription factor into the nucleus.
- Both pathways regulate cell fate decisions in animal development.

Next we examine the *NF-κB signal transduction pathway* that is activated downstream of many important hormones that affect the immune system. In this case, an inhibitor of a transcription factor, rather than a transcription factor itself, is degraded following ubiquitinylation. In the resting state, the transcription factor termed **NF-κB** is sequestered in the cytosol and bound to an inhibitor. Several signal transduction pathways, activated by different hormone receptors, converge on an enzyme that ubiquitinates and thus causes rapid degradation of the inhibitor. Freed of its inhibitor, NF-κB immediately and vigorously moves to the nucleus and activates transcription of multiple genes. In learning how the NF-κB pathway is activated by one class of surface receptors, we also see a very different function of polyubiquitinylation: the formation of a scaffold to assemble a key signal transduction complex.

Wnt Signaling Prevents Destruction of a Transcription Factor by a Cytosolic

Protein Complex

The components of the Wnt signaling pathways have been conserved throughout the evolution of metazoan organisms and were elucidated mainly through genetic analysis of developmental mutants in *Drosophila*. In vertebrates, mutations in these pathways are thought to trigger several types of cancers. In fact, the first vertebrate *Wnt* gene to be discovered, the mouse *Wnt-1* gene, attracted notice because it was overexpressed in certain mouse mammary cancers. When a mouse retroviral DNA, the mammary tumor virus (MMTV) genome, became inserted near the *Wnt-1* gene, the retrovirus LTR promoter (see Figure 7-13) activated inappropriate expression of the *Wnt-1* gene.

The human genome encodes 19 different Wnt proteins, and many Wnts are essential for critical developmental events, including bone formation, the segmented development of the musculoskeletal system in vertebrates, and the formation and maintenance of skin and hair. Yet all Wnt proteins are thought to use the same set of cell-surface receptors and signal transduction proteins.

Wnt proteins are signaling molecules secreted from the cell and are modified by linkage of a monounsaturated fatty acid, palmitoleic acid, to a serine in the middle of the protein. Like other secreted signaling proteins, Wnt proteins interact with several extracellular and cell-surface proteins and activate multiple downstream signal

transduction pathways. The principal signaling receptor for Wnt proteins is *Frizzled (Fz)*, which contains seven transmembrane α helices. Like the glucagon receptor (see Figure 15-15), Fz has a large extracellular domain that is connected to the first membrane-spanning α helix and contains the major Wnt binding site. Unlike the glucagon receptor, however, as far as is known Fz does not activate a G protein. The palmitoleate attached to the Wnt protein binds to a specific site on the Fz extracellular domain and stabilizes the Wnt-Fz complex. Thus this lipid is central to receptor engagement by Wnt proteins and is one of the few known examples of post-translational modification by a lipid that mediates a ligand-receptor interaction.

At least three different signal transduction pathways can be activated by the binding of different Wnt proteins to particular Fz proteins. The most widespread, canonical Wnt signaling pathway has a second transmembrane receptor, LRP (called *Arrow* in *Drosophila*), that associates with Fz such that Wnt binds simultaneously to both (Figure 16-26). Inactivating mutations in the genes encoding Wnt proteins, Fz, or LRP all have similar effects on the development of mouse embryos, indicating that all three proteins are essential for Wnt signaling.

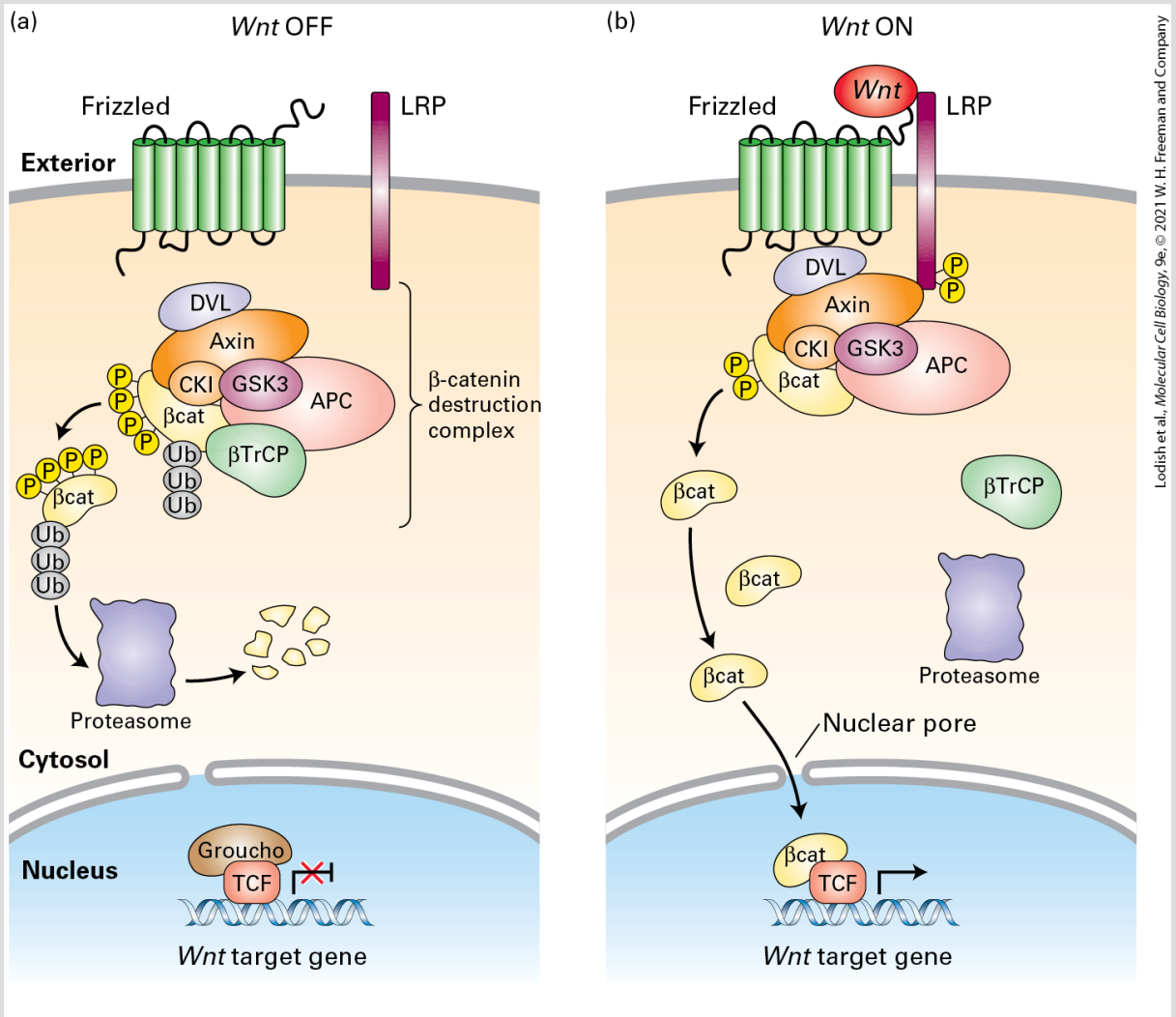


FIGURE 16-26 Canonical Wnt signaling pathway. (a) In the absence of Wnt, the transcription factor TCF is bound to promoters or enhancers of Wnt target genes. Its activator β-catenin is absent from the nucleus, and TCF is associated with transcriptional repressor proteins such as Groucho (Gro) that inhibit gene transcription. The key signaling molecule in the canonical Wnt pathway is the cytoplasmic protein β-catenin, whose stability is controlled by a large multiprotein β-catenin destruction complex. The tumor suppressor protein Axin is the scaffold of the complex, and binds to β-catenin, the tumor suppressor APC, the Dishevelled (Dvl) protein, and two constitutively active protein kinases, glycogen synthase kinase 3β (GSK 3β) and casein kinase 1 (CK1). CK1 and GSK3 sequentially phosphorylate β-catenin at multiple serine and threonine residues. The E3 βTrCP ubiquitin ligase then binds to two phosphorylated residues on β-catenin, leading to ubiquitinylation of β-catenin and its degradation in proteasomes. (b) Wnt proteins bind to a complex of two receptor proteins on the cell surface, Frizzled (Fz) and LRP. Fz proteins have seven

transmembrane segments and an extracellular N-terminal domain that binds Wnt. Binding of Wnt to these receptors induces phosphorylation of the LRP cytosolic domain by several kinases, followed by binding of Axin to phosphorylated LRP and Dvl to Fz. This causes part of the Axin–APC–CK1–GSK3–β-catenin complex to fall apart, preventing phosphorylation of β-catenin by CK1 and GSK3, inhibiting ubiquitination of β-catenin, and allowing β-catenin to accumulate in the cytosol. After translocation to the nucleus, β-catenin binds to TCF, displaces the Gro repressor, and recruits co-activator proteins to activate expression of Wnt target genes. See R. Nusse and H. Clevers, 2017, *Cell* **169**:985; and also the Wnt home page, http://web.stanford.edu/group/nusselab/cgi-bin/wnt.

The central player in the canonical Wnt intracellular signal transduction pathway is a protein called *β-catenin* in vertebrates and *Armadillo* in *Drosophila*. This multi-talented protein functions both as a transcription factor in the Wnt signaling pathway and as a linker protein that connects cell adhesion proteins in the plasma membrane to the actin cytoskeleton (see Figure 20-14). In the absence of a Wnt signal, the β-catenin molecules that are not involved in membrane-cytoskeleton interactions ("free" β-catenin) are targeted for degradation by a cytosolic *destruction complex* that is held together by the scaffold protein Axin. This complex contains the tumor suppressor adenomatous polyposis coli (APC) protein, so named because its loss may result in colorectal cancer. The large size of the destruction complex observed by fluorescence microscopy suggests that it contains many copies of each component.

In the resting state, two constitutively active kinases in the destruction complex, casein kinase 1 (CK1) and GSK3, sequentially

phosphorylate β-catenin on multiple serine and threonine residues; some of these phosphorylated residues serve as binding sites for an E3 ubiquitin-ligase protein named *βTrCP*. β-Catenin is then ubiquitinylated by βTrCP and rapidly degraded by the 26S proteasome (Figure 16-26a; for more on ubiquitinylation, see Figure 3-32). Degradation prevents free β-catenin from entering the nucleus and activating transcription of Wnt target genes.

The complete mechanism by which Wnt signaling blocks the degradation of β-catenin, enabling it to act as a transcription factor, has not yet been fully elucidated. The formation of the Wnt/Fz/LRP complex at the cell surface causes the adaptor protein Dishevelled (Dvl), a component of the destruction complex, to bind to Fz and Axin to bind to LRP. This results in the destabilization and inactivation of the destruction complex — no phosphorylation or subsequent polyubiquitinylation of β-catenin. Thus free β-catenin accumulates; it transfers into the nucleus and activates transcription of Wnt target genes (Figure 16-26b).

Aberrant hyperactive Wnt signaling and abnormally high levels of free β-catenin is implicated in the progression of many cancers; the Wnt signaling pathway is hyperactive in more than 90 percent of human colon cancers (see Chapter 25). This observation provided one of the earliest clues that β-catenin can activate many growth-promoting genes. Inactivating mutations in genes encoding the tumor suppressors APC and Axin, components of the destruction

complex, are found in multiple types of human cancers, as are mutations in β-catenin phosphorylation sites for the destruction complex kinases GSK3 or CK1; these mutations either reduce the formation of the destruction complex or reduce the phosphorylations on β-catenin (see Figure 16-26a). Thus they reduce degradation of β-catenin and allow free β-catenin to activate gene expression even in the absence of the normal Wnt signal. ■

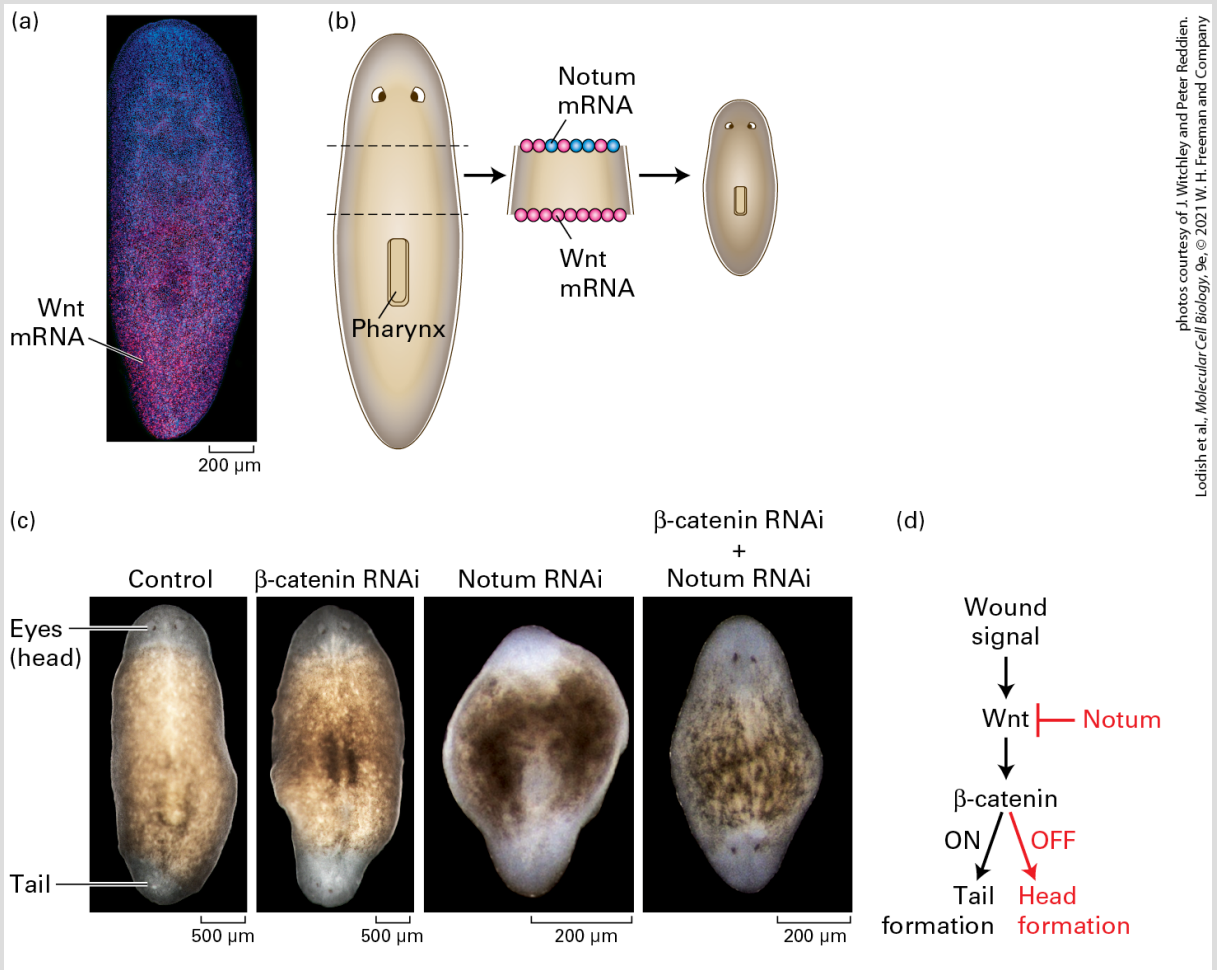
Among the Wnt target genes are many that also control Wnt signaling, indicating a high degree of feedback regulation. The importance of β-catenin stability and localization means that Wnt signals affect a critical balance between the three pools of β-catenin in the cell: at the membrane-cytoskeleton interface, in the cytosol, and in the nucleus.

Concentration Gradients of Wnt Protein Are Essential for Many Steps in Development

Wnts are proteins secreted from the cell; however, in part because of the hydrophobic lipid that is covalently linked to these proteins, these proteins diffuse only a short distance from a signaling cell and generally have only localized effects. As Wnt diffuses farther and farther away from secreting cells, its concentration decreases. Different Wnt concentrations induce different fates in target cells:

cells that detect a large amount of Wnt turn on certain genes and form certain structures; cells that detect a smaller amount turn on different genes and so form different structures. Signals that induce different cell fates depending on their concentration at their target cells are referred to as **morphogens**. In the development of the vertebrate nervous system, for instance, high levels of Wnt signaling specify that a certain progenitor cell becomes a posterior neuron while low levels of Wnt signaling specify that the progenitor cell becomes an anterior neuron.

A striking example of a Wnt gradient regulating tissue pattern occurs in the adult planarian *Schmidtea mediterranea* (see Figure 1-23e); Wnt mRNA and protein are expressed in a posterior-to-anterior gradient (Figure 16-27a). If the planarian's head is cut off, a new one regenerates within 14 days; a normal, albeit smaller, worm is regenerated. Similarly, after removal of the tail, a new one regenerates.



photos courtesy of J. Witchley and Peter Reddien.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 16-27 Gradients of Wnt are essential for normal regeneration of a head and a tail by planaria. (a) As shown by in situ hybridization, the *WntP-2* gene (pink dots) is expressed in a posterior-to-anterior gradient in adult planaria. J. Witchley et al., 2013, *Cell Rep.* **4**:633. (b) As indicated in the diagram, a small piece excised from the middle of the body of the planarian *S. mediterranea* was placed in culture. In situ mRNA hybridization performed after 4 days indicated that Wnt mRNA (pink dots) was expressed in cells at both the anterior and posterior wound sites, but that the Wnt inhibitor Notum (blue dots) was expressed only at the anterior wound site. Thus a posterior-to-anterior gradient of Wnt protein is formed. (c) After 14 days in culture, the excised body piece has regenerated a normal, albeit smaller, worm, with a head, easily visualized by the two eyes, from the anterior wound site and a tail from the posterior. Treatment of the excised body piece with an inhibitory RNA specific for β-catenin results in regeneration of a two-headed planarian, whereas treatment with an inhibitory RNA specific for Notum results in regeneration of a two-tailed planarian. Treatment of the excised body piece with two inhibitory RNAs, one specific for β-catenin and the other for Notum, results in a phenotype similar to that caused

by the loss of β-catenin alone: a two-headed planarian is regenerated. (d) These experiments give rise to a model in which β-catenin, stabilized by addition of Wnt to cells, causes expression of genes that promote tail formation; inhibition of Wnt/β-catenin signaling by Notum causes a head to be formed. [Part (a) photo: Jessica Witchley and Peter Reddien. Part (c, *left two photos*) Republished with permission from AAAS, from C. P. Petersen and P. W. Reddien, 2008, "Smed-β-catenin-1 Is Required for Anteroposterior Blastema Polarity in Planarian Regeneration," *Science* **319**(5861):327–330; photos courtesy of J. Witchley and Peter Reddien. Part (c, *right two photos*) Republished with permission from AAAS, from C. P. Petersen and P. W. Reddien, 2011, "Polarized *notum* Activation at Wounds Inhibits Wnt Function to Promote Planarian Head Regeneration," *Science* **332**(6031):852–855; permission conveyed through Copyright Clearance Center, Inc.; photos courtesy of J. Witchley and Peter Reddien.]

Strikingly, a small body piece taken from the middle of the animal regenerates a normal head from the anterior (head-facing) wound and a tail from the posterior wound (Figure 16-27b). Wnt signaling is required for the choice of head versus tail made at the wounds in these fragments. The gene *Wnt-1* is expressed at all planarian wounds. By contrast, the gene encoding the secreted extracellular enzyme *Notum* is expressed only at anterior-facing wounds. Notum inhibits Wnt signaling by cleaving off palmitoleic acid, the fatty acid that is covalently attached to Wnt and, as we learned above, is essential for Wnt signaling.

RNA interference (RNAi) can be used to inhibit either the *β-catenin* gene, eliminating Wnt signaling, or the *notum* gene, or both genes. As shown in Figure 16-27c, in the absence of all Wnt signaling, a two-headed planarian is regenerated. In contrast, in the absence of Notum, a two-tailed planarian is regenerated; in this case, Wnt is

present and active at both the anterior and posterior wounds, and thus two tails are formed. In the absence of both Wnt signaling and Notum, the phenotype is similar to that caused by the loss of Wnt alone: a two-headed planarian is regenerated. Wnt signaling is absent at both the anterior and posterior wounds, inducing head regeneration at both sites, and the absence of Notum is irrelevant in the absence of Wnt. These experiments show that Wnt/β-catenin signaling promotes tail regeneration and inhibits head regeneration; the absence of Wnt signaling induces head formation at wounds (Figure 16-27d).

In Chapter 22, we will learn that mature planarians contain pluripotent stem cells termed *neoblasts* that can differentiate into any body cell type. Gradients of Wnt protein play a major role in instructing neoblasts to differentiate into the appropriate cell types of tissues that constitute the planarian head or tail. As evidence, RNA interference (RNAi) has been used to specifically inhibit *β-catenin* throughout an uninjured animal. Even in these uninjured animals, inhibition of Wnt signaling by this means causes heads to appear around the body. These are consequences of defects in the pattern of tissue replacement and repair mediated by stem cells (see the chapter opening figure).

Hedgehog Signaling Relieves Repression of Target Gene Expression

The Hedgehog (Hh) signaling pathway is similar to the Wnt pathway in that two membrane proteins, one with seven membrane-spanning segments, are required to receive and transduce a signal. The Hh pathway also involves the disassembly of an intracellular complex containing a transcription factor. Hh signaling differs from Wnt signaling in that its two membrane receptors move between the plasma membrane and intracellular vesicles, and that in mammalian cells, the Hh signaling pathway is restricted to the single primary cilium that protrudes from the cell surface of most vertebrate cells.

Like Wnt proteins, Hh proteins contain a covalently attached lipid and often act as morphogens and signal to influence the fate of nearby cells. Hh signaling plays essential roles in the development of numerous organ systems in animals. For instance, Hh signaling specifies the identities of neurons made along the ventral (belly) to dorsal (back) axis of the embryonic neural tube, which makes the spinal cord and brain. Hh signaling also regulates lung morphogenesis and hair follicle formation. One of the three mammalian Hh proteins, Sonic Hedgehog (Shh), is essential for normal patterning of the limbs; abnormal expression of Shh in an anterior region of the developing limb, in addition to its normal expression in the posterior domain, leads to polydactyly (extra digits).

Processing of Hh Precursor Protein

Hedgehog proteins are formed from a precursor protein with autoproteolytic activity that enables the protein to cut itself in half while still in the endoplasmic reticulum. The cleavage produces an N-terminal fragment, which is subsequently secreted to signal other cells, and a C-terminal fragment, which is degraded. As shown in Figure 16-28, cleavage of the precursor is accompanied by covalent addition of the lipid cholesterol to the new carboxyl terminus of the N-terminal fragment. A second modification to Hedgehog, the addition of a palmitoyl group to the N-terminus, makes the protein even more hydrophobic.

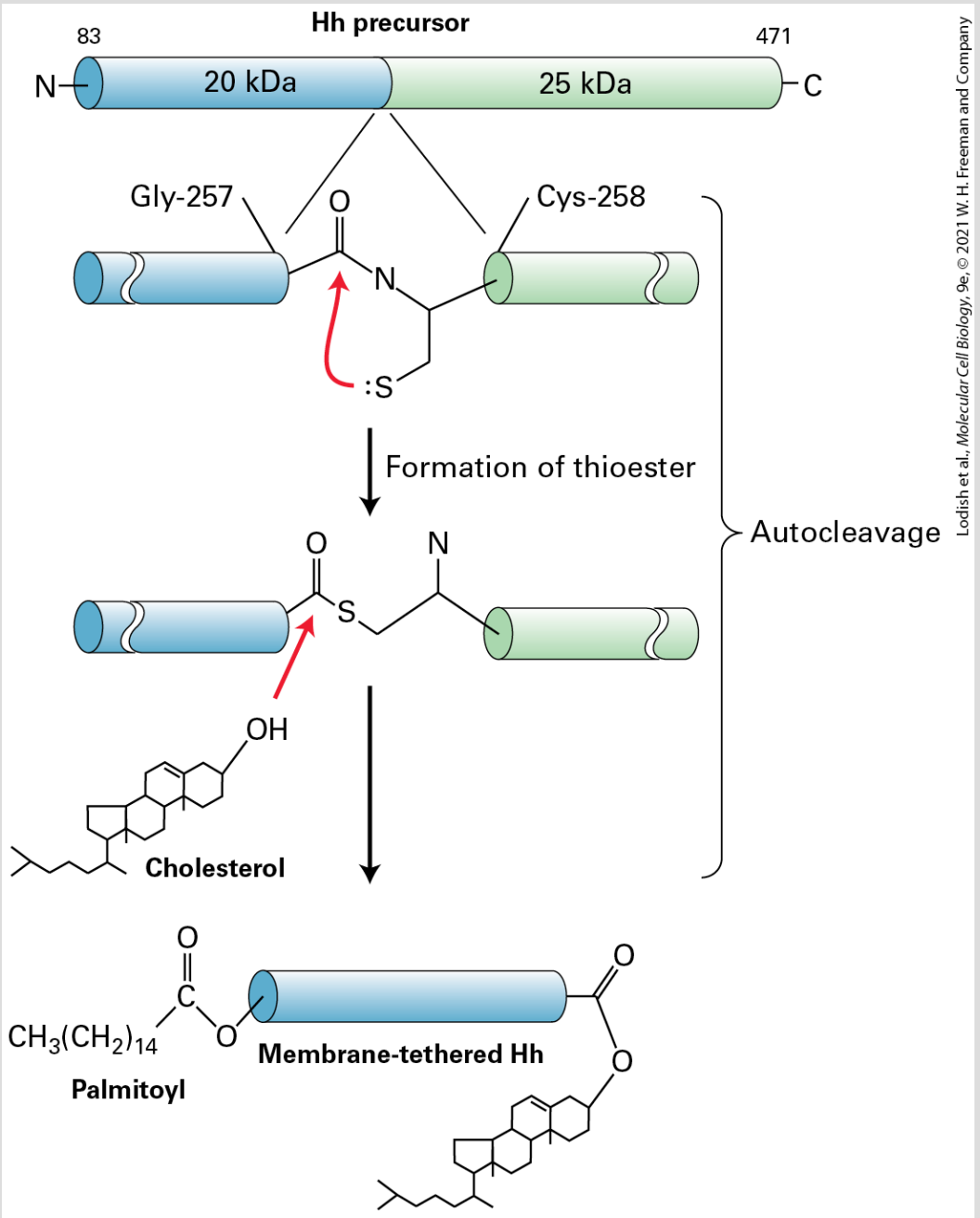


FIGURE 16-28 Processing of Hedgehog (Hh) precursor protein. Cells synthesize a 45-kDa Hh precursor, which in the endoplasmic reticulum undergoes a nucleophilic attack by the thiol side chain of cysteine 258 (Cys-258) on the carbonyl carbon of the adjacent residue glycine 257 (Gly-257), forming a high-energy thioester intermediate. Enzyme activity in the C-terminal domain then catalyzes the formation of an ester bond between the hydroxyl group of cholesterol and glycine 257, cleaving the precursor into two

fragments. The N-terminal signaling fragment (blue) retains the cholesterol group and is also modified by the addition of a palmitoyl group to the N-terminus. Once Hh has been secreted, the two hydrophobic anchors may tether the processed Hh protein to the plasma membrane. See P. Thérond, 2012, *Curr. Opin. Cell Biol.* **24**:173.

Despite its hydrophobicity Hh proteins can travel relatively long distances — up to 300 μm in the developing vertebrate limb. Both Hh and Wnt are often anchored to the phospholipid monolayers of extracellular lipoprotein particles (see Figure 14-27 for the structure of a typical lipoprotein) via their attached lipid groups, and while anchored to these particles they can diffuse in extracellular spaces. However, the majority of the Hh proteins produced by a cell can remain bound to its plasma membrane; in such cases, Hh signals mainly by cell-cell contact. As with Wnt proteins, spatial restriction plays a crucial role in constraining the effects of Hh proteins.

The Hh Receptors Patched and Smoothened and the Downstream Signaling Pathway Were Initially Elucidated by Genetic Studies of *Drosophila* Development

Two membrane proteins, Smoothened (Smo) and Patched (Ptc), are required to receive a Hedgehog signal and transduce it to the cell

nucleus. Smo has seven membrane-spanning α helices, is a member of the GPCR superfamily of proteins, and is related in sequence to the Wnt receptor Fz. Ptc contains 12 transmembrane α helices and is most similar structurally to the Niemann-Pick C1 (NPC1) protein, a cholesterol-transporting member of the ABC superfamily of membrane transport proteins (see Table 11-3). Ptc is the receptor to which Hh binds, but Smo is the membrane protein that initiates the signal to the nucleus. One can think of the pathway as two series of steps: The binding of Hh to Ptc initiates a series of steps that brings Smo to the plasma membrane, and from there Smo initiates another series of steps that frees a transcription factor from a cytosolic complex.

Figure 16-29 depicts the current model of the Hedgehog pathway in *Drosophila*, where the pathway is known in the most detail. Early evidence supporting this model came from the study of fly embryos with loss-of-function mutations in the *hedgehog (Hh)* or *smoothened (Smo)* genes. Both types of mutant embryos have very similar developmental phenotypes; the name *hedgehog* came from the appearance of Hh mutant embryos, which were covered by an array of disorganized hairlike bristles that resembled the spines on hedgehogs, as a result of a defect in the patterning of body segments. Moreover, both the *Hh* and *Smo* genes are required to activate transcription of the same target genes (e.g., *patched* and *wingless*) during embryonic development. In contrast, loss-of-function mutations in the *patched* (*Ptc*) gene produce a quite

different phenotype, one similar to the effect of flooding the embryo with Hedgehog protein. These findings suggested that, in the absence of Hh, Ptc represses target genes by inhibiting a signaling pathway that requires Hh and Smo and is needed for gene activation. The additional observation that Smo is required for the transcription of Hh target genes in mutants lacking *patched* function places Smo downstream of Ptc in the Hh pathway. Together with biochemical experiments showing that Hh binds directly to Ptc, this genetic evidence indicates that, in the absence of Hh, Ptc blocks Smo-mediated activation of the transcription of Hh target genes. Binding of Hh to Ptc relieves this inhibition and permits the transcription of target genes.

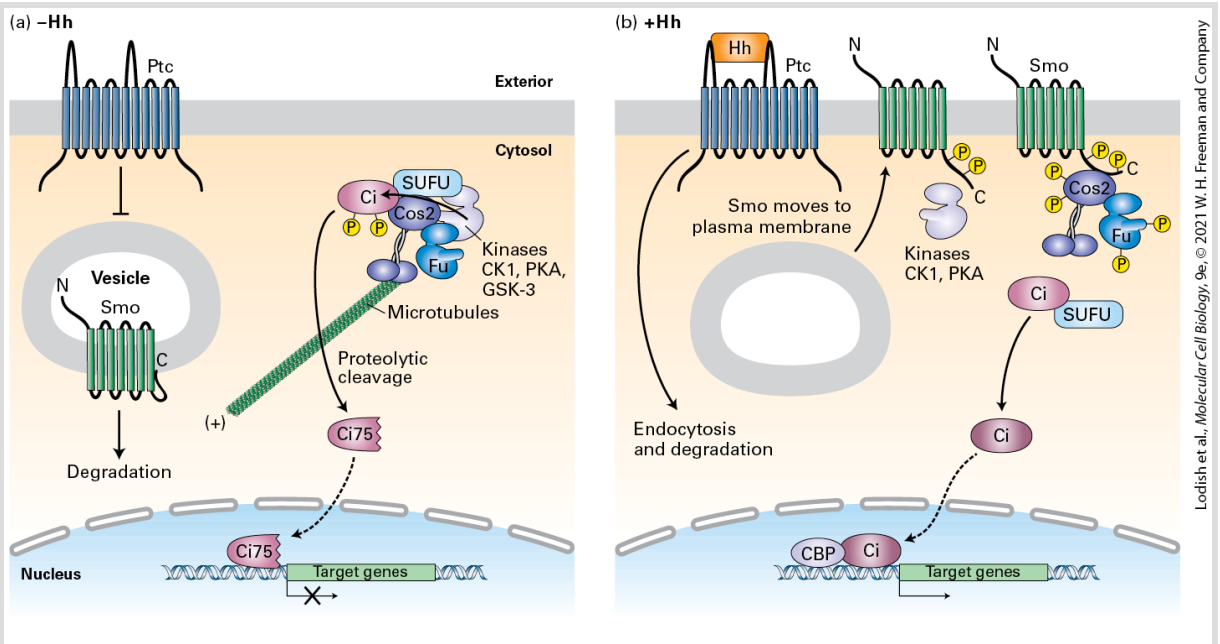


FIGURE 16-29 Hedgehog signaling in *Drosophila*. (a) In the absence of Hedgehog (Hh), the Hh receptor protein Patched (Ptc) inhibits Smoothened (Smo), which is present largely in the membranes of internal vesicles. Smo becomes ubiquitinated and then is degraded in proteasomes. A cytosolic complex containing the kinase Fused (Fu); other kinases including

protein kinase A (PKA), glycogen synthase kinase 3β (GSK3β), and casein kinase 1 (CK1), the zinc-finger transcription factor Cubitis interruptus (Ci), and SUFU (Suppressor of Fu), a protein that binds to and inhibits Ci, is bound to a microtubule by the kinesin-related motor protein Costal-2 (Cos2). In this complex, Ci becomes phosphorylated in a series of steps catalyzed by PKA, GSK3β, and CK1. The phosphorylated Ci is then proteolytically cleaved by the ubiquitin-proteasome pathway, generating an N-terminal fragment Ci75, which is transported into the nucleus and functions as a transcriptional repressor of Hh target genes. (b) Hh binds to Ptc, causing Ptc to be endocytosed from the cell surface and degraded, thereby relieving the inhibition of Smo. Smo then moves to the plasma membrane and its cytosolic C-terminus is phosphorylated by PKA, CK1, and other kinases. Then the C-terminal segment of Smo binds Cos2, and Smo is stabilized from degradation. Both Fu and Cos2 become extensively phosphorylated, and most importantly, the Fu-Cos2-Ci complex dissociates. This leads to the stabilization of the full-length Ci, which moves into the nucleus, displaces the repressor Ci75 from the promoter of target genes, recruits the CREB-binding activator protein (CBP), and induces expression of target genes. The exact membrane compartments in which Ptc and Smo respond to Hh and function are unknown. See S. Goetz and K. Anderson, 2010, *Nat. Rev. Genet.* **11**:331; and R. Heck et al., 2016, *Development* **143**:367.

Subsequent immunostaining and subcellular fractionation studies showed that, in the absence of Hh, Ptc is enriched in the plasma membrane but Smo is found in membranes of internal vesicles (Figure 16-29a). Precisely how Ptc inhibits Smo function is poorly understood.

In the absence of Hh, a microtubule-bound complex containing several different proteins including Fu and Cos2 sequesters in the cytosol the zinc-finger transcription factor called, in *Drosophila*, *Cubitis interruptus (Ci)*. The protein complex contains at least four kinases, including the same two that phosphorylate β-catenin in the

Wnt pathway. Phosphorylation of Ci in this complex by these kinases triggers binding of a component of an E3 ubiquitin ligase complex, which in turn directs ubiquitinylation of Ci and its targeting to proteasomes. There Ci undergoes a specific proteolytic cleavage but not complete degradation to amino acids; the resulting Ci fragment, designated *Ci75*, translocates to the nucleus and *represses* expression of Hh target genes.

Following binding of Hh to the receptor Ptc, the Hh-Ptc complex, like many other receptor-ligand complexes, is endocytosed from the cell surface into internal vesicles and eventually degraded; the binding of Hh to Ptc also inhibits the ability of Ptc to inhibit Smo (Figure 16-29b). The now active Smo moves from internal vesicles to the plasma membrane, where it initiates the signaling pathway that sends the full-length transcription factor Ci to the nucleus. The C-terminal cytosol-facing domain of Smo becomes phosphorylated through the coordinated activities of several kinases. Then Cos2 binds to the phosphorylated C-terminal tail of Smo and itself becomes phosphorylated. This leads to a disruption of the cytosolic complex of Fu, Cos2, and Ci, and its dissociation from microtubules, triggering a reduction in both phosphorylation and proteolytic cleavage of Ci. As a consequence, full-length Ci is released from the dissociated complex and translocates to the nucleus, where it binds to the transcriptional co-activator CREB-binding protein (CBP), activating rather than repressing Hh target genes.

Feedback Regulation of Hh Signaling

As in other signaling pathways, feedback control of the Hh pathway is important because unrestrained Hh signaling can cause cancerous overgrowth or formation of the wrong cell types. In *Drosophila*, one of the genes induced by the Hh signal is *patched*. The subsequent increase in expression of Ptc dampens the Hh signal in large measure by reducing the pool of active Smo protein inside the cell and by sequestering Hh protein on the cell surface. Thus the system is buffered: if during development too much Hh signal is made, a consequent increase in Ptc will compensate; if too little Hh signal is made, the amount of Ptc produced is decreased.

Hedgehog Signaling in Vertebrates Requires Primary Cilia

The Hedgehog signaling pathway in vertebrates shares many conserved features with the *Drosophila* pathway, but there are also some striking differences. First, mammalian genomes contain three *Hh* genes and two *Ptc* genes, which are expressed differentially among various tissues. Second, mammals express three Gli transcription factors, which collectively perform the roles of the single Ci transcription factor in *Drosophila*. All other components of the Hh pathway in *Drosophila* also are conserved in mammals.

The most fascinating aspect of the mammalian Hh pathway is the involvement of the primary cilium. Cilia are long, thin structures enveloped by a plasma membrane that protrude from the cell surface. The roles of cilia and flagella in specialized cell types — such as moving materials along the airway surface formed by tracheal cells and powering the locomotion of sperm — are well known (see Figure 1-15 and Chapter 18). Most vertebrate cells have a single cilium called the *primary cilium*, a slim, nonmotile structure that projects from the surface, but primary cilia are conspicuously absent in almost all invertebrate cell types that have been examined.

As we will learn in Chapter 18, a cilium is extended and maintained by the transport of proteins and particles along a bundle of microtubules in its center; different intraflagellar transport (IFT) motor proteins move proteins and particles along the microtubule bundle from the base of the cilium to the tip and in the opposite direction. Some of the first evidence for a role of primary cilia in Hh signaling came from a screen for mutations in mice that altered early development in a manner similar to that seen in embryos with altered Hh signaling: the mutant phenotypes included losses of certain types of cells in the neural tube that require high concentrations of a particular Hh protein to develop. Many of these mutations were in genes encoding IFT proteins, indicating a role for cilia (or flagella) in Hh signaling.

Further studies showed that in the absence of Hh signaling, Ptc is localized to the membrane of the primary cilium and Smo is located in internal vesicles near the base of the cilium. As in *Drosophila*, in the absence of Hh, kinases in a cytosolic complex phosphorylate Gli, inducing cleavage of Gli by the proteasome and translocation of the Gli fragment (termed Gli^R) into the nucleus, where it binds to regulatory regions of Gli-responsive genes and *represses* their expression.

As a result of Hh binding to Ptc, Smo moves into the membrane of the primary cilium. Gli then accumulates at the tip of the cilium, a process requiring a kinesin motor protein (Chapter 18) to move material along microtubules to the tip of the cilium. At the cilia tip, Gli becomes activated by proteins attached to Smo by a mechanism not yet known in detail, but involving dissociation of Gli from Sufu, an inhibitor of Gli that is a component of the cytosolic complex. Another motor protein, in this case the retrograde motor dynein (Chapter 18), is required to move the full-length and activated Gli, (termed Gli^A) to the base of the cilium. As in *Drosophila*, this active transcription factor then moves into the nucleus, where it activates, rather than represses, expression of multiple target genes. The detailed mechanisms of Hh signaling and Gli activation remain unclear.

Ptc is predicted to transport cholesterol or a structurally related sterol that could either activate or inhibit Smo. The vertebrate Smo

protein can bind cholesterol as well as several kinds of oxysterol molecules, both by a site in its N-terminal extracellular segment and one within its membrane-spanning domain. As cholesterol is sufficient to activate vertebrate Smo, in vertebrates cholesterol or a related sterol, perhaps transported into the cytosol by Ptc, is likely to be the activator of Smo. In contrast, in *Drosophila*, cholesterol is not sufficient to activate Smo, and the identity of the activator for *Drosophila* Smo is completely unknown. Thus much remains to be learned about the mechanism of Smo activation both in flies and man!

Inappropriate activation of Hh signaling is the cause of several types of human tumors, including medulloblastomas (cerebellum tumors) and rhabdomyosarcomas (muscle tumors). Primary cilia are essential for this abnormal Hh signaling, and drugs that inhibit the function of primary cilia are being tested on animal models of these cancers. For instance, expression of a mutant activated form of Smo in the postnatal mouse brain causes medulloblastomas, but these tumors do not form if, simultaneously, a gene encoding an essential ciliary protein is inactivated. ■

Degradation of an Inhibitor Protein Activates the NF-κB Transcription Factor

As we have seen, in the non-stimulated state of the Wnt and Hh signal transduction pathways, a key transcription factor is ubiquitinylated and subjected to limited proteolytic cleavage or complete degradation; activation of the signaling pathway blocks ubiquitinylation and releases the transcription factor in its active state. The NF-κB pathway works in the opposite manner: in the resting state, the NF-κB transcription factor is retained in the cytosol bound to an inhibitor; activation of the signaling pathway results in the inhibitor becoming ubiquitinylated and then degraded, triggering release of the active transcription factor. This mechanism allows cells to respond to a variety of stress signals by quickly activating gene transcription. The steps in the NF-κB pathway were revealed in studies with both mammalian cells and *Drosophila*.

The NF-κB signaling pathway is named for its key transcription factor NF-κB (an acronym for the somewhat unwieldy descriptor "*n*uclear-*f*actor *k*appa-light-chain enhancer of activated *B* cells"). This transcription factor is rapidly activated in many mammalian immune-system cells in response to bacterial and viral infection, inflammation, and a number of other stressful situations, such as ionizing radiation. In certain cells of the immune system, for example, the NF-κB pathway is activated when components of bacterial or fungal cell walls or segments of bacterial DNA bind to certain *Toll-like receptors* on the cell surface or in endosomal membranes (see Figure 24-37). In several types of cells, this pathway is activated by so-called inflammatory cytokines, such as *tumor*

necrosis factor alpha (TNFα) and *Interleukin 1 (IL-1)*, which are released by nearby immune cells in response to infection. In all of these cases, binding of ligand to its receptor induces a signaling pathway that activates the NF-κB transcription factor.

NF-κB was originally discovered on the basis of its transcriptional activation of the gene encoding the light chains of antibodies (immunoglobulins) in B cells. It is now thought to be a master transcriptional regulator of the immune system in mammals. Although flies do not make antibodies, NF-κB homologs in *Drosophila* induce synthesis of a large number of secreted antimicrobial peptides in response to bacterial and viral infection. The existence of these homologs indicates that the NF-κB regulatory system has been conserved during evolution and is more than half a billion years old.

Biochemical studies in mammalian cells and genetic studies in flies have provided important insights into the operation of the NF-κB pathway. The NF-κB transcription factor is a heterodimer formed of two subunits, p65 and p50. These two subunits share a region of homology at their N-termini that is required both for their dimerization and binding to DNA. In cells that are not undergoing a stress or responding to signs of an infection, NF-κB binds directly to an inhibitor called *I-κBα* and by this linkage is sequestered in an inactive state in the cytosol. A single molecule of I-κBα binds to the

paired N-terminal domains of the p50–p65 subunits, thereby masking their nuclear-localization signals (Figure 16-30).

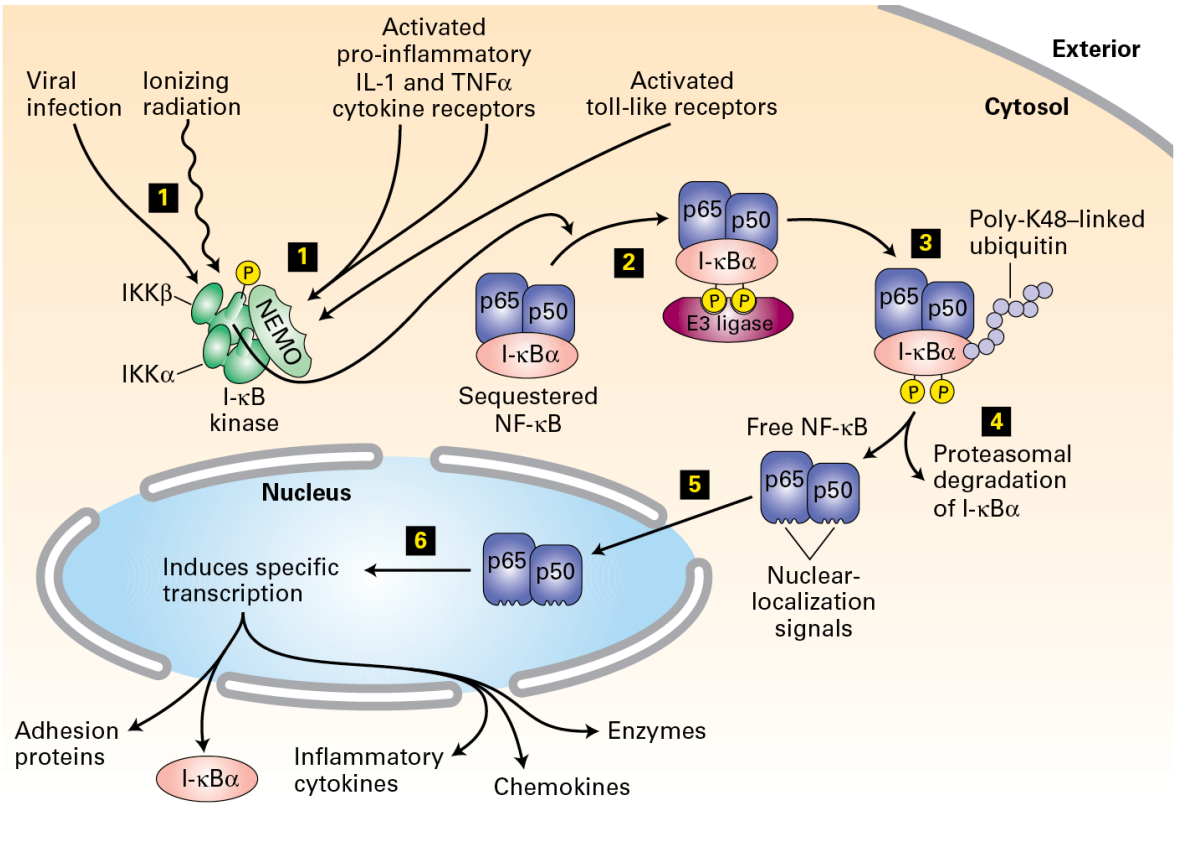


FIGURE 16-30 Activation of the NF-κB signaling pathway. In resting cells, the heterodimeric transcription factor NF-κB, composed of p50 and p65 subunits, is sequestered in the cytosol, bound to the inhibitor protein I-κBα. Step **1**: As indicated by the arrows, the three-subunit I-κB kinase is activated downstream of signal transduction pathways stimulated by diverse agents, including activation of any of several Toll-like receptors (Figure 24-37), and by components of invading viruses, bacteria, or fungi. Step **2**: The β subunit of I-κB kinase phosphorylates the inhibitor protein I-κBα, creating a binding site on I-κBα for an E3 ubiquitin ligase. Steps **3** and **4**: Subsequent lysine 48–linked polyubiquitinylation of I-κBα targets it for degradation by proteasomes. Step **5**: The removal of I-κBα unmasks the nuclear-localization signals in both subunits of NF-κB, allowing their translocation to the nucleus. Step **6**: In the nucleus, NF-κB activates transcription of numerous target genes, including the gene encoding I-κBα, which acts to terminate signaling, and, depending on the

cell, activates genes encoding various inflammatory cytokines and other proteins. See R. Khush et al., 2001, *Trends Immunol.* **22**:260; and J. L. Luo et al., 2005, *J. Clin. Invest.* **115**:2625.

A three-protein complex termed *I-κB kinase* operates immediately upstream of NF-κB and is responsible for releasing it from sequestration. Activation of the kinase activity of the β subunit of I-κB kinase, abbreviated *IKKβ*, is the point of convergence of all of the extracellular signals noted above that activate NF-κB. Within minutes of stimulation of the cell by an infectious agent or inflammatory cytokine, the β subunit of I-κB kinase becomes activated by phosphorylation and then it phosphorylates two N-terminal serine residues on I-κBα (Figure 16-30, steps **1** and **2**). An E3 ubiquitin ligase then binds to these phosphoserines and ubiquitinylates I-κBα. This E3 ubiquitin ligase links the carboxyl terminus of one ubiquitin to lysine 48 (K48) on another, forming poly-K48-linked ubiquitin chains that target the attached I-κBα protein to the proteasome and its immediate degradation (steps **3** and **4**; also see Chapter 3). In cells expressing mutant forms of I-κBα in which these two serines have been changed to alanine and so cannot be phosphorylated, NF-κB is permanently inactive, demonstrating that phosphorylation of I-κBα is essential for pathway activation.

The degradation of I-κBα exposes the nuclear-localization signals on the two NF-κB subunits, which then translocate into the nucleus and activate transcription of a multitude of target genes (Figure 16-30,

steps 5 and 6). NF-κB signaling is eventually turned off by negative feedback regulation because one of the genes whose transcription is immediately induced by NF-κB encodes I-κBα. As its levels increase, the I-κBα protein binds active NF-κB in the nucleus and returns it to the cytosol.

In many immune-system cells, NF-κB stimulates transcription of more than 150 genes, including those encoding cytokines and chemokines; the latter attract other immune-system cells and fibroblasts to sites of infection using the signal transduction pathways we described earlier. NF-κB also promotes expression of receptor proteins that enable neutrophils (a type of white blood cell) to migrate from the blood into the underlying tissue (see Figure 20-42). In certain immune cells, NF-κB stimulates expression of iNOS, the inducible isoform of the enzyme that produces nitric oxide, which is toxic to bacterial cells, as well as expression of several anti-apoptotic proteins, which prevent cell death. Thus this single transcription factor coordinates and activates the body's defense, either directly by responding to pathogens and stress or indirectly by responding to signaling molecules released from other infected or wounded tissues and cells.

Enormous Signalsomes with Polyubiquitin Chain Scaffolds Link Many Cell Surface Receptors to

Downstream Proteins in the NF-κB Pathway

As we have just seen, activation of the β subunit of I-κB kinase is the point of convergence for extracellular signals transmitted through multiple receptors, including Toll-like and IL-1 receptors. Since the cytosolic domains of the Toll-like and IL-1 receptors have no enzyme activity, it was a mystery for many years how activation of these receptors led to phosphorylation and activation of the I-κB kinase β subunit. Recent work showed that huge multiprotein complexes termed *signalsomes* are formed by activation of these receptors and that several kinases incorporated into these complexes phosphorylate and thus activate I-κB kinase.

Oligomerization occurs at multiple levels of the signal transduction cascade, as exemplified by the IL-1 signalsome (Figure 16-31). An MyD88 *adapter* protein binds to the cytosolic domain of the ligand-activated IL-1 receptor, triggering formation of long helical filaments formed by interactions between MyD88 proteins and the kinases IRAK 2 (*I*nterleukin 1 *r*eceptor-*a*ssociated *k*inase 2) and IRAK4. TRAF6 proteins form homotrimers each of which then bind to three molecules of IRAK 2, the proteins at the ends of these filaments, and the tips of adjacent TRAF6 trimers bind together forming large two-dimensional lattices containing thousands of copies of TRAF6 and the other signalsome proteins.

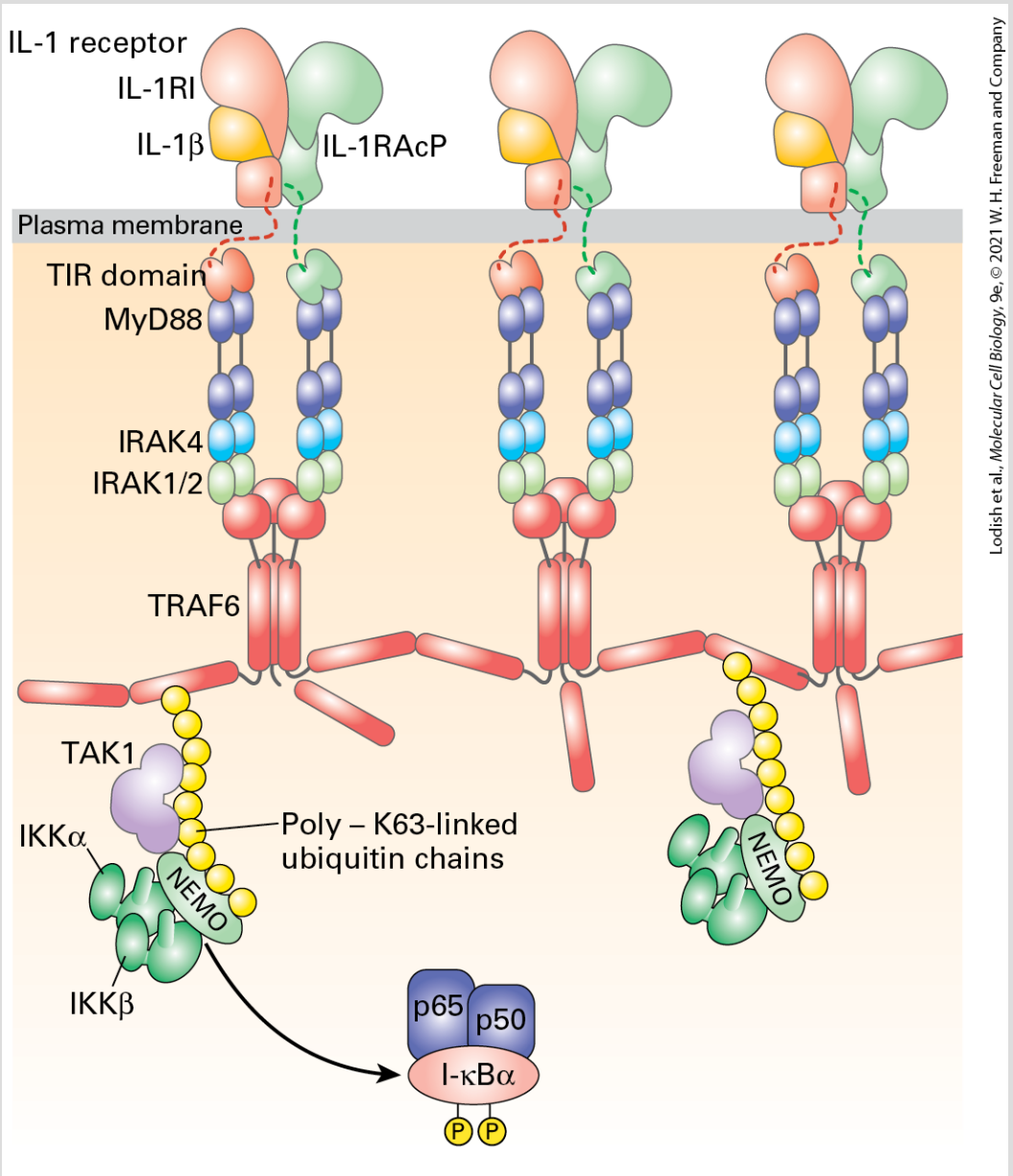


FIGURE 16-31 The giant interleukin 1 (Il-1) signalsome and activation of NF-κB. Like the cytosolic domains of the Toll-like receptors (Figure 24-37), the cytosolic domains of both subunits of the interleukin-1β (IL-1β) receptors contain a globular TIR domain; binding of IL-1β to interleukin 1 receptors triggers dimerization. MyD88 (red) proteins then bind to the receptors through interactions with the TIRs, and this in turn facilitates the successive binding of many copies of the kinases IRAK4 (green) and IRAK1 or IRAK2 (orange), leading to phosphorylation and activation of their kinase activities. Activated IRAK1 or IRAK2, in turn, binds TRAF6, an E3 ubiquitin ligase, to the signalsome, and

promotes the formation of an extended two-dimensional lattice formed by TRAF6-TRAF6 interactions. TRAF6 also catalyzes synthesis of long lysine-63-linked polyubiquitin chains that are covalently linked to TRAF6 itself. Several proteins including the protein kinase TAK1 and the protein NEMO bind to these K-63 linked polyubiquitin chains. NEMO (a component of the I-κB kinase complex depicted in Figure 16-30), in turn, recruits IKK kinases α and β to the complex, forming the three subunit I-κB kinase complex. IKK α and β kinases then become activated by phosphorylation by TAK1 or other active kinases in the supramolecular complex. As depicted in Figure 16-30, the β subunit of I-κB kinase then phosphorylates the inhibitor I-κB, leading to proteasome destruction of I-κBα and activation of the NF-κB signaling pathway. Only a small portion of an actual signalsome is depicted in this figure as the free ends of the TRAF6 proteins will bind to other TRAF6 molecules, expanding the area of the signalsome just inside the plasma membrane.

[Data from R. Ferrao et al., 2012, *Sci. Signal.* **5**:re3; B. Skaug et al., 2009, *Annu. Rev. Biochem.* **78**:769; and J. Napetschnig and H. Wu, 2013, *Annu. Rev. Biophys.* **42**:443.]

TRAF6 is an E3 ubiquitin ligase that synthesizes polyubiquitin chains. At the time TRAF6 was discovered all polyubiquitinylation was thought to signal degradation by proteasomes, so researchers looked for ubiquitinylated target proteins that were quickly destroyed. Not finding these, scientists looked for other possible roles for polyubiquitin and soon found that, depending on the specific E3 ubiquitin ligase, multiple types of polyubiquitin chains are formed that have different structures and biological functions. The E3 ligase TRAF6 links the carboxyl terminus of one ubiquitin to lysine 63 (K63) on another ubiquitin (see Figure 3-39), forming a linear poly-K63 ubiquitin chain. Rather than targeting the protein to which it is attached for degradation, these ubiquitin chains act as

scaffolds that bind proteins that have a *poly-K63 ubiquitin-binding domain*. One of these proteins is the protein kinase TAK1, which becomes activated by binding to the polyubiquitin chain; another is the NEMO subunit of the I-κB kinase. Binding to adjacent sites on the poly-K63 ubiquitin brings the TAK1 kinase and its target, the β subunit of I-κB kinase, into proximity so that TAK1 can phosphorylate and thereby activate this downstream kinase (see Figure 16-31). I-κB kinases can also become activated by phosphorylation by other already active I-κB kinases in the complex. As noted above, I-κB kinase then phosphorylates I-κBα, leading to activation of the NF-κB factor. Thus two different types of polyubiquitin chains participate in very different ways in transmitting the IL-1 signal to activation of the NF-κB transcription factors.

What is the advantage of such huge signalsome complexes? Why did they evolve as part of signal transduction pathways downstream of Toll-like receptors and receptors for many inflammatory hormones, including IL-1 and TNFα? One hypothesis is that many hormone receptors must cooperate to form signalsomes, leading to a threshold all-or-none response: a low concentration of an extracellular signal would not trigger assembly of a stable functional signalsome whereas just a slightly higher concentration would trigger a maximum response. This prevents the immune system from mounting a potentially damaging immune response in the absence of significant danger. Depending on the TNFα

concentration, formation of the signalsome by the TNFα receptor can take between 20 and 50 minutes, and this delay allows formation of a signalsome to be reversed if the initial stimulation does not persist. Such a fail-safe mechanism might have evolved to allow a response only when there is sufficient danger, as signaled by a persistent and high dose of stimulation.

KEY CONCEPTS OF SECTION 16.7

Signal Transduction Pathways ThatUtilize Proteasomal Degradation of Signaling Components: Wnt, Hedgehog, and the Many Hormones That Activate NF-κB

- Many signaling pathways involve ubiquitinylation and proteolysis of target proteins and so are irreversible or only slowly reversible. These target proteins can be either a transcription factor or an inhibitor of a transcription factor.
- Wnt signaling controls numerous critical developmental events, such as brain development, limb patterning, and organogenesis. Like Wnt, Hedgehog (Hh) also functions as a morphogen during development. Activating mutations in both pathways can cause cancer.
- Both Hh and Wnt are secreted proteins that contain lipid anchors that reduce their signaling ranges. The fatty acid covalently attached to Wnt is essential for binding to its Frizzled (Fz) receptor.
- Wnt signals act through two cell-surface proteins, the receptor Fz and co-receptor LRP, and an intracellular complex containing β-catenin (see Figure 16-26). Binding of Wnt promotes the stability and nuclear localization of β-catenin, which either directly or indirectly promotes activation of the TCF transcription factor.
- Gradients of Wnt protein concentration are essential for many steps in development, including regeneration of a head and tail in planaria (see Figure 16-27).
- The Hh signal also acts through two cell-surface proteins, Smoothened (Smo) and Patched (Ptc), and an intracellular complex containing the Cubitis interruptus (Ci) transcription factor (see Figure 16-29). An activating form of Ci is generated in the presence of Hh; a repressing Ci fragment is generated in the absence of Hh. Both Ptc and Smo change their subcellular location in response to Hh binding to Ptc.

- Hh signaling in vertebrates requires primary cilia and intraflagellar transport proteins. Ptc localizes to the ciliary membrane in the absence of Hh, and Smo moves from internal membranes to the cilium membrane when Hh is present.
- The NF-κB transcription factor regulates many genes that permit immune system cells to respond to infection and inflammation.
- In unstimulated cells, NF-κB is localized to the cytosol, bound to the inhibitor protein I-κBα. In response to many types of extracellular signals, I-κBα is phosphorylated and then ubiquitinylated, targeting it for degradation by proteasomes. The destruction of I-κBα releases active NF-κB, which translocates to the nucleus (see Figure 16-30).
- Signal transduction by the interleukin 1 and Toll-like receptors involves formation of huge multiprotein complexes termed *signalsomes*. Polyubiquitin chains formed by the TRAF6 protein in this complex form a scaffold that brings the TAK1 kinase near its substrate, the β subunit of the I-κB kinase, activating the I-κB kinase that phosphorylates the inhibitor I-κBα. The resultant destruction of I-κBα allows signals to be transmitted from the receptor to downstream components of the NF-κB pathway (see Figure 16-31).

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

activation loop
Akt
cytokine receptors
cytokines
enhancer
erythropoietin (Epo)
granulocyte colony–stimulating factor (G-CSF)
Hedgehog (Hh)
HER
MAP kinase
morphogen
NF-κB

phosphoinositides
prolactin
promoter
protein kinase B (PKB)
scaffold proteins
Smads
transforming growthfactor β (TGF-β)
Wnt

Review the Concepts

1. Even though GRB2 lacks intrinsic enzyme activity, it is an essential component of the epidermal growth factor (EGF) signaling pathway that activates MAP kinase. What is the function of GRB2? What roles do the SH2 and SH3 domains play in the function of GRB2? Many other signaling proteins possess SH2 domains. What determines the specificity of SH2 interactions with other molecules?
2. Once an activated signaling pathway has elicited the proper changes in target-gene expression, the pathway must be inactivated. Otherwise, pathological consequences may result, as exemplified by persistent growth factor–initiated signaling in many cancers. Many signaling pathways possess intrinsic negative feedback by which a downstream event in a pathway turns off an upstream event. Describe the negative feedback that down-regulates signals induced by (a) erythropoietin and (b) TGF-β.

3. A mutation in the Ras protein renders Ras constitutively active (Ras^D). What is constitutive activation? How does constitutively active Ras promote cancer? What type of mutation might render the following proteins constitutively active: (a) Smad3, (b) MAP kinase, and (c) NF-κB?
4. The enzyme Ste11 participates in several distinct MAP kinase signaling pathways in the budding yeast *Saccharomyces cerevisiae*. What is the substrate for Ste11 in the mating factor signaling pathway? When a yeast cell is stimulated by mating factor, what prevents the induction of osmolytes required for survival in high-osmotic-strength media, given that Ste11 also participates in the MAP kinase pathway initiated by high osmolarity?
5. Describe the events required for full activation of protein kinase B.
6. Describe the function of the PTEN phosphatase in the PI-3 kinase signaling pathway. Why does a loss-of-function mutation in PTEN promote cancer? Predict the effect of constitutively active PTEN on cell growth and survival.
7. Name three features common to the activation of cytokine receptors and receptor tyrosine kinases. Name one difference with respect to the enzyme activity of these receptors.
8. Erythropoietin (Epo) is a hormone that is produced naturally in the body in response to low O_2 levels in the blood. The intracellular events that occur in response to Epo binding to its cell-surface receptor are well characterized. What molecule translocates from the cytosol

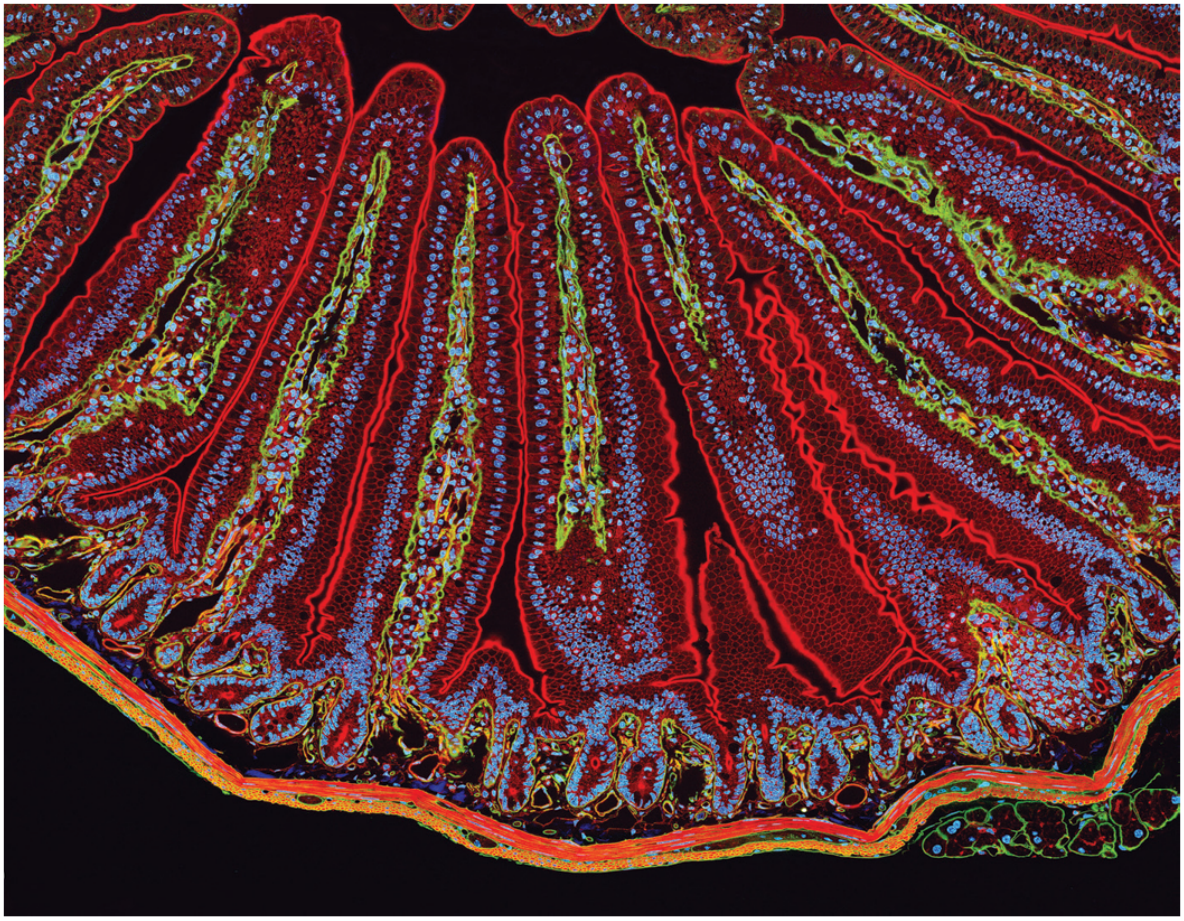
to the nucleus after (a) JAK2 activates STAT5 and (b) GRB2 binds to the Epo receptor? Why did some endurance athletes use Epo to improve their performance ("blood doping") until it was banned by most sports?

9. Explain how expression of a dominant-negative mutant of JAK blocks the erythropoietin (Epo)-cytokine signaling pathway.
10. Binding of TGF-β to its receptors can elicit a variety of responses in different cell types. For example, TGF-β induces plasminogen activator inhibitor 1 in epithelial cells and specific immunoglobulins in B cells. In both cell types, Smad3 is activated. Given the conservation of the signaling pathway, what accounts for the diversity of the response to TGF-β in various cell types?
11. How is the signal generated by binding of TGF-β to cell-surface receptors transmitted to the nucleus, where changes in target-gene expression occur? What activity in the nucleus ensures that the concentration of active Smads closely reflects the level of activated TGF-β receptors on the cell surface?
12. What feature of Delta ensures that only neighboring cells are signaled?
13. What biochemical reaction is catalyzed by γ-secretase?
14. The extracellular signaling protein Hedgehog can remain anchored to cell membranes. What modifications to Hedgehog enable it to be membrane bound? Why is this property useful?

15. Explain why loss-of-function *hedgehog* and *smoothened* mutations yield the same phenotype in flies, but a loss-of-function *patched* mutation yields the opposite phenotype.
16. Why is the signaling pathway that activates NF-κB considered to be relatively irreversible compared with cytokine or RTK signaling pathways? Nonetheless, NF-κB signaling must be down-regulated eventually. How is the NF-κB signaling pathway turned off?
17. Describe two roles for polyubiquitinylation in the NF-κB signaling pathway.

CHAPTER 17

Cell Organization and Movement I: Microfilaments



Courtesy Thomas Deerinck and Mark Ellisman.

A section of mouse intestine stained for actin (red), the extracellular matrix protein laminin (green), and DNA (blue). Each blue dot of DNA indicates the presence of a cell. Actin in the microvilli on the apical end of the epithelial cells can be seen lining the surface facing the lumen (*top*). Actin can also be seen prominently in the smooth muscle that surrounds the intestine (*bottom*).

OUTLINE

17.1 Microfilaments and Actin Structures
17.2 Dynamics of Actin Filaments
17.3 Mechanisms of Actin Filament Assembly
17.4 Organization of Actin-Based Cellular Structures
17.5 Myosins: Actin-Based Motor Proteins
17.6 Myosin-Powered Movements
17.7 Cell Migration: Mechanism, Signaling, and Chemotaxis

When we look through a microscope at the diversity of cells in nature, the variety of cell shapes and movements we see is astonishing. We may notice that some cells, such as vertebrate sperm, ciliates such as *Tetrahymena*, or flagellates such as *Chlamydomonas*, swim rapidly, propelled by cilia and flagella. Other cells, such as amoebae and human macrophages, move more sedately, propelled not by external appendages, but by coordinated movement of the cell itself. If we examine tissues under the microscope, we might notice that some cells connect to form a pavement-like sheet, whereas other cells — neurons, for example — have processes up to 3 feet in length that make selective contacts with other cells. When we look at the internal organization of cells, we see that organelles have characteristic locations. For example, the Golgi complex is generally near the nucleus, which itself is found in the center of the cell. How is this diversity of shape, cellular organization, and motility achieved? Why is it important for cells to have a distinct shape and well-defined internal organization?

To begin to answer these questions, let us consider two very different cells: epithelial cells and macrophages.

The epithelial cells that line the intestine form a tight, pavement-like layer of brick-shaped cells, known as an *epithelium* (Figure 17-1a, b). The epithelium imports nutrients from the intestinal lumen across the apical (*top*) domain of the plasma membrane and exports the nutrients across the basolateral (*bottom-side*) domain of the plasma membrane toward the bloodstream. To perform this directional transport, the apical and basolateral domains of the plasma membranes of epithelial cells contain different transport proteins.

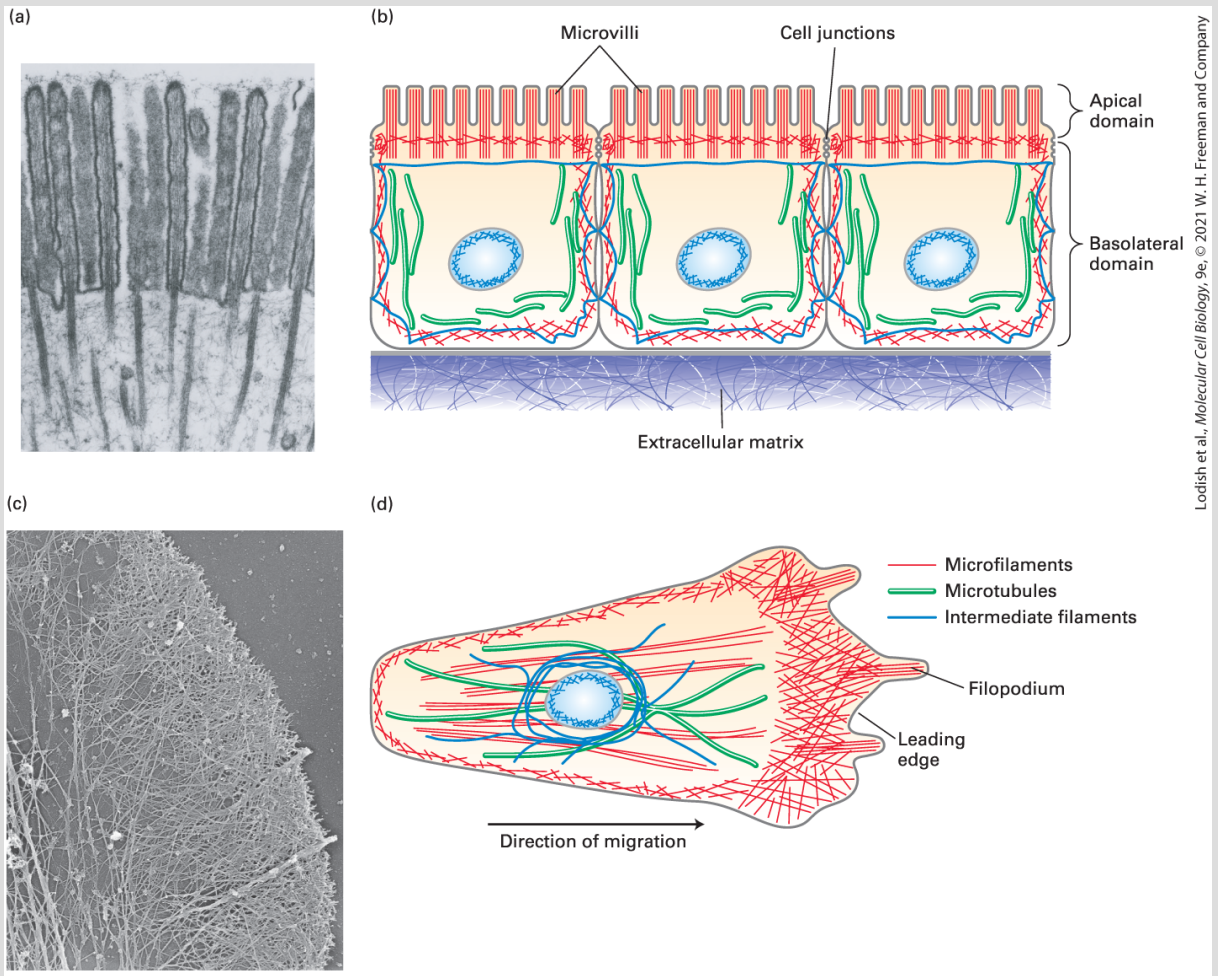


FIGURE 17-1 Overview of the cytoskeletons of an epithelial cell and a migrating cell. (a) Transmission electron micrograph of a thin section of an epithelial cell from the small intestine, showing the core bundles of microfilaments that provide support to the microvilli. (b) Epithelial cells are highly polarized, with distinct apical and basolateral domains. An intestinal epithelial cell transports nutrients into the cell through the apical domain and out of the cell across the basolateral domain. (c) Transmission electron micrograph of part of the leading edge of a migrating cell. The cell was treated with a mild detergent to dissolve the membranes, which also allows solubilization of most cytoplasmic components. The remaining cytoskeleton was shadowed with platinum and visualized in the electron microscope. Note the network of actin filaments visible in this micrograph. (d) A migrating cell, such as a fibroblast or a macrophage, has morphologically distinct domains, with a leading edge at the front. Microfilaments are indicated in red, microtubules in green, and intermediate filaments in dark blue. The position of the nucleus (light blue oval) is also shown.

[Part (a) M. S. Mooseker and L. G. Tilney, 1975, *J. Cell Biol.* **67**(3):725–743; https://doi.org/10.1083/jcb.67.3.725. Part (c) T. M. Svitkina and G. G. Borisy, 1999, *J. Cell Biol.* **145**(5):1009–1026; https://doi.org/10.1083/jcb.145.5.1009.]

Epithelial cells are sealed together by cell junctions (discussed in Chapter 20), which create a physical barrier between the apical and basolateral domains of the membrane. This separation allows the cell to place the proper transport proteins in the appropriate places: either the apical or the basolateral membrane. In addition, the apical membrane domain has a unique morphology, with numerous fingerlike projections called microvilli that dramatically increase the area of the plasma membrane available for nutrient absorption. So we see that the distinct functions of different regions of the intestinal epithelium are due to specific morphological features. To achieve these features (such as membrane domains, microvilli, and cell junctions), epithelial cells must have an internal structure that gives the cells shape, promotes cell-to-cell adhesion, and aids in delivering the right proteins to the right membrane domain.

Now consider the macrophage, a type of white blood cell that seeks out infectious agents and destroys them by an engulfing process called *phagocytosis*. For example, bacteria release chemicals that attract the macrophage and guide it to the site of infection. As the macrophage follows the chemical gradient, twisting and turning to get to the bacteria and phagocytose them, it must constantly reorganize its cell locomotion machinery. As we will see, the internal machinery of macrophages that allows them to crawl across

a substrate is always oriented in the direction in which they crawl (Figure 17-1c, d).

These are just two examples of **cell polarity**, the ability of a cell to generate functionally distinct regions within itself. In fact, as you think about various types of cells, you might realize that most of them have some form of cell polarity. A universal example of cell polarity is the ability of cells to divide — they must first select an axis for cell division and then set up the machinery to segregate their organelles along that axis.

The shape, internal organization, and functional polarity of a cell are due to a network of filamentous proteins called the **cytoskeleton**. We can visualize the cytoskeleton in different ways. For example, the cytoskeleton can be visualized by electron microscopy after treating cells with gentle detergents that solubilize the plasma membrane and internal organelles, releasing most of the cytoplasm (see Figure 17-1c). The cytoskeleton can also be visualized by immunofluorescence microscopy, as shown in Figure 17-2. We know from various studies that the cytoskeleton extends throughout the cell and is attached to the plasma membrane and to internal organelles. Thus the cytoskeleton provides a framework for organizing cellular components and processes. The term *cytoskeleton* may imply a fixed structure like a bony skeleton in vertebrates. In fact, the cytoskeleton can be very dynamic, with components reorganizing in less than a minute under some conditions. Under

different conditions, it can be stable for hours. As we will see throughout this chapter, the lengths of cytoskeletal filaments and the dynamics of filament assembly vary both in different regions of the cell and depending on the needs of the cell. The result is that cytoskeletal filaments are assembled into many types of structures and are regulated within different regions of the cell.

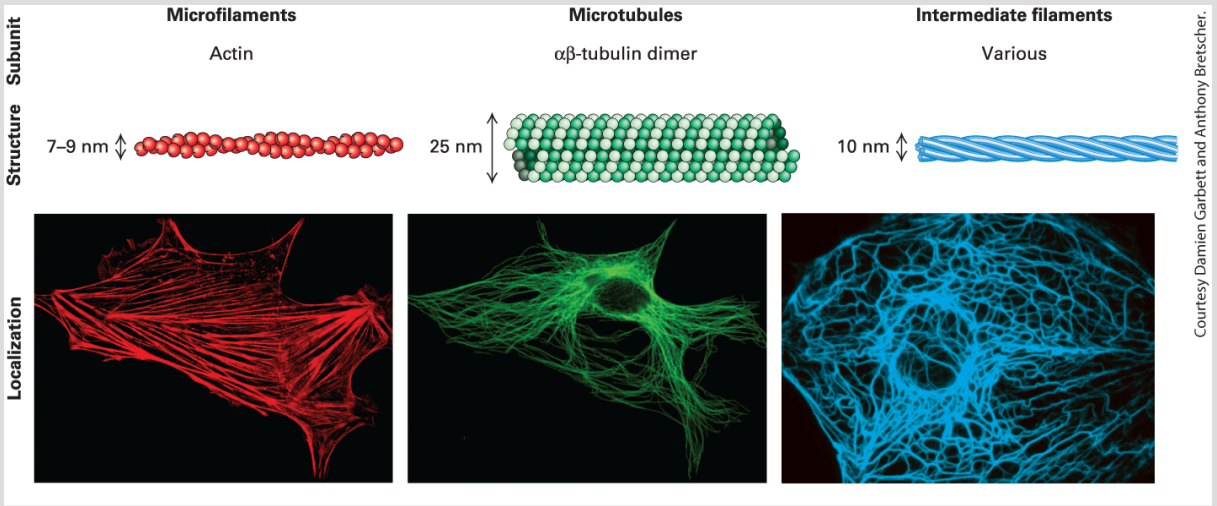


FIGURE 17-2 The components of the cytoskeleton. Each filament type is assembled from specific subunits in a reversible process so that cells can assemble and disassemble filaments as needed. Bottom panels show localization of the three filament systems in cultured cells as seen by immunofluorescence microscopy of actin, tubulin, and an intermediate filament protein, respectively. The image of microtubules shows their distribution in an interphase cell. In mitosis, they make up the machinery for chromosome segregation, as we discuss in Chapter 18.

The cytoskeleton is composed of three different filament systems, shown in Figures 17-1b, 17-1d, and 17-2. The filaments are polymers whose assembly and disassembly are regulated. This gives the cell flexibility to assemble and disassemble different types of cytoskeletal structures as needed. The three filament systems are:

- **Microfilaments** are polymers of the protein *actin* that are organized into functional bundles and networks by actin-binding proteins. Microfilaments are especially important in organizing the plasma membrane, giving shape to surface structures such as microvilli. Microfilaments can provide mechanical support on their own or can serve as tracks for ATP-powered myosin **motor proteins**. Myosin can ferry cargo along microfilaments, and myosin and microfilaments working together provide a contractile function (as in muscle).
- **Microtubules** are polymeric tubes formed from the protein *tubulin* and organized by microtubule-associated proteins. Microtubules make up the mitotic spindle, the machine for segregating duplicated chromosomes at mitosis. In interphase cells, as shown in Figure 17-2, microtubules often extend throughout the cell, providing a framework for the organization of organelles and structural support to cilia and flagella. Motor proteins called *kinesins* and *dyneins* transport cargo along microtubules, and like myosins, they are powered by ATP hydrolysis.
- **Intermediate filaments** are also polymers, but assembled from tissue-specific subunits that lend structural support to the nuclear membrane, provide structural integrity to cells in tissues, and serve structural and barrier functions in skin, hair, and nails. Unlike microfilaments and microtubules, intermediate filaments are not used as tracks by motor proteins.

As we can see in Figure 17-1, cells can organize their cytoskeletons in distinct ways. How cells organize the cytoskeleton depends on the

cells' ability to sense signals — from soluble factors bathing the cell, from adjacent cells, or from the extracellular matrix — and interpret them (Figure 17-3). These signals are detected by cell-surface receptors that activate signal transduction pathways, which ultimately converge on factors that regulate cytoskeletal organization.

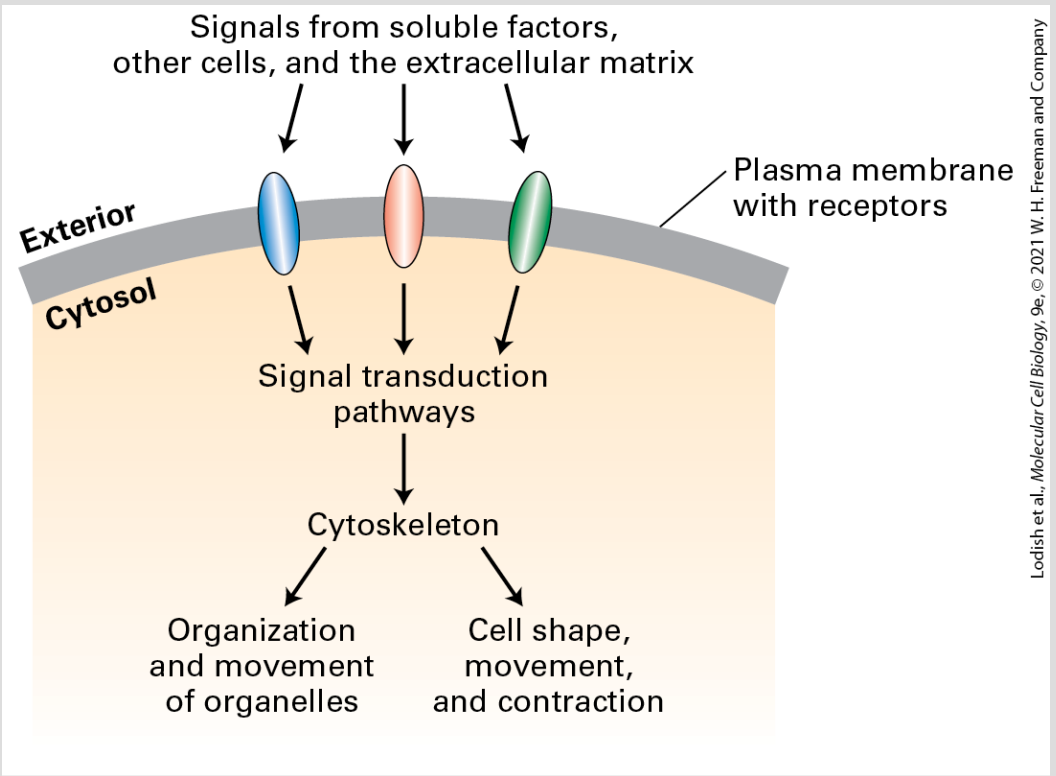


FIGURE 17-3 Regulation of cytoskeleton function by cell signaling. Cells use cell-surface receptors to sense external signals from the extracellular matrix, other cells, or soluble factors. These signals are transmitted across the plasma membrane and activate specific cytosolic signaling pathways. Signals — often integrated from more than one receptor — lead to the organization of the cytoskeleton so as to provide cells with their shape as well as to determine organelle distribution and movement. In the absence of external signals, cells still organize their internal structure, but not in a polarized manner.

The importance of the cytoskeleton for normal cell function and motility is evident when a defect in a cytoskeletal component — or in cytoskeletal regulation — causes disease. For example, about 1 in 500 people has a cytoskeletal defect that affects the contractile apparatus of the heart resulting in various cardiomyopathies (diseases of the heart muscle). Many defects in red blood cytoskeletal components that support these cells' plasma membrane are known to cause disease. Metastatic cancer cells break away from their tissue of origin and migrate to new locations because of misregulation of the cytoskeleton.

In this and the following chapter, we discuss the structure, function, and regulation of the cytoskeleton. We look at how a cell arranges its cytoskeleton to generate cell shape and cell polarity to provide organization and motility to its organelles and to be the structural framework for such processes as cell swimming and cell crawling. We discuss how cells assemble the three different filament systems and how signal transduction pathways regulate these structures. Regulation of the cytoskeleton during the cell cycle is discussed in Chapter 19. In Chapter 20, we examine how the cytoskeleton participates in the functional organization of tissues. Our focus in this current chapter is on microfilaments and actin-based structures. Although we initially examine the three cytoskeletal systems separately, in the next chapter we will see that microfilaments cooperate with microtubules and intermediate filaments in the normal functioning of cells.

17.1 Microfilaments and Actin Structures

Microfilaments can assemble into a wide variety of structures within a cell (Figure 17-4a). Each of these structures is critical for particular cellular functions. For example, microfilaments can form a tight bundle of filaments making up the core of the fingerlike **microvilli**, but they can also be found in a less ordered network just beneath the plasma membrane, in a region of cytoplasm known as the *cell cortex*, where they provide support and organization to membrane proteins. In epithelial cells, microfilaments form a contractile band around the cell, the *adherens belt*, that provides strength to the epithelium. In migrating cells, a network of microfilaments is found at the front of the cell in the **leading edge**, or **lamellipodium**, from which bundles of filaments called **filopodia** may protrude. Many cells have contractile microfilaments called **stress fibers**, which attach to the external substratum as cells migrate (discussed in Chapter 20). Specialized cells such as macrophages use contractile microfilaments to engulf and internalize pathogens though phagocytosis. Highly dynamic, short bursts of actin filament assembly can power the movement of *endocytic vesicles* away from the plasma membrane. At a late stage of cell division in animals, after all the organelles have been duplicated and segregated, a *contractile ring* forms and constricts to generate two daughter cells in the process of *cytokinesis*. Thus cells use actin filaments in many

ways: in structural roles, as contractile mechanisms, by harnessing actin polymerization and depolymerization to do work, and to move vesicles around within a cell. Multiple arrangements of microfilaments often coexist within a single cell, as shown in a migrating fibroblast (Figure 17-4b).

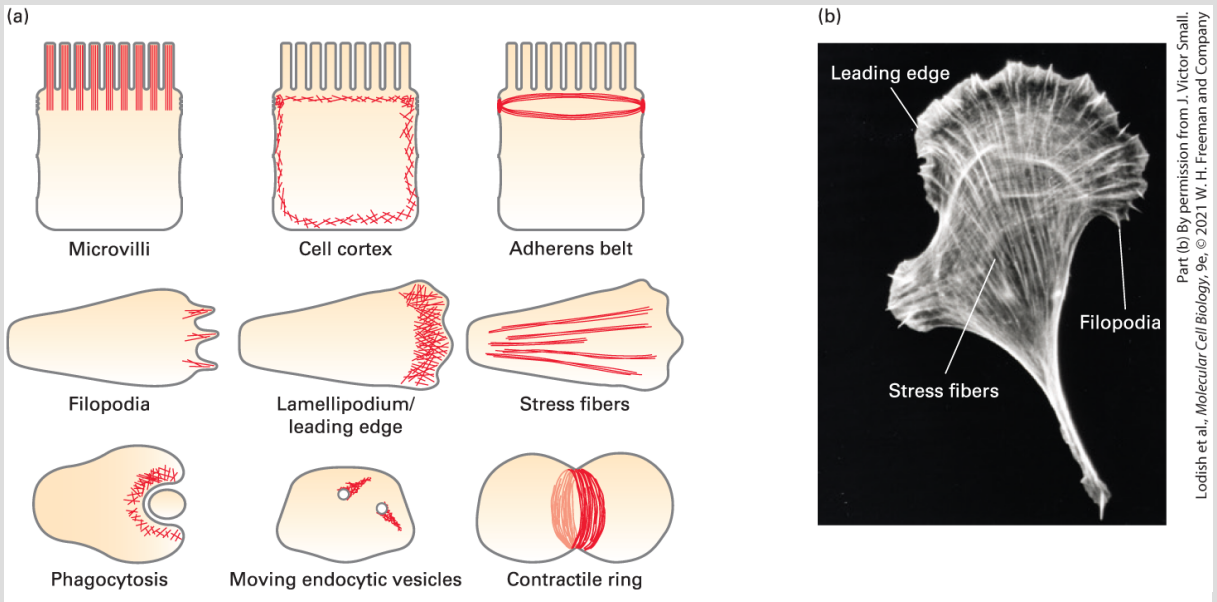


FIGURE 17-4 Examples of microfilament-based structures. (a) In each panel, microfilaments are depicted in red. (b) A cell moving toward the top of the page, stained for actin with fluorescent phalloidin, a drug that specifically binds F-actin. Note how many different organizations of microfilaments can exist in one cell.

The basic building block of microfilaments is **actin**, a protein that has the remarkable capacity to reversibly assemble into a polarized filament with functionally distinct ends. These filaments are molded into the various structures described in the previous paragraph by actin-binding proteins. The name *microfilament*, which arose from the very thin filaments seen by electron microscopists in thin-

section preparations of cells, refers to actin in its polymerized form, with its associated proteins. In this section, we look at the actin protein itself and the filaments into which it assembles.

Actin Is Ancient, Abundant, and Highly Conserved

Actin is an abundant intracellular protein in eukaryotic cells. In muscle cells, for example, actin constitutes 10 percent by weight of the total cellular protein; even in nonmuscle cells, actin makes up 1–5 percent of the cellular protein. The cytosolic concentration of actin in nonmuscle cells ranges from 0.1 to 0.4 mM. In structures such as microvilli, however, the local actin concentration can be as high as 5 mM. To grasp how much actin is present in cells, consider a typical liver cell, which has 2×10^4 insulin receptor molecules but approximately 5×10^8, or half a billion, actin molecules. Because they form structures that extend across large parts of the cell interior, cytoskeletal proteins are among the most abundant proteins in a cell.

Actin is encoded by a gene family that gives rise to some of the most conserved proteins within and across species. The protein sequences of actins from amoebae and from animals are identical at 80 percent of their amino acid positions, despite about a billion years of evolution. The multiple actin genes found in modern

eukaryotes are related to a bacterial gene, *MreB*, whose product forms filaments that are important in bacterial cell-wall synthesis. Some single-celled eukaryotes, such as yeasts and amoebae, have one or two actin genes, whereas multicellular organisms often contain multiple actin genes. For instance, humans have 6 actin genes, while *Arabidopsis* has 8 to 10, and maize has 21 actin genes. Each functional actin gene encodes a different isoform of the protein. In vertebrates, four actin isoforms are present in specific types of muscle cells, and two other isoforms are found in nonmuscle cells. These six isoforms differ at only about 25 of the 375 residues in the complete protein or show about 93 percent identity. Actin isoforms have historically been classified into three groups based on their overall charge: the α-actins, β-actins, and γ-actins. Alpha-actins are associated with various contractile structures, β-actins are enriched in the cell cortex and leading edge of motile cells, and γ-actin is found in stress fibers.

G-Actin Monomers Assemble into Long, Helical F-Actin Polymers

Actin exists as a globular monomer called *G-actin* and as a filamentous polymer called *F-actin*, which is a linear chain of actin subunits. Each actin molecule contains a Mg^{2+} ion complexed with either ATP or ADP. In fact, actin is an ATPase, which hydrolyzes ATP

to ADP and P_i. The importance of the interconversion between the ATP and the ADP forms of actin is discussed below.

X-ray crystallographic analysis reveals that the G-actin monomer is separated into two lobes by a deep cleft (Figure 17-5a). At the base of the cleft is the *ATPase fold*, which is the site where ATP and Mg^{2+} are bound. The ATPase fold has structural similarity to the GTP-binding cleft of the GTPase molecular switches (see Figure 15-7). The floor of the cleft in G-actin acts as a hinge that allows the lobes to flex relative to each other. When ATP or ADP is bound to G-actin, the nucleotide affects the conformation of the molecule (in fact, without a bound nucleotide, G-actin denatures very quickly).

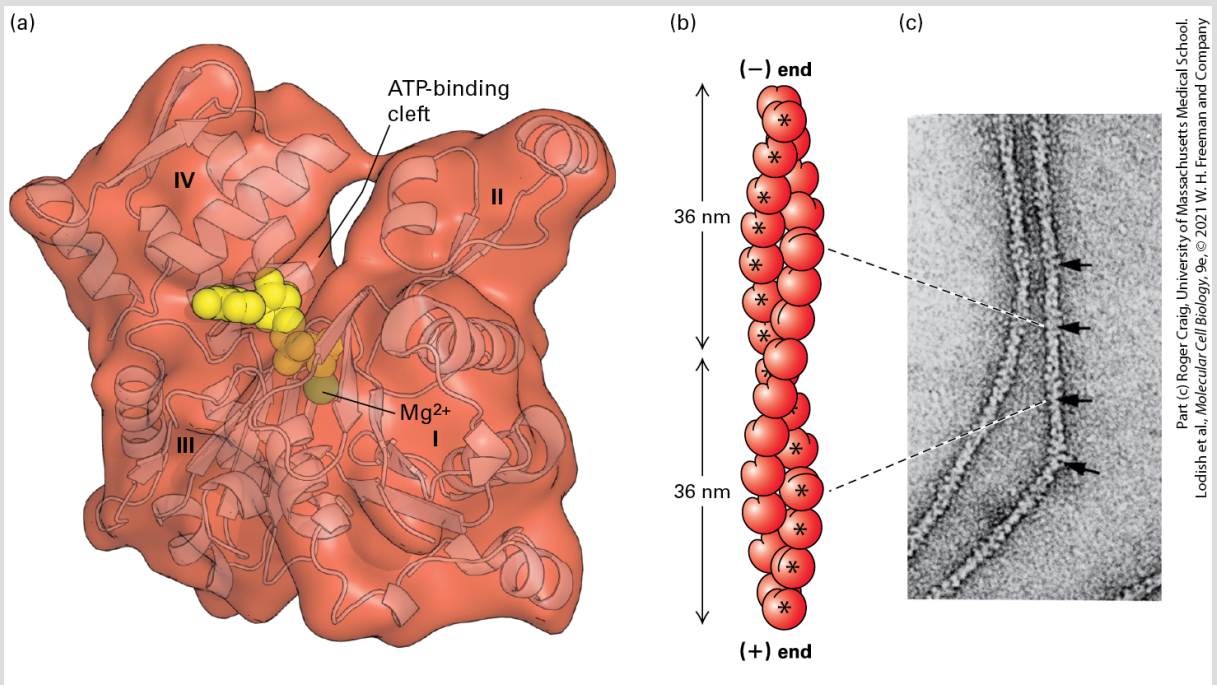


FIGURE 17-5 Structures of monomeric G-actin and F-actin filaments. (a) Structure of the actin monomer (measuring $5.5 \times 5.5 \times 3.5\,\mathrm{nm}$), which is divided by a central cleft into two approximately equal-sized lobes and four subdomains, numbered I–IV. ATP (yellow)

binds at the bottom of the cleft and contacts both lobes (the green ball represents Mg^{2+}). The N- and C-termini lie in subdomain I. (b) An actin filament appears as two strands of subunits. One repeating unit consists of 28 subunits (14 in each strand, indicated by * for one strand), covering a distance of 72 nm. The ATP-binding cleft of every actin subunit is oriented toward the same end of the filament. The end of a filament with an exposed binding cleft is the $(-)$ end; the opposite end is the $(+)$ end. (c) In the electron microscope, negatively stained actin filaments appear as long, flexible, and twisted strands of beaded subunits. Because of the twist, the filament appears alternately thinner (7-nm diameter) and thicker (9-nm diameter) (arrows). (The microfilaments visualized in a cell by electron microscopy are F-actin filaments plus any bound proteins.)

[Part (a) Data from C. E. Schutt et al., 1993, *Nature* **365**:810, PDB ID 2btf, courtesy of M. Rozycki.]

G-actin can polymerize, in a reversible reaction, into F-actin. The F-actin filaments that form in vitro are indistinguishable from microfilaments seen in cells, indicating that F-actin is the major component of microfilaments.

From the results of x-ray diffraction studies of actin filaments and from the actin monomer structure shown in Figure 17-5a, scientists have determined that the subunits in an actin filament are arranged in a helical structure (Figure 17-5b). In this arrangement, the filament can be considered as two helical strands wound around each other. Each subunit in the structure contacts one subunit above it and one below it in its own strand as well as two subunits in the other strand. The subunits in a single strand wind around the back of the other strand and repeat after 72 nm, or 14 actin subunits. Since there are two strands, the actin filament appears to repeat every 36 nm (see Figure 17-5b). When F-actin is viewed by electron

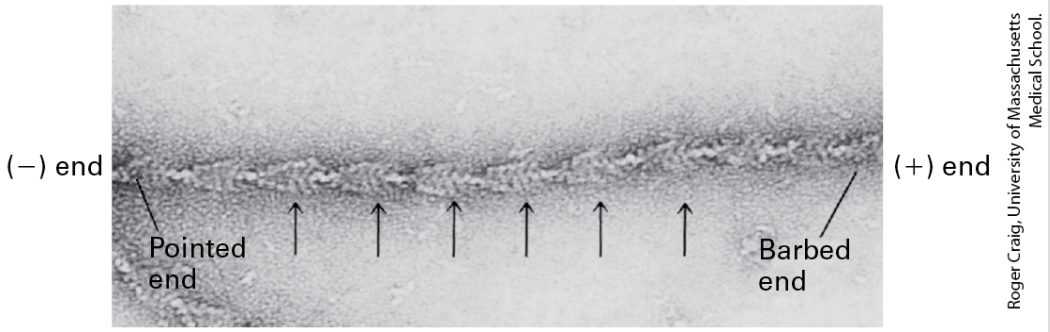
microscopy after negative staining with uranyl acetate, it appears as a twisted string whose diameter varies between 7 and 9 nm (Figure 17-5c).

F-Actin Has Structural and Functional Polarity

All subunits in an actin filament are oriented the same way. A consequence of this subunit orientation is that the filament as a whole exhibits polarity; that is, one end of the filament differs from the other. As we will see, a result of this subunit orientation is that one end of the filament is favored for the addition of actin subunits and is designated the (+) end, whereas the other end is favored for removal of subunits and is designated the (−) end. At the (+) end, the ATP-binding cleft of the terminal actin subunit contacts the neighboring subunit, whereas on the (−) end, the cleft is exposed to the surrounding solution (see Figure 17-5b).

The cleft in an actin subunit, and therefore the polarity of a filament, is not detectable without the atomic resolution afforded by x-ray crystallography. However, the polarity of actin filaments can be demonstrated by electron microscopy in so-called decoration experiments, which exploit the ability of the motor protein myosin (discussed in Section 17.5) to bind specifically to actin filaments. In this type of experiment, an excess of myosin S1, a proteolytic

cleavage product of myosin containing the actin-binding head domain (see Figure 17-22), is mixed with actin filaments under conditions where binding takes place. Myosin attaches to the sides of a filament and binds with a slight tilt. When all the actin subunits are bound by myosin, the filament appears as if decorated with arrowheads that all point toward the (−) end of a filament (Figure 17-6). Thus the (−) end is often called the *pointed end* of an actin filament. The (+) end is known as the *barbed end*. Because myosin binds to actin filaments and not to microtubules or intermediate filaments, arrowhead decoration is one criterion by which actin filaments can be definitively identified among the other cytoskeletal fibers in electron micrographs of cells.



EXPERIMENTAL FIGURE 17-6 Myosin S1 decoration demonstrates the identity and polarity of an actin filament. Myosin S1 head domains bind to actin subunits in a particular orientation. When bound to all the subunits in a filament, S1 appears to spiral around the filament. This coating of myosin heads produces a series of arrowhead-like decorations (arrows). The polarity in decoration defines a pointed (−) end and a barbed (+) end.

KEY CONCEPTS OF SECTION 17.1

Microfilaments and Actin Structures

- Microfilaments can be assembled into diverse structures, many of which are associated with the plasma membrane (see Figure 17-4a).
- Actin, the basic building block of microfilaments, is an abundant protein of eukaryotic cells and is highly conserved among species.
- Actin can reversibly assemble into filaments that consist of two helices of actin subunits.
- The actin subunits in a filament are all oriented in the same direction, with the nucleotide-binding site exposed on the $(-)$ end (see Figure 17-5).

17.2 Dynamics of Actin Filaments

The actin cytoskeleton is a dynamic structure, in which the assembly and disassembly of actin filaments, and their association into functional structures, are exquisitely regulated. In some cytoskeletal structures, microfilaments are stable for hours, whereas in others they grow or shrink in seconds. These changes in the organization of actin filaments can generate forces that cause large changes in the shape of a cell or drive movement of intracellular structures. In this section, we first consider actin's intrinsic ability to polymerize and depolymerize and then look at how these processes are modulated by accessory proteins. We will see how actin-binding proteins make important contributions to the stability and regulated disassembly of filaments. In subsequent sections, we turn to the mechanisms that cells use to assemble and organize actin filaments, how the location of assembly within a cell is regulated by signaling pathways, and how together with myosins they can drive motility processes. A summary of the actin-binding proteins discussed in this chapter is presented in Table 17-1.

Table 17-1 • Summary of Actin-Binding Proteins Discussed in This Chapter

Name	Primary Function
Profilin	G-actin nucleotide exchange
Cofilin	F-actin disassembly

Thymosin-β_4	G-actin monomer sequestering
CapZ $(+\text{end})$ and Tropomodulin $(-\text{end})$	F-actin end-binding
Gelsolin	F-actin severing
Formins, Arp2/3	Actin nucleating
Fimbrin, α-actinin, Filamin, Spectrin	F-actin cross-linking
Ankyrin, Band 4.1, ERM proteins, Dystrophin	F-actin-membrane linking
Myosins	F-actin-based molecular motors
Nebulin	F-actin length regulation
Tropomyosin	F-actin stabilizing

Actin Polymerization In Vitro Proceeds in Three Steps

The in vitro polymerization of G-actin monomers to form F-actin filaments can be monitored by viscometry, sedimentation, fluorescence spectroscopy, or fluorescence microscopy (see Chapter 4). When actin filaments grow long enough to become entangled, the viscosity of the solution, which is measured as a decrease in its flow rate in a viscometer, increases. The basis of the sedimentation assay is the ability of ultracentrifugation (100,000*g* for 30 minutes) to sediment F-actin, but not G-actin. The third assay makes use of G-

actin covalently labeled with a fluorescent dye; the fluorescence spectrum of the labeled G-actin monomer changes when it polymerizes into F-actin. Finally, growth of the fluorescently labeled filaments can be imaged in real time with fluorescence microscopy. These four assays are useful for kinetic studies of actin polymerization and for characterizing actin-binding proteins to determine how they affect actin dynamics or how they cross-link actin filaments.

The mechanism of actin assembly has been studied extensively. Remarkably, one can purify G-actin at a high protein concentration without it forming filaments — provided it is maintained in a buffer with ATP and low levels of cations. The addition of cations (e.g., to 100 mM K^+ and 2 mM Mg^{2+}) to a solution of G-actin will induce polymerization into F-actin filaments, with the kinetics of the reaction depending on the starting concentration of G-actin. The process is reversible: F-actin depolymerizes into G-actin when the ionic strength of the solution is lowered. The polymerization of pure G-actin in vitro proceeds in three sequential phases (Figure 17-7a).

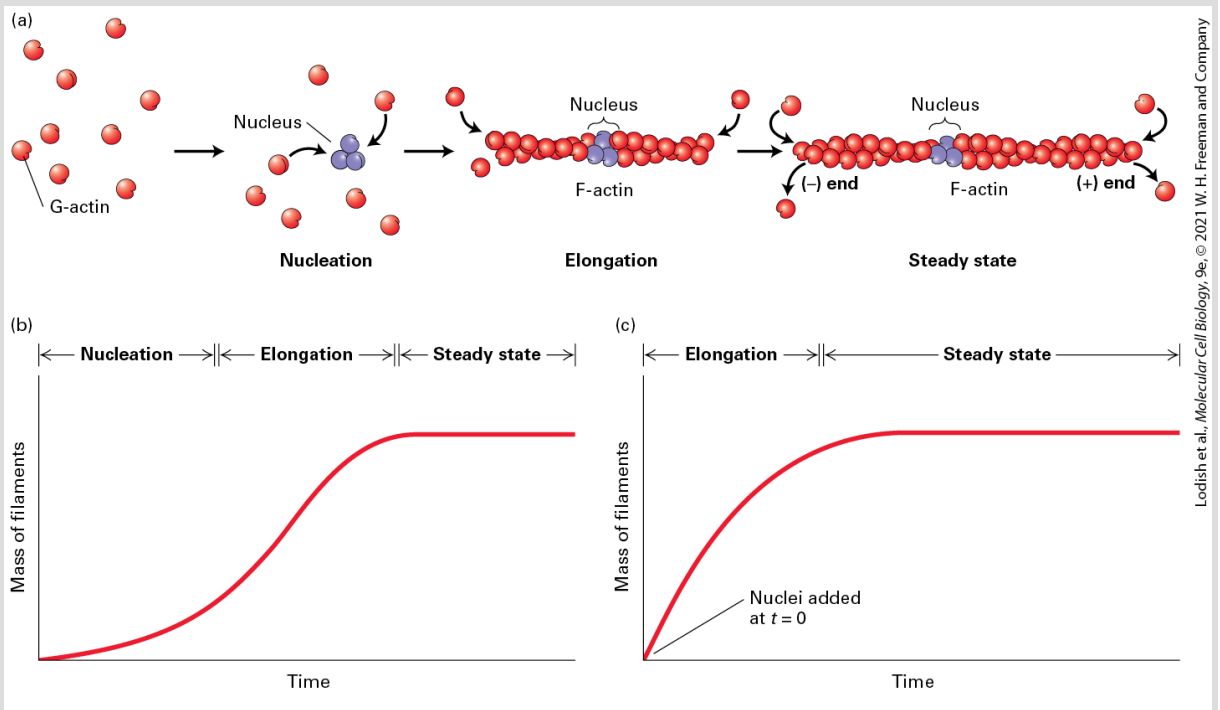


FIGURE 17-7 The three phases of in vitro G-actin polymerization. (a) In the initial nucleation phase, ATP–G-actin monomers (red) slowly form stable complexes of actin (purple). These nuclei are rapidly elongated in the second phase by the addition of subunits to both ends of the filament. In the third phase, the ends of actin filaments are in equilibrium with monomeric G-actin. (b) Time course of the in vitro polymerization reaction reveals the initial lag period associated with nucleation, the elongation phase, and the steady state. (c) If some short, stable actin filament fragments are added at the start of the reaction to act as nuclei, elongation proceeds immediately, without any lag period.

1. The *nucleation phase* is marked by a lag period in which G-actin subunits combine into an oligomer of two or three subunits. When the oligomer reaches three subunits in length, it can act as a seed, or nucleus, for the next phase.
2. During the *elongation phase*, the short oligomer rapidly increases in length by the addition of actin monomers to both of its ends. As F-actin filaments grow, the concentration of G-actin monomers decreases until equilibrium is reached

between the filament ends and monomers, and a steady state is reached.

3. In the *steady-state phase*, G-actin monomers exchange with subunits at the filament ends, but there is no net change in the total length of filaments.

The kinetic curves in Figure 17-7b, c show the filament mass during each phase of polymerization. We can show experimentally that the lag period is due to nucleation, because the lag period can be eliminated by the addition of a small number of F-actin nuclei — consisting of very short filaments — to a solution of G-actin (Figure 17-7c).

How much G-actin is required for nucleation to occur? Scientists have placed various concentrations of G-actin under polymerizing conditions and found that, below a certain concentration, filaments cannot assemble (Figure 17-8). Above this concentration, filaments begin to form; when steady state is reached, the incorporation of more free subunits is balanced by the dissociation of subunits from filament ends to yield a mixture of filaments and monomers. The minimal concentration of monomers at which filaments form is known as the **critical concentration (C_c)**. Below C_c, filaments will not form; above C_c, filaments form. At steady state, the concentration of monomeric actin remains at the C_c (Figure 17-8).

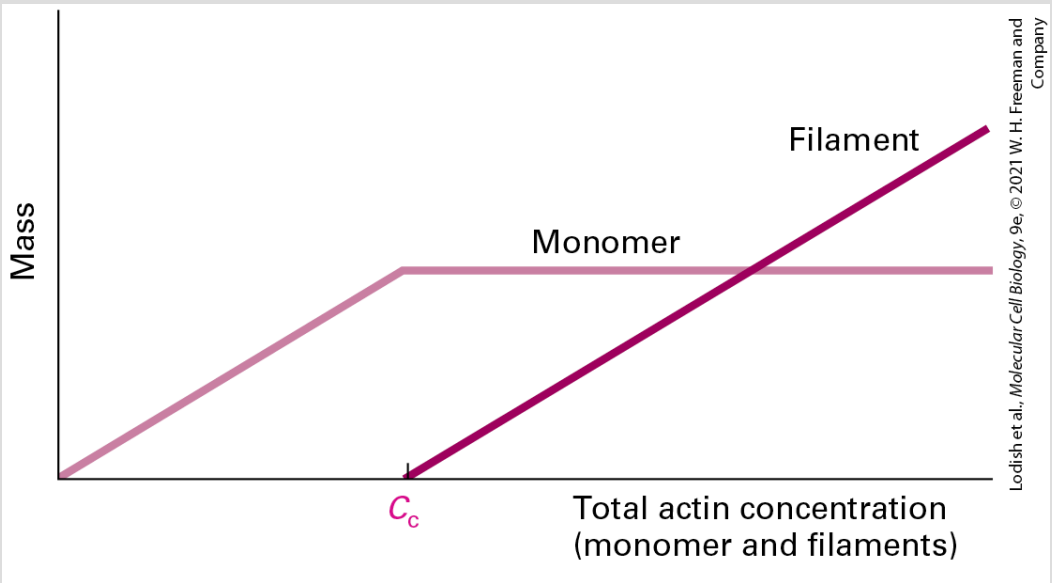
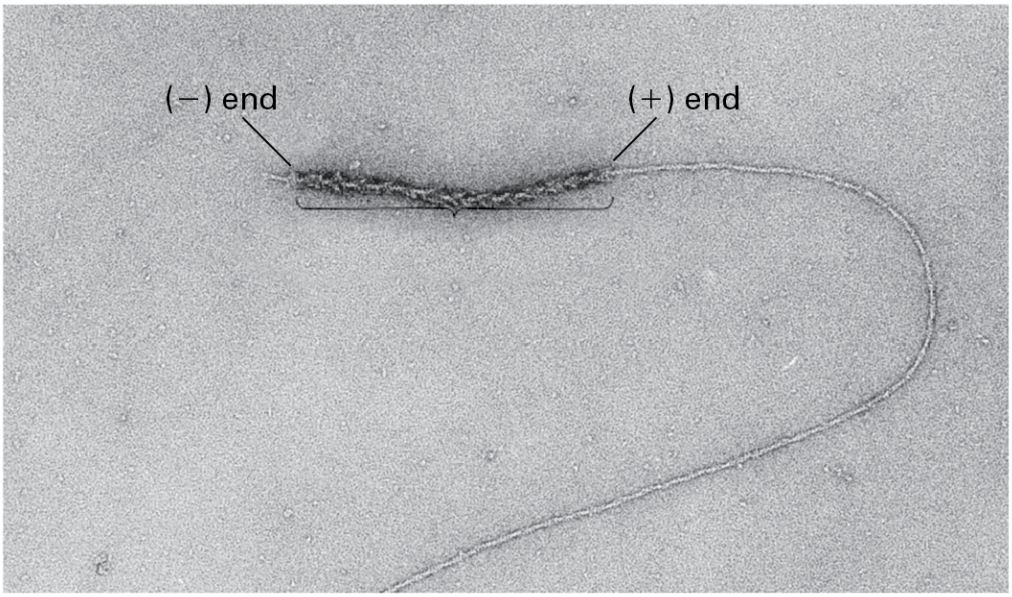


FIGURE 17-8 Determination of the critical concentration (C_c) of G-actin above which filaments are formed. When different concentrations of actin are placed under polymerizing conditions and the reaction is allowed to come to steady state, the C_c is the concentration at which actin filaments begin to form. At initial monomer concentrations below the C_c, no polymerization takes place. At monomer concentrations above the C_c, filaments assemble until steady state is reached with the monomer concentration remaining at C_c.

Actin Filaments Grow Faster at (+) Ends than at (−) Ends

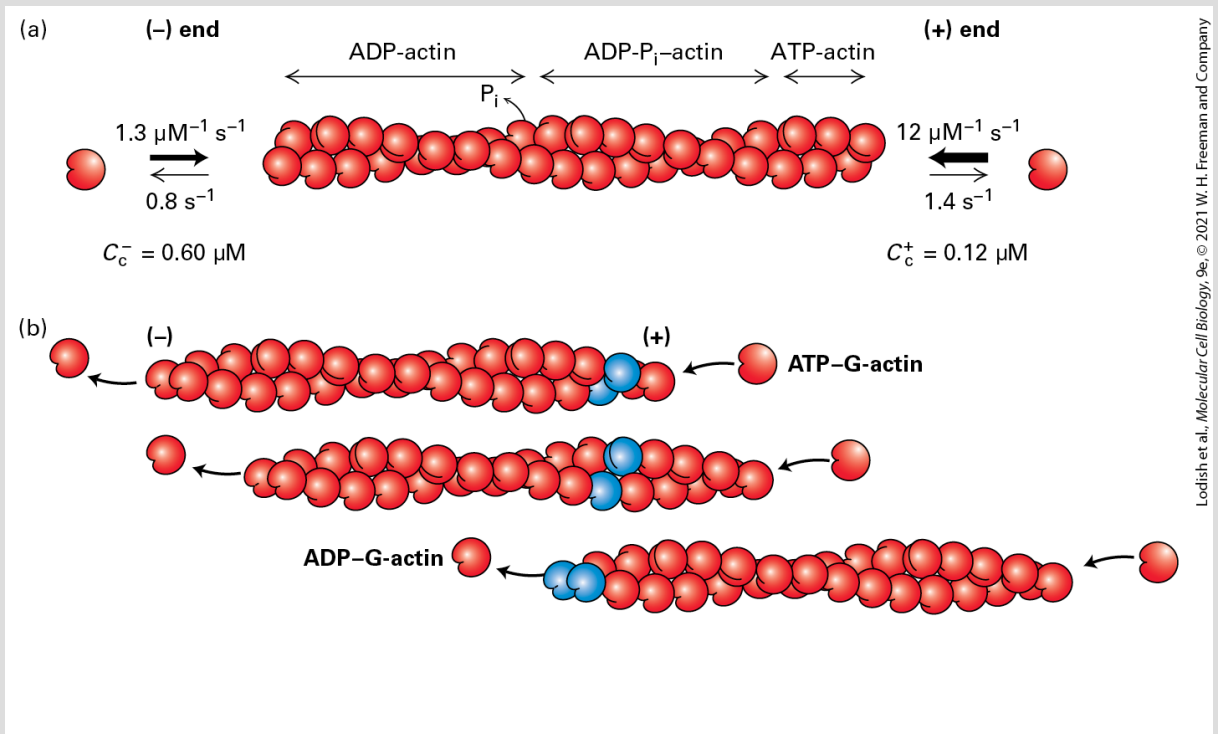
Using myosin S1 decoration experiments, we saw earlier that F-actin has an inherent structural polarity that is due to the uniform orientation of subunits in the filament (see Figures 17-5 and 17-6). As we described above, G-actin can bind ATP or ADP. When ATP–G-actin assembles into a filament, the ATP is rapidly hydrolyzed to ADP-P_i–actin, and then the P_i is more slowly released, generating a

filament mostly composed of ADP-P_i–actin and ADP-actin. Moreover, this gives rise to two different ends: the (+) end is composed of ATP-actin or ADP-P_i–actin, whereas the (−) end has ADP–G-actin. If free ATP–G-actin is added to such a preexisting myosin-decorated filament, the two ends grow at very different rates (Figure 17-9). In fact, the rate of addition of ATP–G-actin is nearly 10 times faster at the (+) end than at the (−) end. Kinetic experiments have shown that the rate of addition is about $12\,\mu M^{-1}\,s^{-1}$ at the (+) end and about $1.3\,\mu M^{-1}\,s^{-1}$ at the (−) end (Figure 17-10a). This means that if 1 μM of free ATP–G-actin is added to preformed filaments, 12 subunits, on average, will be added to the (+) end every second, whereas only 1.3 will be added to the (−) end every second. Note that the rate of addition (the association rate) is a function of the concentration of free ATP–G-actin. What about the rate of subunit loss from each end? By contrast with the rates of addition, the rates of dissociation of G-actin subunits from the two ends are quite similar: about $1.4\,s^{-1}$ from the (+) end and $0.8\,s^{-1}$ from the (−) end. Since this dissociation is simply the rate at which subunits leave the filament ends, it does not depend on the concentration of free ATP–G-actin.



Courtesy Thomas Pollard.

EXPERIMENTAL FIGURE 17-9 The two ends of a myosin-decorated actin filament grow at different rates. When short actin filaments are decorated with myosin S1 heads to reveal the orientation of the filaments and then used to nucleate actin polymerization, G-actin monomers are added much more efficiently to the (+) ends than to the (−) ends of the nucleating filaments.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 17-10 Actin treadmilling. ATP-actin subunits are added faster at the $(+)$ end than at the $(-)$ end of an actin filament, resulting in a lower critical concentration at the $(+)$ end and treadmilling at steady state. (a) The rate of addition of ATP–G-actin is much faster at the $(+)$ end than at the $(-)$ end, whereas the rate of dissociation of actin is similar at the two ends. This difference results in a lower critical concentration at the $(+)$ end. At steady state, ATP-actin is added preferentially at the $(+)$ end, giving rise to a short region of the filament containing ATP-actin and regions containing $\mathrm{ADP\text{-}P_i\text{–}actin}$ and ADP-actin toward the $(-)$ end. (b) At steady state, ATP–G-actin subunits add preferentially to the $(+)$ end, while ADP–G-actin subunits disassemble from the $(-)$ end, giving rise to treadmilling of subunits. For reference, two actin subunits are colored blue, to highlight that over time subunits appear to move from the plus end toward the minus end within the filament.

How do these association and dissociation rates impact actin dynamics? As we noted above, the association rate depends on the free ATP–G-actin concentration, whereas the dissociation rate does not. Thus when the free ATP–G-actin concentration is high, subunits will be added to the filament ends more quickly than subunits will be lost, so the filament will grow. As the concentration is lowered, a point will be reached where the rate of addition is balanced by the rate of loss, and no net filament growth occurs. Consider just the $(+)$ end for the moment. The free G-actin concentration at which rates of addition and loss are equal is called the *critical concentration for the* $(+)$ *end*, C_c^+. We can calculate C_c^+ by setting the rate of assembly equal to the rate of disassembly. Thus at the critical concentration the rate of assembly is C_c^+ times the measured rate of addition of $12\,\mu\mathrm{M}^{-1}\,\mathrm{s}^{-1}(C_c^+ \times 12\,\mathrm{s}^{-1})$, whereas the rate of disassembly is independent of the free actin concentration, namely, $1.4\,\mathrm{s}^{-1}$. Setting these two rates equal to each other yields $C_c^+ = 1.4\,\mathrm{s}^{-1}/12\,\mu\mathrm{M}^{-1}\,\mathrm{s}^{-1}$,

or $0.12\,\mu\text{M}$, for the $(+)$ end. Above this free ATP–G-actin concentration, subunits are added to the $(+)$ end and net growth occurs, whereas below it, there is a net loss of subunits, and shrinkage occurs.

Now let's consider the $(-)$ end. Because the rate of addition is much lower, $1.3\,\mu\text{M}^{-1}\,\text{s}^{-1}$, yet the rate of dissociation is about the same, $0.8\,\text{s}^{-1}$, we expect the critical concentration C_c^- at the $(-)$ end to be higher than C_c^+. Indeed, as we just did for the $(+)$ end, we can calculate C_c^- to be about $0.8\,\text{s}^{-1}/1.3\,\mu\text{M}^{-1}\,\text{s}^{-1}$, or $0.6\,\mu\text{M}$. Thus at less than $0.6\,\mu\text{M}$ free ATP–G-actin — say, $0.3\,\mu\text{M}$—the $(-)$ end will lose subunits. Notice that at this concentration the $(+)$ end will grow, because $0.3\,\mu\text{M}$ is above C_c^+. Because the critical concentrations for the two ends are different, at steady state the free ATP–G-actin will be intermediate between C_c^+ and C_c^-, so the $(+)$ end will grow and the $(-)$ end will lose subunits. This phenomenon is known as **treadmilling** because particular subunits, such as those shown in blue in Figure 17-10b, appear to move through the filament.

The treadmilling of actin filaments is powered by hydrolysis of ATP. When ATP–G-actin binds to the $(+)$ end of a filament, ATP is hydrolyzed to ADP and P_i. Since the P_i is slowly released from the subunits in the filament, the filament becomes organized into three regions: a short region of ATP-actin subunits at the $(+)$ end, ADP-P_i–actin in the middle of the filament, and, following P_i release,

ADP-actin subunits toward the (−) end (Figure 17-10a). During release of P_i from subunits in a filament, actin undergoes a conformational change that is responsible for the different association and dissociation rates at the two ends. Our analysis relies on a plentiful supply of ATP–G-actin, which, as we will see, turns out to be the case in vivo. Thus actin can use the power generated by hydrolysis of ATP to treadmill, and treadmilling filaments can do work in vivo, as we will see later.

Actin Filament Treadmilling Is Accelerated by Profilin and Cofilin

Actin treadmilling in vivo can occur at rates several times higher than happens with pure actin in vitro under physiological conditions. Why is treadmilling faster in vivo, and how does the cell recharge the ADP-actin dissociating from the (−) end into ATP-actin that adds at the (+) end? The answers involve two actin-binding proteins.

The first of these proteins is *profilin*, a small protein that binds G-actin on the side opposite the nucleotide-binding cleft (see the profilin cycle in Figure 17-11). When profilin binds ADP-actin, it opens the cleft and greatly enhances the loss of ADP. Because [ATP] in the cell is much greater than [ADP], ATP readily binds to G-actin, forming a profilin–ATP-actin complex. This complex cannot bind to

the (−) end because profilin blocks the sites on G-actin for (−) end assembly. However, the profilin–ATP-actin complex can bind efficiently to the (+) end. Once the new actin subunit is bound to the filament, profilin dissociates. Profilin activity does not enhance the treadmilling rate, but it does ensure a constant supply of ATP-actin formed from released ADP-actin. As a consequence, essentially all of the free G-actin in a cell has bound ATP.

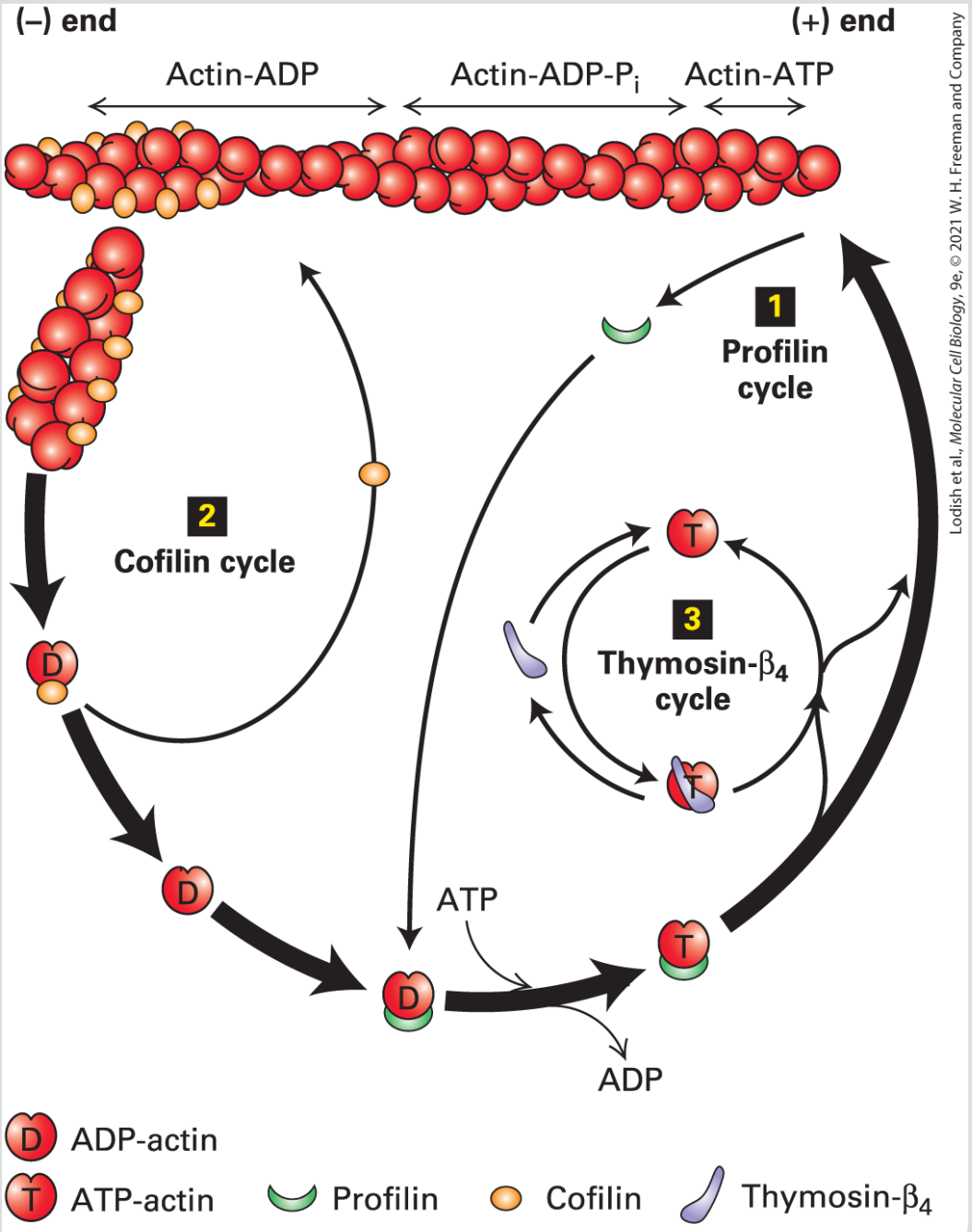


FIGURE 17-11 Regulation of filament turnover by actin-binding proteins. Actin-binding proteins regulate the rate of assembly and disassembly of actin filaments as well as the availability of G-actin for polymerization. In the profilin cycle (step 1), profilin binds ADP–G-actin and catalyzes the exchange of ADP for ATP. The profilin–ATP–G-actin complex can deliver actin to the (+) end of a filament with dissociation and recycling of profilin. In the cofilin cycle (step 2), cofilin binds preferentially to regions of filaments containing ADP-actin,

inducing them to fragment and thus enhancing depolymerization by making more filament ends. In the thymosin-β_4 cycle (step **3**), ATP–G-actin made available by the profilin cycle is bound by thymosin-β_4, which sequesters it from polymerization. As the free G-actin concentration is lowered by polymerization, G-actin–thymosin-β_4 dissociates to make free G-actin available for association with profilin and further polymerization.

Profilin has another important property: it can bind proteins with sequences rich in proline residues at the same time it is bound to actin. We will see later how this property can enhance the rate of actin filament assembly.

The second of the actin-binding proteins is **cofilin**, a small protein that binds specifically to ADP-actin subunits within F-actin. Recall from Figure 17-10a, these are the subunits found toward the (−) end of F-actin. Cofilin binds by bridging two actin monomers and inducing a small change in the twist of the filament. This small twist destabilizes the filament between regions with and without cofilin, breaking it into short pieces. By breaking the filament in this way, cofilin generates many more free (−) ends and therefore greatly enhances the net disassembly of the filament (see the cofilin cycle in Figure 17-11). The released ADP-actin subunits are then recharged by profilin and added to the (+) end as described above.

When purified profilin and cofilin are mixed with purified G-actin under conditions where filaments will form, rates of treadmilling in vitro increase more than tenfold compared to treadmilling in the

absence of these two actin-binding proteins. These enhanced levels of in vitro treadmilling are similar to what is seen in vivo.

Thymosin-β_4 Provides a Reservoir of Actin for Polymerization

It has long been known that cells often have a very large pool of G-actin, sometimes constituting as much as half the actin in the cell. Since cellular actin levels can be as high as 100–400 μM, this means that there can be as much as 50–200 μM unpolymerized actin in cells. Since the critical concentration in vitro is about 0.2 μM, why doesn't all this actin polymerize? The answer is, at least in part, due to G-actin sequestering proteins. One of these proteins is **thymosin-β_4**. When thymosin-β_4 binds to ATP–G-actin, it inhibits addition of the actin subunit to either end of a filament. As an example, let's consider human blood platelets. These disk-shaped cell fragments are very abundant, and when they are activated during blood clotting, they undergo a burst of actin assembly. Platelets are rich in actin; they are estimated to have a total concentration of about 550 μM actin, of which about 220 μM is in the unpolymerized form. They also contain about 500 μM thymosin-β_4, which sequesters much of the free actin. However, as in any protein-protein interaction, free actin and free thymosin-β_4 are in a dynamic equilibrium with actin–thymosin-β_4. When some free actin is incorporated into actin filaments, then more actin–thymosin-β_4 will

dissociate, thus maintaining a steady supply of actin subunits available for polymerization (see the thymosin-β_4 cycle in Figure 17-11). Thus thymosin-$\beta_4$behaves as a buffer of unpolymerized actin, making ATP-actin subunits available as they are needed.

Capping Proteins Block Assembly and Disassembly at Actin Filament Ends

Cells further regulate treadmilling and actin filament dynamics with *capping proteins* that bind to filament ends. If this were not the case, actin filaments would continue to grow and disassemble in an uncontrolled manner. As one might expect, two classes of proteins have been discovered: ones that bind the (+) end and ones that bind the (−) end (Figure 17-12).

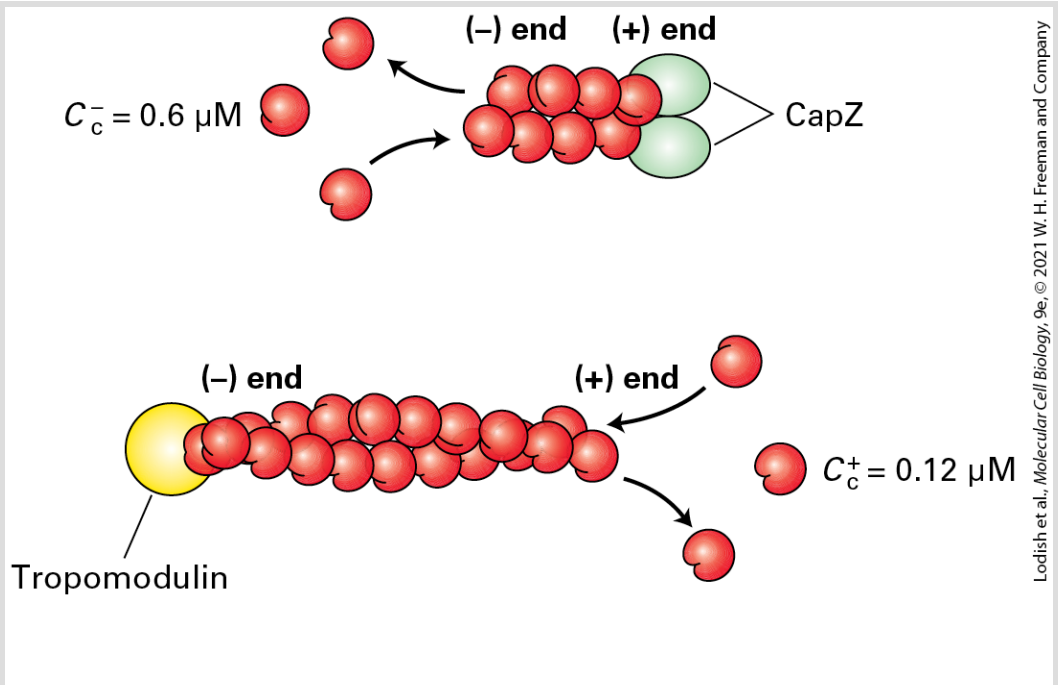


FIGURE 17-12 Filament capping proteins. Capping proteins block assembly and disassembly at filament ends. CapZ blocks the (+) end, which is where filaments normally grow, so its function is to limit actin dynamics to the (−) end. The capping protein tropomodulin blocks (−) ends, where filament disassembly normally occurs; thus the major function of tropomodulin is to stabilize filaments.

A protein known as **CapZ**, consisting of two closely related subunits, binds with a very high affinity ($\sim 0.1\,\mathrm{nM}$) to the (+) end of an actin filament, thereby inhibiting subunit addition or loss. The concentration of CapZ in cells is generally sufficient to rapidly cap any newly formed (+) ends. So how can filaments grow at their (+) ends? At least two mechanisms regulate the activity of CapZ. First, the capping activity of CapZ is inhibited by the regulatory phospholipid phosphatidylinositol 4,5-bisphosphate [$\mathrm{PI(4,5)P_2}$], found in the plasma membrane (see Chapter 16). Second, recent work has shown that certain regulatory proteins are able to bind the (+) end and protect it from CapZ while still allowing assembly there. Thus cells have evolved an elaborate mechanism to block assembly of actin filaments at their (+) ends except when and where assembly is needed.

Tropomodulin is one of the regulatory proteins that inhibit filament assembly and disassembly by binding to the (−) end of an actin filament. This protein is found predominantly in cells in which actin filaments need to be stabilized for long periods of time, such as the short actin filaments in the cortex of red blood cells and the actin

filaments in muscle. In both of these cells, tropomodulin works with the actin-binding protein tropomyosin to stabilize actin filaments. We will look at these protein interactions in more detail later in this chapter.

In addition to CapZ, another class of proteins can cap the $(+)$ ends of actin filaments. These proteins can also sever actin filaments. One member of this family, *gelsolin*, is regulated by Ca^{2+} ion concentration. On binding Ca^{2+}, gelsolin undergoes a conformational change that allows it to bind to the side of an actin filament and then insert itself between subunits of the helix, thereby breaking the filament. It then remains bound to and caps the $(+)$ end, generating a new $(-)$ end that can disassemble. As we discuss in a later section, actin cross-linking proteins can provide linkages between individual actin filaments to turn a solution of F-actin into a gel. If gelsolin is added to such a gel and the level of Ca^{2+} is elevated, gelsolin will sever the actin filaments and turn the gel back into a liquid solution. This ability to turn a gel into a sol is why the protein was named *gelsolin*.

KEY CONCEPTS OF SECTION 17.2

Dynamics of Actin Filaments

- The rate-limiting step in actin filament assembly is the formation of a short actin oligomer (nucleus) that can then be elongated into filaments.
- The critical concentration (C_c) is the concentration of G-actin below which actin filaments will not form.

- When the concentration of G-actin is above the C_c, the filament end will grow; when it is less than the C_c, the filament will shrink (see Figure 17-8).
- ATP–G-actin is added much faster at the $(+)$ end than at the $(-)$ end, resulting in a lower critical concentration at the $(+)$ end than at the $(-)$ end.
- At steady state, actin subunits treadmill through a filament. Treadmilling means ATP-actin is added at the $(+)$ end, ATP is hydrolyzed to ADP and P_i, P_i is lost slowly, and ADP-actin dissociates from the $(-)$ end.
- The length and rate of turnover of actin filaments is regulated by specialized actin-binding proteins (see Figure 17-11). Profilin enhances the exchange of ADP for ATP on G-actin; cofilin binds regions of ADP–F-actin to break the filament and generate new $(-)$ ends to enhance the rate of disassembly of ADP-actin from filament $(-)$ ends, and thymosin-β_4 binds G-actin to provide reserve G-actin when it is needed. Capping proteins bind to filament ends, blocking assembly and disassembly.

17.3 Mechanisms of Actin Filament Assembly

The rate-limiting step of actin polymerization is the formation of an initial actin nucleus from which a filament can grow (see Figure 17-7a). In cells, this step is used as a control point to determine where actin filaments are assembled and what types of actin structures are generated (see Figures 17-1 and 17-4). Two major classes of *actin-nucleating proteins*, the **formin** protein family and the **Arp2/3 complex**, nucleate actin assembly under the control of signal transduction pathways. Moreover, they nucleate the assembly of different actin structures: formins lead to the assembly of long actin filaments, whereas the Arp2/3 complex leads to branched networks of filaments. Here we discuss each separately and see how the power of actin polymerization can drive motile processes in a cell. We then touch on other specialized actin-nucleating factors.

Formins Assemble Unbranched Filaments

Formins are found in nearly all eukaryotic cells as a diverse family of proteins; seven different classes are present in vertebrates. Although they are diverse, all formin family members have two adjacent domains in common, the so-called FH1 and FH2 domains (formin-homology domains 1 and 2). Two FH2 domains from two

individual formin monomers associate to form a doughnut-shaped complex (Figure 17-13a). This complex has the ability to nucleate actin assembly by binding two actin subunits, holding them so that the (+) end of the nascent filament is toward the FH2 domains. The filament can now grow at the (+) end while the FH2 domain dimer remains attached to it. How is this possible? As we saw earlier, an actin filament can be thought of as two intertwined strands of subunits. The FH2 dimer can bind to the two terminal subunits. It then probably rocks between the two end subunits, letting go of one to allow addition of a new subunit and then binding the newly added subunit and freeing up space for the addition of another subunit to the other strand. In this way, rocking between the two subunits on the end, it can remain attached while simultaneously allowing growth at the (+) end (Figure 17-13a).

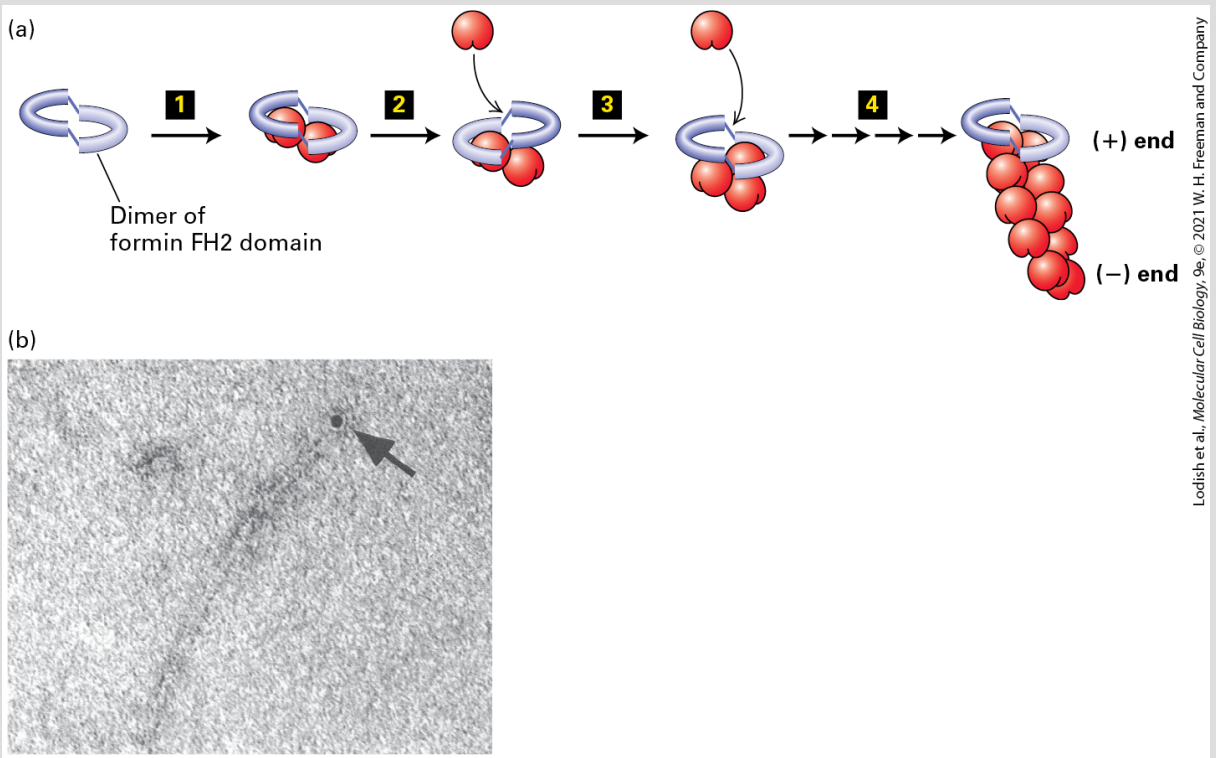


FIGURE 17-13 Actin nucleation by the formin FH2 domain. (a) Formins have a domain called *FH2* that can form a dimer and nucleate filament assembly. The dimer binds two actin subunits (step **1**) and, by rocking back and forth (steps **2**–**4**), can allow insertion of additional subunits between the FH2 domain and the (+) end of the growing filament. The FH2 domain protects the (+) end from being capped by capping proteins. (b) The FH2 domain of a formin was labeled with colloidal gold (black dot) and used to nucleate assembly of an actin filament. The resulting filament was visualized by electron microscopy after staining with uranyl acetate. Formins assemble long unbranched filaments.

[Part (b) Republished with permission from AAAS, from D. Pruyne et al., 2002, "Role of Formins in Actin Assembly: Nucleation and Barbed-End Association," *Science* **297**(5581): 612–615; permission conveyed through Copyright Clearance Center, Inc.]

The FH1 domain adjacent to the FH2 domain also makes an important contribution to actin filament growth (Figure 17-14). This domain is rich in proline residues, which serve as sites for the binding of several profilin molecules. We discussed earlier how

profilin can exchange the ADP nucleotide on G-actin for ATP to generate profilin–ATP-actin. The FH1 domain behaves as a landing site for profilin–ATP–G-actin to increase the local concentration of these complexes. The actin from these profilin-actin complexes is then fed into the FH2 domain to add actin to the (+) end of the filament, and the profilin is released (Figure 17-14).

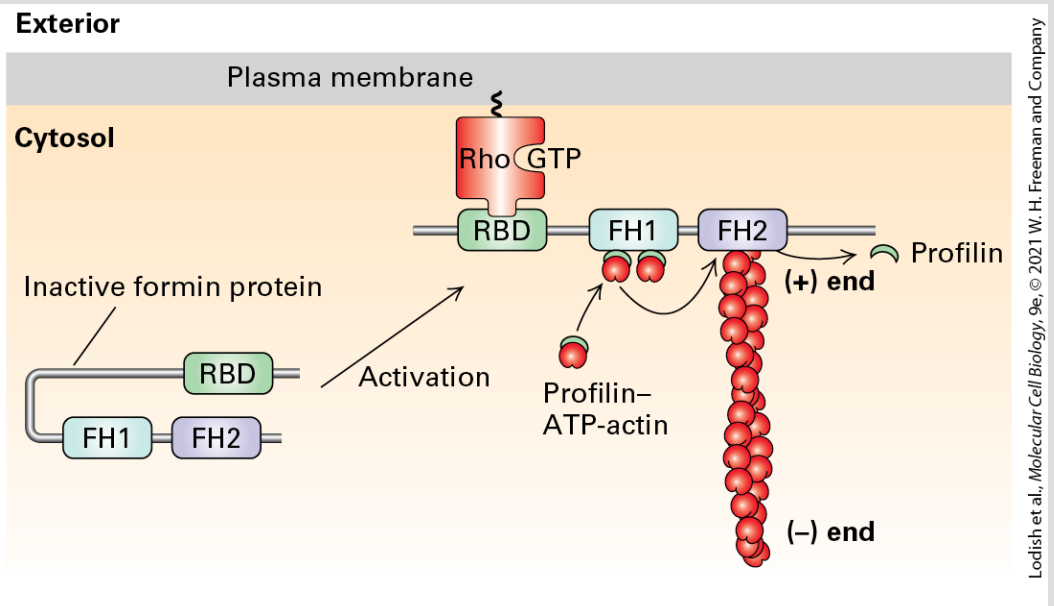


FIGURE 17-14 Regulation of formin activity by an intramolecular interaction. Some of the formin classes found in vertebrates are regulated by an intramolecular interaction. In its inactive state, the formin folds back on itself to inhibit the activity of the FH2 domain. The inactive formin is activated when its Rho-binding domain (RBD) binds to membrane-bound active Rho-GTP, resulting in exposure of the formin's FH2 domain, which can then nucleate the assembly of a new actin filament. All formins have an FH1 domain adjacent to the FH2 domain; the proline-rich FH1 domain is a site for recruitment of profilin–ATP–G-actin complexes that can then be fed into the growing (+) end. For simplicity of representation, a single formin protein is shown, but as shown in Figure 17-13, the FH2 domain functions as a dimer to nucleate actin assembly. Regulation of the Rho family of small GTPases is detailed in Figures 17-40 and 17-42.

Since formin allows rapid addition of actin subunits to the (+) end, long filaments with a formin at their (+) end are generated (Figure 17-13b). Thus formins not only nucleate actin assembly, but they remain bound to the (+) end and, in conjunction with profilin, they facilitate rapid filament assembly. To ensure continued growth of a filament, formins bind to the (+) end in such a way that precludes binding of a (+) end capping protein such as CapZ, which would terminate assembly.

To be useful to a cell, formin activity must be regulated. Many formins exist in a folded, inactive conformation as a result of an interaction between the first half of the protein and its C-terminal tail. These formins are activated by membrane-bound Rho-GTP, a Ras-related small GTPase (discussed in Section 17.7). When Rho switches from its inactive Rho-GDP form to its activated membrane-bound Rho-GTP state, it can bind and activate these formins (see Figure 17-14).

Formins are responsible for the assembly of long actin filaments such as those found in muscle cells, stress fibers, filopodia, and the contractile ring that forms during cytokinesis (see Figure 17-4). The different formin classes in animals differ in their ability to nucleate actin assembly, and some formins can also bundle actin filaments.

The Arp2/3 Complex Nucleates Branched Filament Assembly

The Arp2/3 complex is a protein machine consisting of seven subunits, two of which are *a*ctin-*r*elated *p*roteins (Arp; Figure 17-15a). It is found in essentially all eukaryotes, including plants, yeasts, and animal cells. The Arp2/3 complex alone is a very poor nucleator. To nucleate the assembly of branched actin filaments, Arp2/3 needs to be activated by a **nucleation promoting factor (NPF)**. The major NPF family is characterized by the presence of a region called *WCA* (*W*H2, *c*onnector, *a*cidic for the acidic amino acid residues it contains). Experiments have shown that if the WCA domain is added to an actin assembly assay together with preformed actin filaments, Arp2/3 becomes a potent nucleator of further actin assembly.

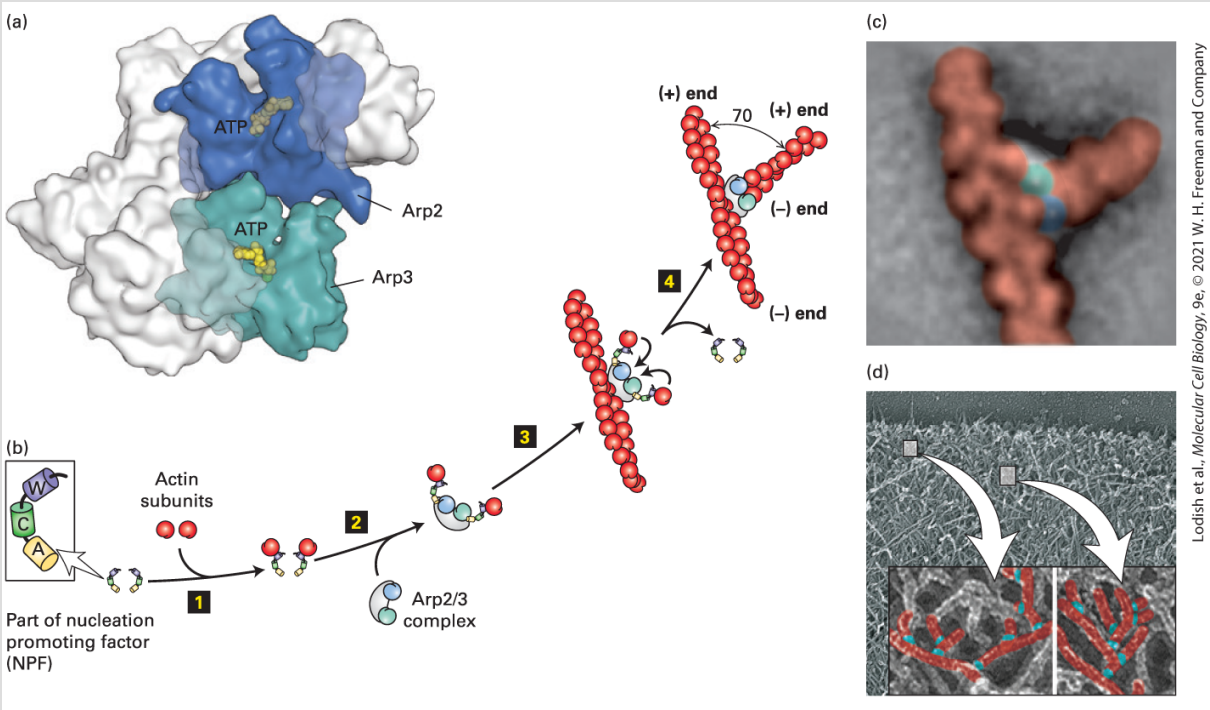


FIGURE 17-15 Actin nucleation by the $\mathrm{Arp2/3}$ complex. (a) X-ray crystallographic structure of the $\mathrm{Arp2/3}$ complex, with five of the subunits in gray and the Arp2 and Arp3 subunits in green and blue. (b) To nucleate actin assembly efficiently, $\mathrm{Arp2/3}$ must interact with the activating part of an NPF, shown here with its W (WH2), C (connector), and A (acidic) domains. Step **1** involves binding of an actin subunit to the W domain of each NPF. Two NPF-actin complexes then bind the $\mathrm{Arp2/3}$ complex (step **2**). This interaction induces a conformational change in the $\mathrm{Arp2/3}$ complex. After binding of the $\mathrm{Arp2/3}$ complex to the side of an actin filament, the actin subunits delivered by the W domains bind to the $\mathrm{Arp2/3}$ complex (step **3**), which then initiates the assembly of an actin filament at the available $(+)$ end (step **4**). The $\mathrm{Arp2/3}$ branch makes a characteristic 70° angle between the old and new filaments. (c) Averaged image compiled from several electron micrographs of $\mathrm{Arp2/3}$ at an actin branch. (d) Image of actin filaments in the leading edge of a motile cell, with a magnification and coloring of individual branched filaments.

[Part (a) Data from B. J. Nolen and T. D. Pollard, 2007, *Mol. Cell* **26**:449–457, PDB ID 2p9l. Part (c) C. Egile et al., 2005, "Mechanism of Filament Nucleation and Branch Stability Revealed by the Structure of the Arp2/3 Complex at Actin Branch Junctions." PLoS Biol. 3(11):e383; https://doi.org/10.1371/journal.pbio.0030383. Part (d) T. M. Svitkina and G. G. Borisy, 1999, *J. Cell Biol.* **145**:1009–1026; https://doi.org/10.1083/jcb.145.5.1009.]

How do the Arp2/3 complex and NPFs nucleate the assembly of actin filaments? Two NPFs each bind an actin subunit at their WH2 domains, and together, they activate the Arp2/3 complex through its interaction with their connector and acidic domains. In the inactive Arp2/3 complex, the two actin-related polypeptides — Arp2 and Arp3 — are in the wrong configuration to nucleate filament assembly (see Figure 17-15b, step **2**). When activated by the NPFs, Arp2 moves into a new configuration permitting the complex to bind to the side of a preexisting actin filament. The actin subunits brought in by the WH2 domains of the NPFs bind to the Arp2/3 template to nucleate filament assembly at the (+) end (Figure 17-15b, step **3**). The NPFs are released, and the new (+) end then grows as long as ATP–G-actin is available or until it is capped by a (+) end capping protein such as CapZ. The angle between the old filament and the new one is 70° (Figure 17-15b, c). This is also the angle observed experimentally in branched filaments at the leading edges of motile cells, which are believed to be formed by the action of the activated Arp2/3 complex (Figure 17-15d). As we discuss in subsequent sections, the Arp2/3 complex can be used to drive actin polymerization that powers intracellular motility.

Actin nucleation by the Arp2/3 complex is finely controlled, and NPFs are part of that regulatory process. One NPF is called **WASp**, as it is defective in patients with Wiskott-Aldrich syndrome, an X-linked disease characterized by eczema, low platelet count, and immune deficiency. WASp exists in a folded inactive conformation

that makes the WCA domain unavailable (Figure 17-16). To ensure that WASp is activated only at the plasma membrane, its activation requires two signals. One signal is binding between WASp and the regulatory phospholipid $PI(4,5)P_2$,which is characteristically enriched in the plasma membrane (Chapter 16). WASp binds $PI(4,5)P_2$ through its basic domain. The second signal is binding of the activated form of the small GTP-binding protein Cdc42 to the RBD region of WASp. GTP-Cdc42 is itself activated in response to certain signaling pathways (discussed in Section 17.7). This type of two-signal input, called *coincidence detection*, ensures that the protein is activated only at the right place — at the plasma membrane — and by the right signaling pathway. Once bound to the two input signals, the conformation of WASp opens up, and the WCA domain becomes accessible.

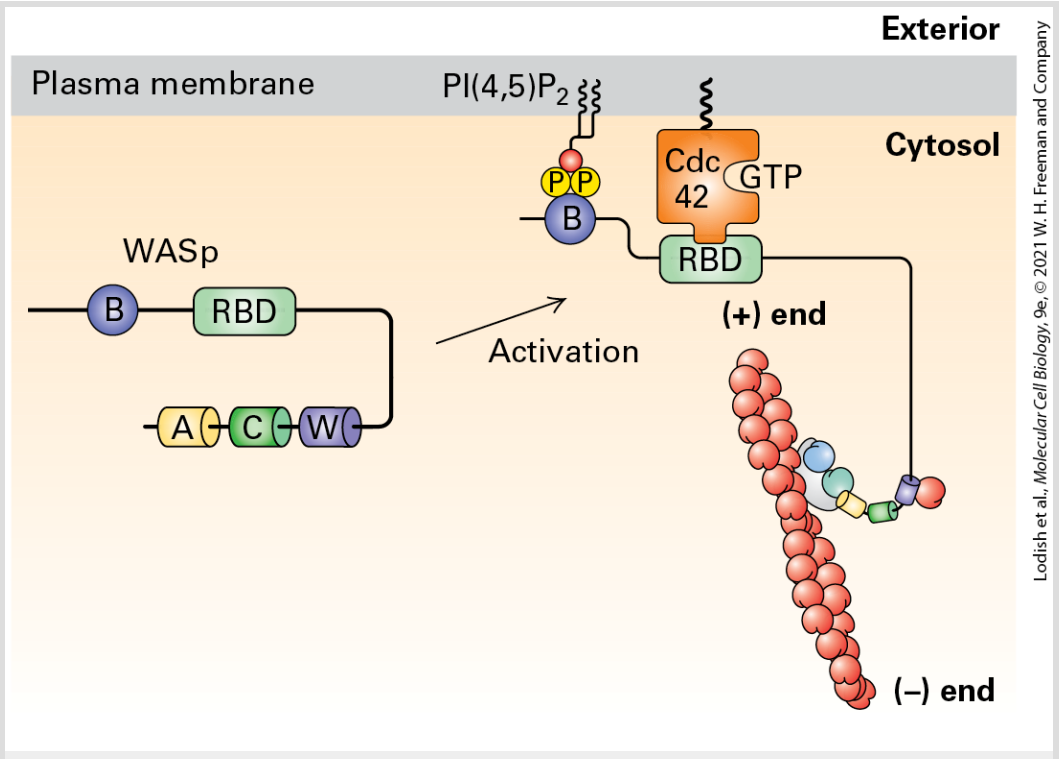


FIGURE 17-16 Regulation of the $\mathrm{Arp2/3}$ complex by WASp and PI(4,5)P2. The NPF WASp is inactive due to an intramolecular interaction that masks the WCA domain. It is activated by a coincidence detection mechanism: it must bind both the regulatory phospholipid PI(4,5)P2 though its basic domain (B) and the membrane-bound active small G protein Cdc42-GTP (a member of the Rho family) through its Rho-binding domain (RBD). When activated in this way, the intramolecular interaction in WASp is relieved, allowing the W domain to bind actin and the acidic A domain to activate the $\mathrm{Arp2/3}$ complex. For simplicity, only a single NFP-$\mathrm{Arp2/3}$ interaction is shown. Regulation of the Rho family of small GTPases is detailed in Figures 17-40 and 17-42.

A second NPF is a large protein complex called *WAVE*, which also has a WCA domain that can activate the $\mathrm{Arp2/3}$ complex. WAVE is also activated by coincidence detection, specifically by binding to acidic phospholipids and to the active form of another small GTP-binding protein, Rac1. As we discuss in Section 17.7, activation of the $\mathrm{Arp2/3}$ complex by Cdc42 through WASp and by Rac1 through WAVE induces the formation of different microfilament-based structures.

Although formins and the $\mathrm{Arp2/3}$ complex are found in fungi, plants, and animals, additional actin nucleators have been discovered in animal cells. One of these, called *Spire*, has four tandem WH2 domains, so it can bind four actin monomers. It does this in a manner that allows the actins to assemble into a filament, although the detailed mechanism remains to be understood. Given that actin filaments perform so many functions in cells, it is not

surprising that additional NPF proteins and actin nucleators have recently been described.

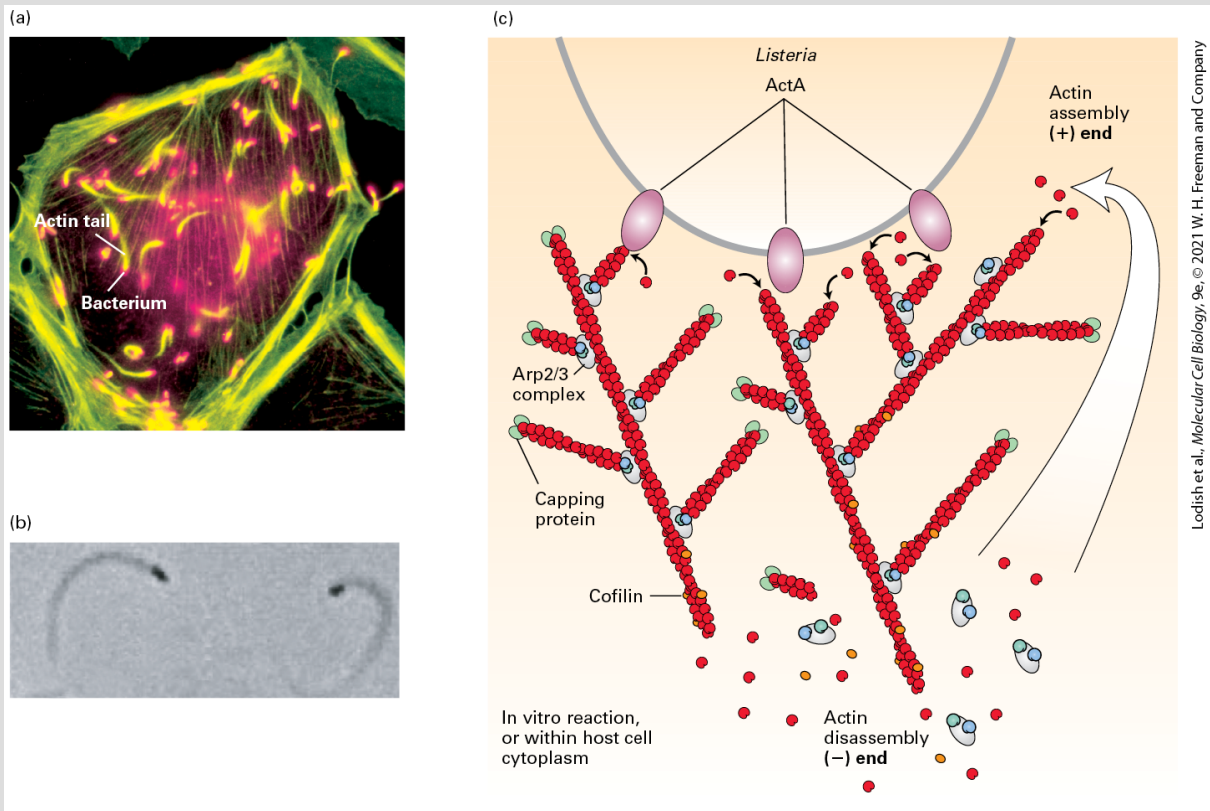
Intracellular Movements Can Be Powered by Actin Polymerization

How can actin polymerization be harnessed to do work? As we have seen, actin polymerization involves the hydrolysis of ATP–G-actin to ADP–G-actin, which allows actin filaments to grow preferentially at the (+) end and disassemble at the (−) end. If a collection of actin filaments were to become immobilized in the network of the cytoskeleton and you could bind to and ride on the assembling (+) ends, you would be transported across the cell. This is just what the intracellular bacterial parasite *Listeria monocytogenes* does to get around the cell. The study of *Listeria* motility was, in fact, the way the nucleating activity of the Arp2/3 protein was discovered. As we will see shortly, *Listeria* has hijacked a normal cell motility process for its own purposes.

Listeria is a food-borne bacterial pathogen that causes mild gastrointestinal symptoms in most adults but can be fatal to elderly or immunocompromised individuals. It enters animal cells and divides in the cytoplasm. Its rapid movement around the cell is powered by massive local polymerization of actin to resemble a comet tail (Figure 17-17a), and when it runs into the plasma

membrane, it continues to use the power of actin polymerization to push its way into the adjacent cell. Since the comet tail is embedded in the stationary cytoskeletal matrix of the cell, the bacterium is pushed forward, ahead of the polymerizing actin. Cytosolic factors then disassemble the tail so that actin subunits can be recycled for further polymerization. To do this, *Listeria* needs to direct the assembly of host-cell actin at one end of the bacterium and at the same time confine assembly there so that it efficiently pushes the bacterium forward. How does it do it? *Listeria* has on its surface a protein called *ActA*, which mimics an NPF by having an actin-binding site and an acidic region that efficiently activates the Arp2/3 complex. Researchers have reconstituted *Listeria* motility in the test tube using purified proteins to see what are the minimal requirements for *Listeria* motility. Remarkably, the bacteria will move when just four proteins are added: ATP–G-actin, the Arp2/3 complex, CapZ, and cofilin (see Figure 17-17b, c). We have discussed the role of actin and Arp2/3, but why are CapZ and cofilin needed? As we have seen earlier, CapZ rapidly caps the free (+) end of actin filaments, so when a growing filament no longer contributes to bacterial movement, it is rapidly capped and inhibited from further elongation. In this way, assembly occurs mostly adjacent to the bacterium, where ActA is stimulating the Arp2/3 complex. Cofilin is necessary to accelerate the disassembly of the (−) end region of the actin filament, regenerating free actin to keep the polymerization cycle going (see Figure 17-11). This minimal rate of motility can be increased by the inclusion of VASP and profilin. The ActA protein

binds the host protein VASP, which has three important properties. First, VASP has a proline-rich region that can bind profilin–ATP-actin and thus enhance ATP-actin assembly into the newly formed (+) ends generated by the Arp2/3 complex. Second, it can hold onto the end of the newly formed filament. Third, it can protect the (+) end of the growing filament from capping by CapZ. These properties allow VASP to enhance actin assembly and confine it to the rear of the bacterium.



EXPERIMENTAL FIGURE 17-17 Listeria uses the power of actin polymerization for intracellular movement. (a) Fluorescence microscopy of a cultured cell stained with an antibody to a bacterial surface protein (red) and fluorescent phalloidin to localize F-actin (green). Behind each *Listeria* bacterium is an actin "comet tail" that becomes embedded in the stationary cytoskeletal matrix and so propels the bacterium forward by actin polymerization. When the bacterium runs into the plasma membrane, it pushes the

membrane out into a structure like a filopodium, which protrudes into a neighboring cell. (b) *Listeria* motility can be reconstituted in vitro with bacteria and just four proteins: ATP–G-actin, $\mathrm{Arp2/3}$ complex, CapZ, and cofilin. This phase-contrast micrograph shows bacteria (black), behind which are the phase-dense actin tails. (c) A model of how *Listeria* moves using just four proteins. The ActA protein on the bacterial cell surface activates the $\mathrm{Arp2/3}$ complex to nucleate new filament assembly from preexisting filaments. Filaments grow at their $(+)$ end until capped by CapZ. Actin is recycled through the action of cofilin, which enhances depolymerization at the $(-)$ end of the filaments. In this way, polymerization is confined to the back of the bacterium and propels it forward.

[Part (a) Courtesy Julie Theriot and Timothy Michison. Part (b) Reprinted with permission from Nature Publishing Group, from T. P. Loisel et al., 1999, "Reconstitution of Actin-Based Motility of *Listeria* and *Shigella* Using Pure Proteins," *Nature* **401**:613–616; permission conveyed through Copyright Clearance Center, Inc.]

To move inside cells, the *Listeria* bacterium hijacks a normal, regulated cellular process involved in cell locomotion. As we discuss in more detail later (Section 17.7), moving cells have a thin sheet of cytoplasm at the front of the cell called the *leading edge* (see Figures 17-1c, 17-4, and 17-15d). This thin sheet of cytoplasm consists of a dense network of actin filaments that are continually elongating at the front of the cell to push the membrane forward. Factors in the leading-edge membrane activate the $\mathrm{Arp2/3}$ complex to nucleate these filaments. Thus the power of actin assembly pushes the membrane forward to contribute to cell locomotion. It is the leading-edge machinery that is co-opted by *Listeria* as it moves within and between host cells.

Microfilaments Function in Endocytosis

As we saw in Chapter 14, endocytosis is the process that cells use to take up particles, molecules, or fluid from the external medium by enclosing them in plasma membrane and then internalizing them. The uptake of molecules or liquid is called *receptor-mediated* or *fluid-phase endocytosis*, and the uptake of large particles is called *phagocytosis* ("cell eating"). Microfilaments participate in both of these processes.

Fluid-phase endocytosis is a multistep process, involving collection of cargo proteins in a clathrin-coated pit, deformation and invagination of the membrane by the clathrin coat, pinching off the membrane by the dynamin GTPase, and movement of the endosome into the cell interior (see Figures 14-20 and 14-29). The power of actin assembly contributes to this process. Endocytosis assembly factors recruit NPFs so that as the endocytic vesicles invaginate and pinch off from the membrane, they are driven into the cytoplasm, powered by a rapid and short-lived burst (a few seconds in duration) of actin polymerization driven by the Arp2/3 complex (Figure 17-18a). Although the physical mechanism that drives the endocytic vesicle into the cytoplasm is not fully resolved, the actin-based machinery is similar to that used in leading-edge formation and *Listeria* motility. This actin-based movement of endocytic vesicles can be reconstituted in vitro (Figure 17-18b).

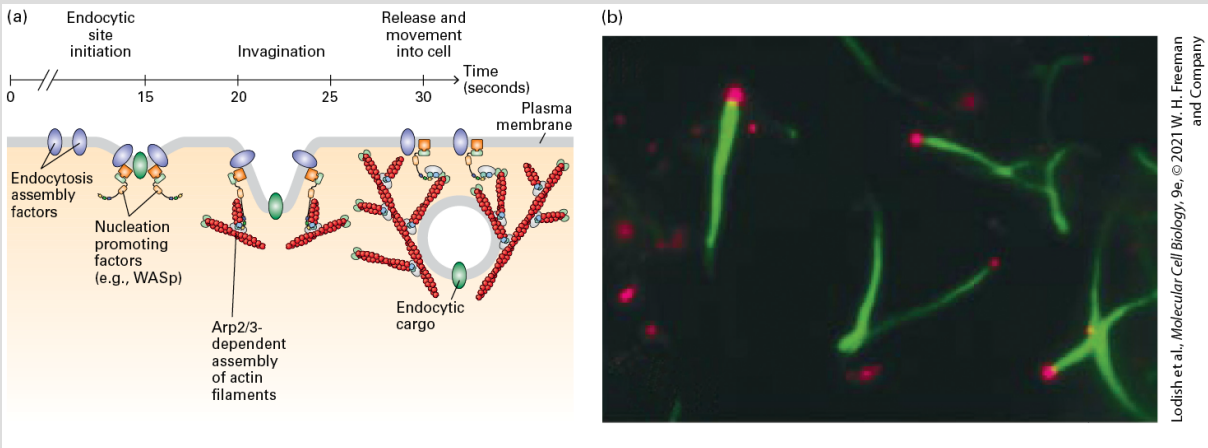
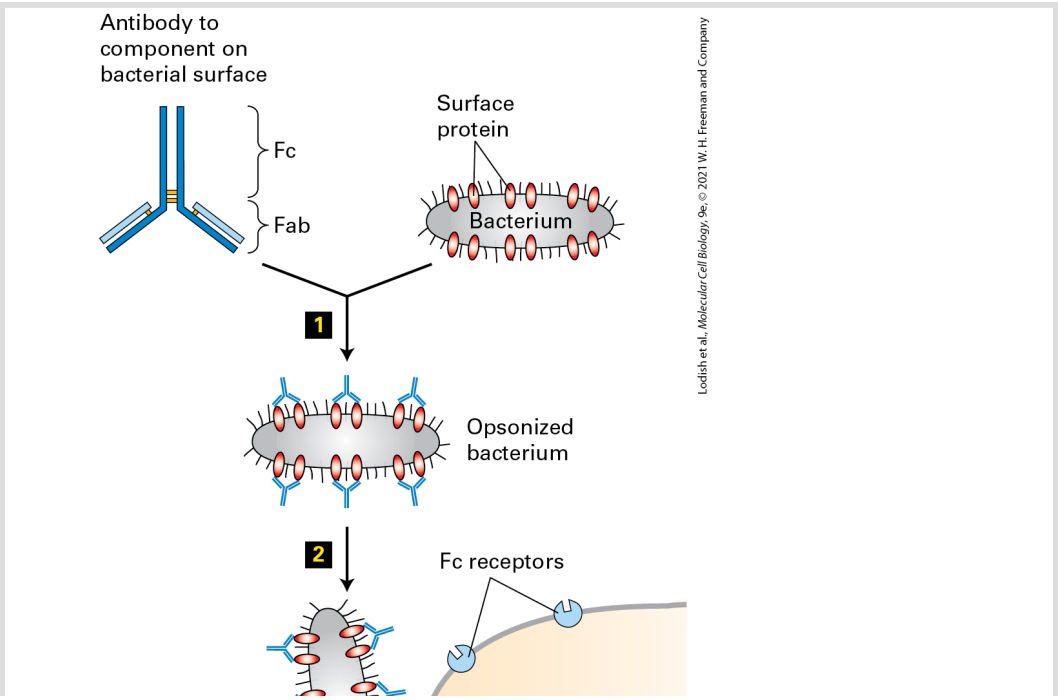


FIGURE 17-18 $\mathrm{Arp2/3}$-dependent actin assembly during endocytosis. (a) Clathrin-mediated endocytosis is a rapid and ordered process. It has been best studied in yeast, in which the temporal order of specific steps has been delineated. In vivo imaging has shown that endocytosis assembly factors recruit NPFs that activate the $\mathrm{Arp2/3}$ complex. A burst of $\mathrm{Arp2/3}$-dependent actin assembly drives internalization of endocytic vesicles away from the plasma membrane. For clarity, other factors involved in endocytosis, such as clathrin, the adaptin AP2 and dynamin are not shown (see Figure 14-20). (b) Endosome movement can be reconstituted in vitro. Endosomes isolated from cells that had taken up fluorescently labeled transferrin (red) were added to a cell extract containing fluorescently labeled actin (green). The endosomes bind WASp, which then activates the $\mathrm{Arp2/3}$ complex to assemble actin tails that propel them through the cytoplasm.

[Part (b) J. Taunton et al., 2000, *J. Cell Biol*. **148**(3):519–530; https://doi.org/10.1083/jcb.148.3.519.]

Phagocytosis by leukocytes (white blood cells) is a vital process in the recognition and removal of pathogens, such as bacteria. The immune system identifies a bacterium as foreign and makes antibodies that recognize surface features of the bacterium. As we discussed in Chapter 3, each antibody has a region called the *Fab domain* that binds specifically to its antigen, in this case a component on the bacterial cell surface. As antibodies coat the

bacterium through interaction between their Fab domains and the cell-surface antigen, a second antibody domain, known as the *Fc domain*, is exposed. This process is known as *opsonization* (Figure 17-19, step **1**; see Chapter 24). Leukocytes have a receptor on their cell surface, called the *Fc receptor*, that binds the Fc portion of the bacterium-bound antibody (step **2**). Initial binding between the leukocyte and opsonized bacterium signals the leukocyte to engulf the bacterium (steps **3** and **4**) inducing assembly of microfilaments at the site of interaction with the bacterium. The assembled microfilaments, together with myosin motor proteins, provide the force necessary to draw the bacterium into the cell, ultimately fully enclosing the pathogen in plasma membrane (step **4**). Once internalized, the newly formed phagosome fuses with lysosomes, and the pathogen is killed and degraded by lysosomal enzymes.



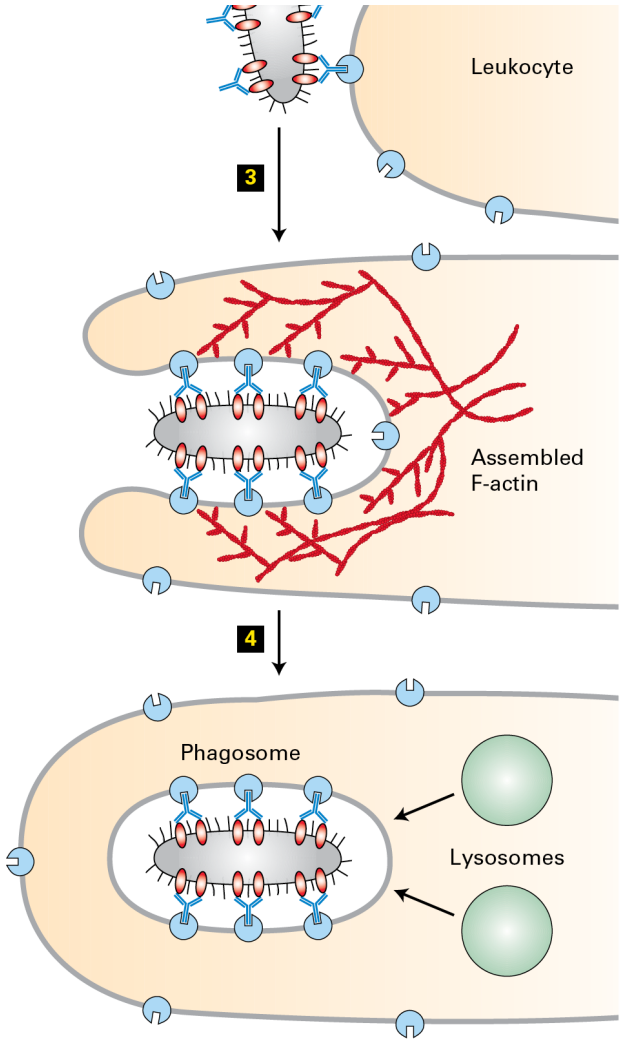


FIGURE 17-19 Phagocytosis and actin dynamics. Actin assembly and contraction drives the phagocytic internalization of particles. Shown here is the phagocytosis and degradation of a bacterium by a leukocyte. An invading bacterium is coated by specific antibodies to a cell-surface protein in a process known as *opsonization* (step 1). The Fc region of the bound antibodies is displayed on the bacterial surface and recognized by a specific receptor, the Fc receptor, on the leukocyte surface (step 2). This interaction signals the cell to assemble a contractile actin structure that results in the internalization and engulfment of the bacterium (step 3). Once it has been internalized into a phagosome, the bacterium is killed and degraded by enzymes delivered from lysosomes (step 4).

Actin assembly has been implicated in other steps of the endocytic pathway as well as in the secretory pathway. For example, an NPF called *WASH* has been implicated in the Arp2/3-dependent nucleation of actin filament on endosomes, regulating their shape and contributing to transport events. Another NPF, called *WHAMM*, localizes to the Golgi complex and is believed to direct the Arp2/3-dependent assembly of actin participating in membrane traffic from the endoplasmic reticulum to the Golgi. The emerging view is that actin assembly facilitates transport between different compartments of the secretory and endocytic pathways.

Toxins That Perturb the Pool of Actin Monomers Are Useful for Studying Actin Dynamics

Certain fungi and sponges produce toxins that target the polymerization cycle of actin and are therefore toxic to animal cells but useful for the study of actin dynamics. Two types of these toxins have been characterized. The first class is represented by two unrelated toxins, cytochalasin D and latrunculin, that promote the depolymerization of filaments, though by different mechanisms. Cytochalasin D, a fungal alkaloid, depolymerizes actin filaments by binding to the (+) end of F-actin, where it blocks further addition of subunits. Latrunculin, a toxin secreted by sponges, binds and sequesters G-actin, inhibiting it from adding to a filament end.

Exposure to either toxin thus increases the monomer pool. When cytochalasin D or latrunculin is added to live cells, the actin cytoskeleton disassembles and cell movements such as locomotion and cytokinesis are inhibited. These observations were among the first to implicate actin filaments in cell motility. Latrunculin is especially useful because it binds actin monomers and prevents any new actin assembly. Thus if latrunculin is added to a cell, the rate at which actin-based structures disappear reflects their normal rate of turnover. This method has revealed that the stability of actin structures varies over a wide range. For example, experiments with latrunculin show that the leading edges of motile cells turn over every 30–180 seconds, and that stress fibers turn over every 5–10 minutes.

In the other class of toxins, the monomer-polymer equilibrium is shifted in the direction of filaments. Examples from this class of toxins are jasplakinolide, another sponge toxin, and phalloidin, which is isolated from *Amanita phalloides* (the "angel of death" mushroom). Jasplakinolide enhances nucleation by binding and stabilizing actin dimers and thereby lowering the critical concentration. Phalloidin binds at the interface between subunits in F-actin, locking adjacent subunits together and preventing actin filaments from depolymerizing. Even when actin is diluted below its critical concentration, phalloidin-stabilized filaments will not depolymerize. Because many actin-based processes depend on actin filament turnover, the introduction of phalloidin into a cell

paralyzes these systems, and the cell dies. Since phalloidin binds only to F-actin, fluorescently labeled phalloidin is commonly used to stain actin filaments for light microscopy (see Figure 17-4).

KEY CONCEPTS OF SECTION 17.3

Mechanisms of Actin Filament Assembly

- Actin assembly is nucleated by two classes of proteins: formins nucleate the assembly of unbranched filaments (see Figure 17-13), whereas the $\mathrm{Arp2/3}$ complex nucleates the assembly of branched actin networks (see Figure 17-15). The activities of formins and $\mathrm{Arp2/3}$ are regulated by signal transduction pathways.
- Functionally different actin-based structures are assembled by formins and by $\mathrm{Arp2/3}$ nucleators. Formins drive the assembly of stress fibers and the contractile ring, whereas $\mathrm{Arp2/3}$ nucleates the assembly of the branched actin filaments found in the leading edge of motile cells.
- The power of actin polymerization can be harnessed to do work, as is seen in the $\mathrm{Arp2/3}$-dependent intracellular movement of pathogenic bacteria (see Figure 17-17) and the inward movement of endocytic vesicles (see Figures 17-18 and 17-19).
- Several toxins affect the dynamics of actin polymerization; some, such as latrunculin, bind and sequester actin monomers, whereas others, such as phalloidin, stabilize filamentous actin. Fluorescently labeled phalloidin is useful for staining actin filaments for their visualization by microscopy.

17.4 Organization of Actin-Based Cellular Structures

We have seen that actin filaments are assembled into a wide variety of different arrangements and that many actin-associated proteins nucleate actin assembly and regulate filament turnover. Dozens of proteins in a vertebrate cell organize these filaments into diverse structures, which the cell uses for an array of functions. In this section, we discuss a few of these proteins, including typical types of actin cross-linking proteins and adapter proteins involved in making links between actin and membrane proteins. One fascinating problem, about which very little is known, is how cells assemble different actin-based structures within the cytoplasm of the same cell. Some of this organization is due to local regulation of actin assembly, a topic we come to at the end of the chapter.

Cross-Linking Proteins Organize Actin Filaments into Bundles or Networks

When actin filaments are assembled in a test tube, they form a tangled network. In cells, however, actin filaments are found in organized structures, such as the highly ordered filament bundles in microvilli or the network characteristic of the leading edge of a motile cell (see Figure 17-4a). These structures are erected with the

help of the filament assembly mechanisms discussed previously and by the presence of **actin cross-linking proteins**.

To be able to organize actin filaments, an actin cross-linking protein must have two F-actin–binding sites (Figure 17-20a). Actin cross-linking proteins can have two actin-binding sites within a single polypeptide, as with *fimbrin*, which builds bundles of filaments all having the same polarity, as seen in the stereocilium (a giant microvillus) in Figure 17-20b. Like fimbrin, *filamin* has dual actin-binding sites in a single polypeptide, but unlike fimbrin, which is short and stiff, filamin has a flexible region between the binding sites that functions like a molecular leaf spring. This enables filamin to form stabilizing cross-links between filaments in a network, such as the ones found in the leading edge of a motile cell.

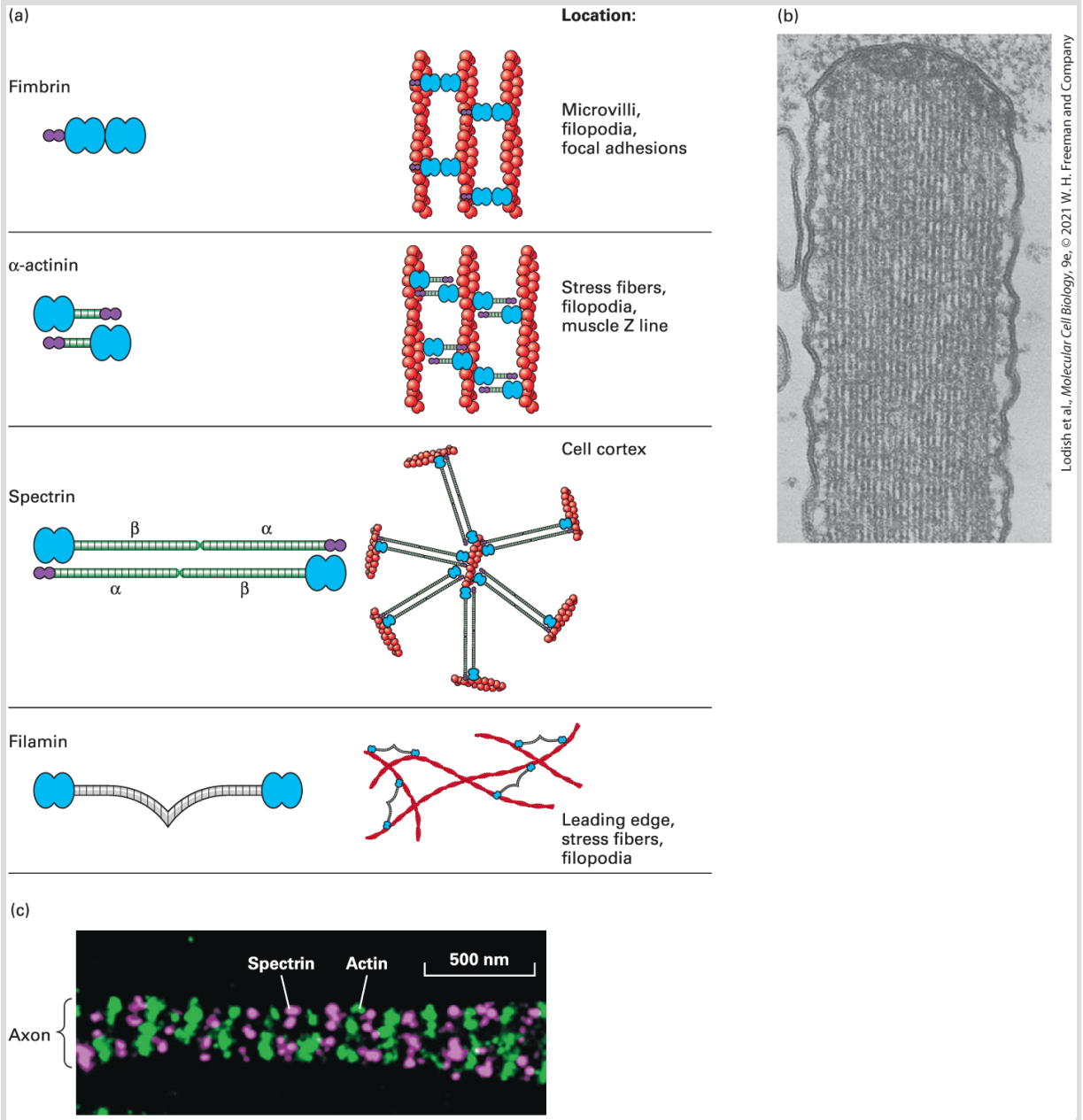


FIGURE 17-20 Actin cross-linking proteins. Actin cross-linking proteins mold F-actin filaments into diverse structures. (a) Examples of four F-actin cross-linking proteins, all of which have two domains (blue) that bind F-actin. Some have a Ca^{2+}-binding site (purple) that inhibits their activity at high levels of free Ca^{2+}. (b) Transmission electron micrograph of a thin section of a stereocilium (an unfortunate name, since it is really a giant microvillus) on a sensory hair cell in the inner ear. This structure contains a bundle of actin filaments cross-linked by fimbrin, a small cross-linking protein that allows for close and regular interaction of actin filaments. (c) Super-resolution image of the repeating structure of actin

(green) and spectrin (magenta) that form a cagelike network to support the plasma membrane of an axon.

[Part (b) L. G. Tilney et al., 1983, *J. Cell Biol.* **96**(3):822–834; https://doi.org/10.1083/jcb.96.3.822. Part (c) Courtesy of Xiaowei Zhuang and Ke Xu.]

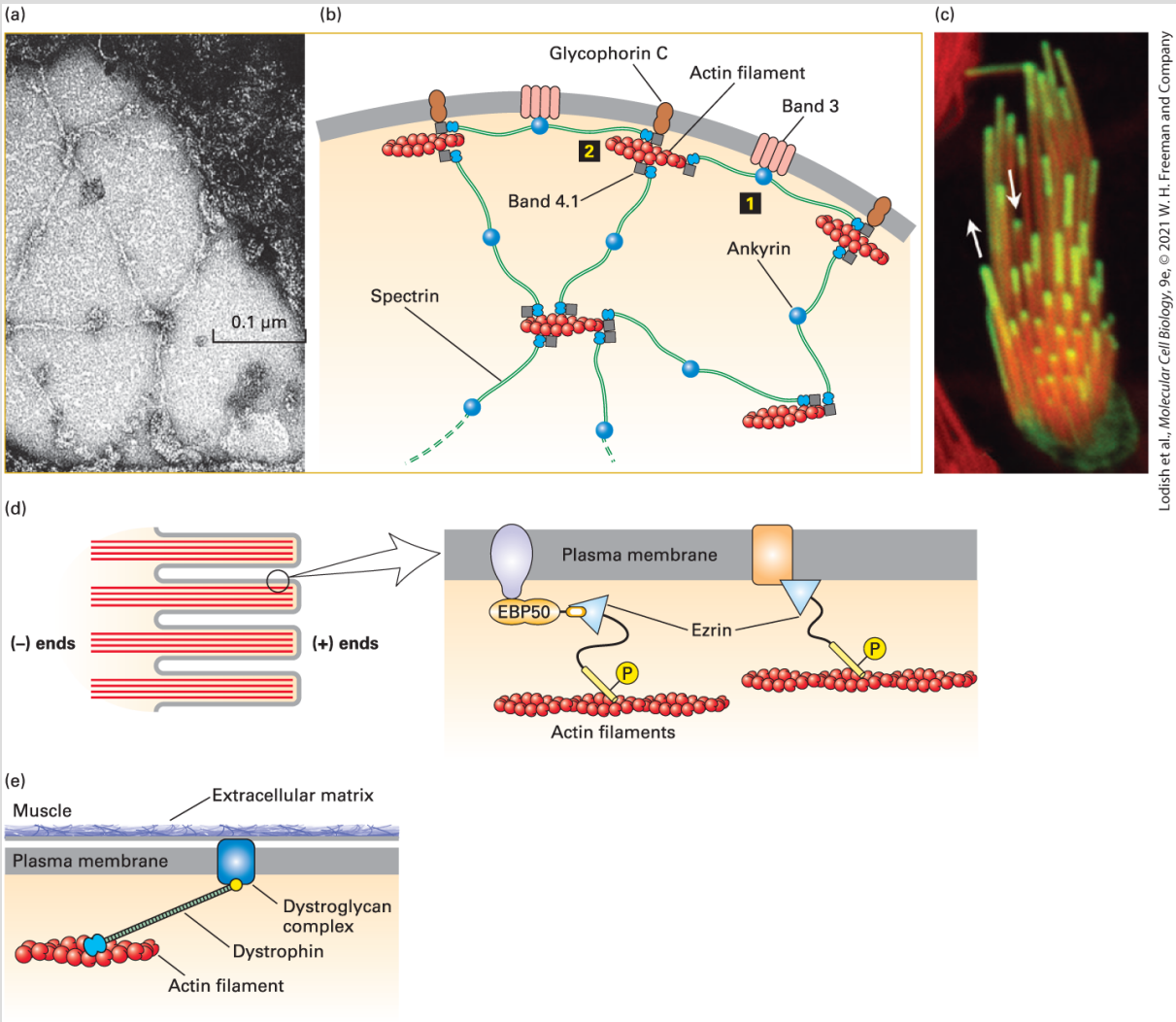
Other actin cross-linking proteins possess a single actin-binding site per polypeptide chain, but the functional protein is a dimer. These dimeric cross-linking proteins assemble to generate a rigid rod with a binding site at either end, as happens with *α-actinin*. Like fimbrin, α-actinin bundles parallel actin filaments, but keeps them farther apart than fimbrin. Another protein, called *spectrin*, is a tetramer with two polypeptide chains containing actin-binding sites and two chains that extend the length of the cross-linking protein. Spectrin spans a much greater distance between actin filaments than does fimbrin or α-actinin. Spectrin assists in forming microfilament networks beneath the plasma membrane (shown in Figure 17-20c and discussed in the next section). The Arp2/3 complex (which we discussed in terms of its ability to nucleate actin filament assembly) is also an important cross-linking protein, attaching the (−) end of one filament to the side of another filament (see Figure 17-15).

Adapter Proteins Link Actin Filaments to Membranes

Actin-based structures are not useful to a cell unless they are attached to specific components by *adapter proteins*. In this section, we describe how adapter proteins attach microfilaments to membranes to provide mechanical support, or attach a contractile apparatus to the membrane. While we discuss some specific examples, essentially all eukaryotic cells have actin-based filaments associated with the plasma membrane as part of a structure known as the *cell cortex*. In these various roles, actin filaments can interact with membranes either laterally or at their ends.

Our first example of actin filaments attached to a membrane involves the human erythrocyte — the red blood cell. The erythrocyte consists essentially of a plasma membrane enclosing a high concentration of the protein hemoglobin. It transports oxygen from the lungs to tissues and brings carbon dioxide from tissues back to the lungs. Erythrocytes must be able to survive the raging torrents of blood flow in the heart, then flow down arteries and survive squeezing through narrow capillaries before being cycled past the lungs and back to the heart. To survive this grueling journey for thousands of cycles, an erythrocyte has an actin-based network underlying the plasma membrane that gives the cell both the tensile strength and the flexibility necessary for its journey. This network is made of short actin filaments about 14 subunits in length, stabilized on their sides by tropomyosin (discussed in more detail in Section 17.6) and by the capping protein tropomodulin on the (−) end. These short filaments serve as hubs for binding about four to six

flexible spectrin molecules, generating a fishnet-like structure (Figure 17-21a) that provides strength and flexibility to the cell cortex. Spectrin molecules are attached to membrane proteins in two ways. One is via the protein *ankyrin*, which binds about halfway along a spectrin molecule and also binds to the transmembrane bicarbonate transporter (also known as *band 3*). The other adapter is *band 4.1*, which binds F-actin, the actin-binding domain of spectrin, and *glycophorin C*, another transmembrane protein (Figure 17-21b). Spectrin-based networks, similar to the one found in erythrocytes, occur in many types of cells. For example, an ankyrin-spectrin attachment links the Na^+/K^+ ATPase to the actin cytoskeleton on the basolateral membrane of epithelial cells, and an ordered repeating actin/spectrin network supports the axonal plasma membrane of neurons (Figure 17-20c).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 17-21 Lateral attachment of microfilaments to membranes. (a) Electron micrograph of the erythrocyte membrane showing the spoke-and-hub organization of the cortical cytoskeleton supporting the plasma membrane in human erythrocytes. The long spokes are composed mainly of spectrin and can be seen to intersect at the hubs, or membrane-attachment sites. The darker spots along the spokes are ankyrin molecules, which link spectrin to integral membrane proteins. (b) Diagram of the erythrocyte cytoskeleton, showing the two main types of membrane attachments: step **1** through ankyrin to band 3 and step **2** through band 4.1 to glycophorin C. (c) Actin is incorporated into the tips of stereocilia (giant microvilli). Cells with stereocilia were transfected to express GFP-labeled actin for a short period of time and then counterstained with rhodamine-phalloidin to stain all the F-actin. The experiment shows that new actin is incorporated at the tips of the stereocilia. (d) Ezrin, a member of the ezrin-radixin-moesin (ERM) family, links actin filaments laterally to the plasma membrane in surface structures such as microvilli;

attachment can be direct or indirect. Ezrin exists in an inactive, closed cytosolic form. Activation to an open form is mediated by phosphorylation and requires both the presence of the plasma membrane lipid PI(4,5)P2 and a specific ezrin kinase. Ezrin, activated by phosphorylation (P), links directly to the cytoplasmic region of transmembrane proteins (*right*) or indirectly through a scaffolding protein such as EBP50 (*left*). (e) Dystrophin, which has an actin-binding site on its N-terminal end and a C-terminal domain, binds the membrane protein dystroglycan, which itself associates with the extracellular matrix. See Figure 20-41 for more detail.

[Part (a) From T. J. Byers and D. Branton, 1985, *Proc. Nat'l. Acad Sci. USA* **82**:6153; courtesy Daniel Branton. Part (c) A. K. Rzadzinska et al., 2004, *J. Cell Biol.* **164**(6):887–897; https://doi.org/10.1083/jcb.200310055.]

Genetic defects in proteins of the red blood cell cytoskeleton can result in cells that rupture easily, giving rise to diseases known as *hereditary spherocytic anemias* (*spherocytic* because the cells are rounder, *anemias* because there is a shortage of red blood cells) and hence a shorter life span. In human patients, mutations in spectrin, band 4.1, and ankyrin can cause these diseases. ■

Microfilaments also support cell-surface structures such as microvilli and membrane ruffles. If we look at a microvillus, it is clear that its actin filaments must have an end-on attachment at the tip and lateral attachments down its length. What is the orientation of actin filaments in microvilli? Decoration experiments show that the (+) end is at the tip. Moreover, when fluorescent actin is added to a cell, it is incorporated at the tip of a microvillus, showing not only that the (+) end is there, but also that actin filament assembly occurs there (Figure 17-21c). It is not yet known how actin filaments

are attached at the microvillus tip or how assembly is regulated there. This (+) end orientation of actin filaments with respect to the plasma membrane is found almost universally — not just in microvilli, but also, for example, in the leading edges of motile cells. The lateral attachments to the plasma membrane are provided, at least in part, by the *ERM (ezrin-radixin-moesin) family* of proteins. This family consists of regulated proteins that exist in a folded, inactive form. They are activated after being phosphorylated by a specific kinase in a reaction requiring the plasma membrane-localized regulatory phospholipid phosphatidylinositol 4,5-bisphosphate [$PI(4,5)P_2$]. The F-actin–binding and membrane protein–binding sites of the ERM protein are exposed to provide a lateral linkage between membrane and actin filaments (Figure 17-21d). This type of coincident-detection mechanism — requiring both a regulatory lipid and localized kinase — ensures that ERM proteins are only activated at appropriate locations. ERM proteins can link actin filaments to the cytoplasmic domain of membrane proteins directly or indirectly through scaffolding proteins (see Figure 17-21d).

The types of actin-membrane linkages we have discussed so far do not involve areas of the plasma membrane attached directly to other cells in a tissue or to the extracellular matrix, but such linkages do exist. Contact between epithelial cells is mediated by highly specialized regions of the plasma membrane called *adherens junctions* (see Figure 17-1b). Other specialized regions of association

called *focal adhesions* mediate attachment of cells to the extracellular matrix. In turn, these specialized attachments connect to the cytoskeleton, as will be described in more detail when we discuss cell migration (Section 17.7) and cells in the context of tissues (see Chapter 20).

Muscular dystrophies are genetic diseases that are often characterized by the progressive weakening of skeletal muscle. One of these genetic diseases, Duchenne muscular dystrophy, affects the protein dystrophin, whose gene is located on the X chromosome, so the disease is much more prevalent in males than in females. *Dystrophin* is a modular protein whose function is to link the cortical actin network of muscle cells to a complex of membrane proteins that link to the extracellular matrix. Thus dystrophin has an N-terminal actin-binding domain, followed by a series of spectrin-like repeats, and terminates in a domain that binds the transmembrane dystroglycan complex to the extracellular matrix protein laminin (see Figure 17-21d, e and for more detail Figure 20-41). In the absence of dystrophin, the plasma membrane of muscle cells becomes weakened by cycles of muscle contraction and eventually ruptures, resulting in death of the muscle fibers. ■

KEY CONCEPTS OF SECTION 17.4

Organization of Actin-Based Cellular Structures

- Actin filaments are organized by cross-linking proteins that have two F-actin–binding sites. Actin cross-linking proteins can be long or short, rigid or flexible (see Figure 17-20).
- Actin filaments are attached to the plasma membrane by adapter proteins, as seen in erythrocytes and in cell-surface structures such as microvilli (see Figure 17-21).
- Attachment to the plasma membrane can be regulated, as seen in the activation of ERM proteins by phosphorylation.
- Hereditary diseases, such as spherocytic anemias and muscular dystrophy, have been traced to defects in the attachment of the microfilament-based cortical cytoskeleton to the plasma membrane.

17.5 Myosins: Actin-Based Motor Proteins

In Section 17.3, we discussed how actin polymerization nucleated by the Arp2/3 complex can be harnessed to do work, such as moving vesicles during endocytosis, moving cells along a substrate at their leading edges, and propelling a *Listeria* bacterium through the cytoplasm of a host cell. In addition to this actin polymerization–based motility, cells have a large family of motor proteins called **myosins** that can perform work by moving along actin filaments.

The first myosin discovered, *myosin II*, was isolated from skeletal muscle. For a long time, biologists thought that this was the only type of myosin found in nature. However, they then discovered other types of myosins and began to ask how many functional classes might exist. Today we know that there are multiple families of myosins, in addition to the myosin II of skeletal muscle, that move along actin. Indeed, with the discovery and analysis of all these microfilament-based motors and the corresponding microtubule-based motors described in the next chapter, our former relatively static view of cells has been replaced with the realization that the cytoplasm is incredibly dynamic — like a busy freeway system with motors busily ferrying components around.

Myosins have the amazing ability to convert the energy released by ATP hydrolysis into mechanical work (movement along actin). While

all myosins convert the chemical energy of ATP hydrolysis into mechanical work, different myosins perform very different cellular functions. For example, many molecules of myosin II pull together on actin filaments to bring about muscle contraction, whereas myosin V binds to cargo-filled vesicles and transports them along actin filaments. Other types of myosin move organelles around cells or aid in cell motility.

To begin to understand myosins, we first discuss their general structure. Next we explore the diversity of myosin classes and describe in more detail some of those that are common in eukaryotes. We investigate the mechanism that converts the energy released by ATP hydrolysis into work and then see how this mechanism is modified to tailor the properties of specific myosin classes to their specific functions.

Myosins Have Head, Neck, and Tail Domains with Distinct Functions

Much of what we know about myosins comes from studies of myosin II isolated from skeletal muscle. In skeletal muscle, hundreds of individual myosin II molecules are assembled into bipolar bundles called *thick filaments* (Figure 17-22a). In a later section, we will discuss how these myosin filaments interdigitate with actin

filaments to bring about muscle contraction. Here we investigate the properties of the myosin molecule itself.

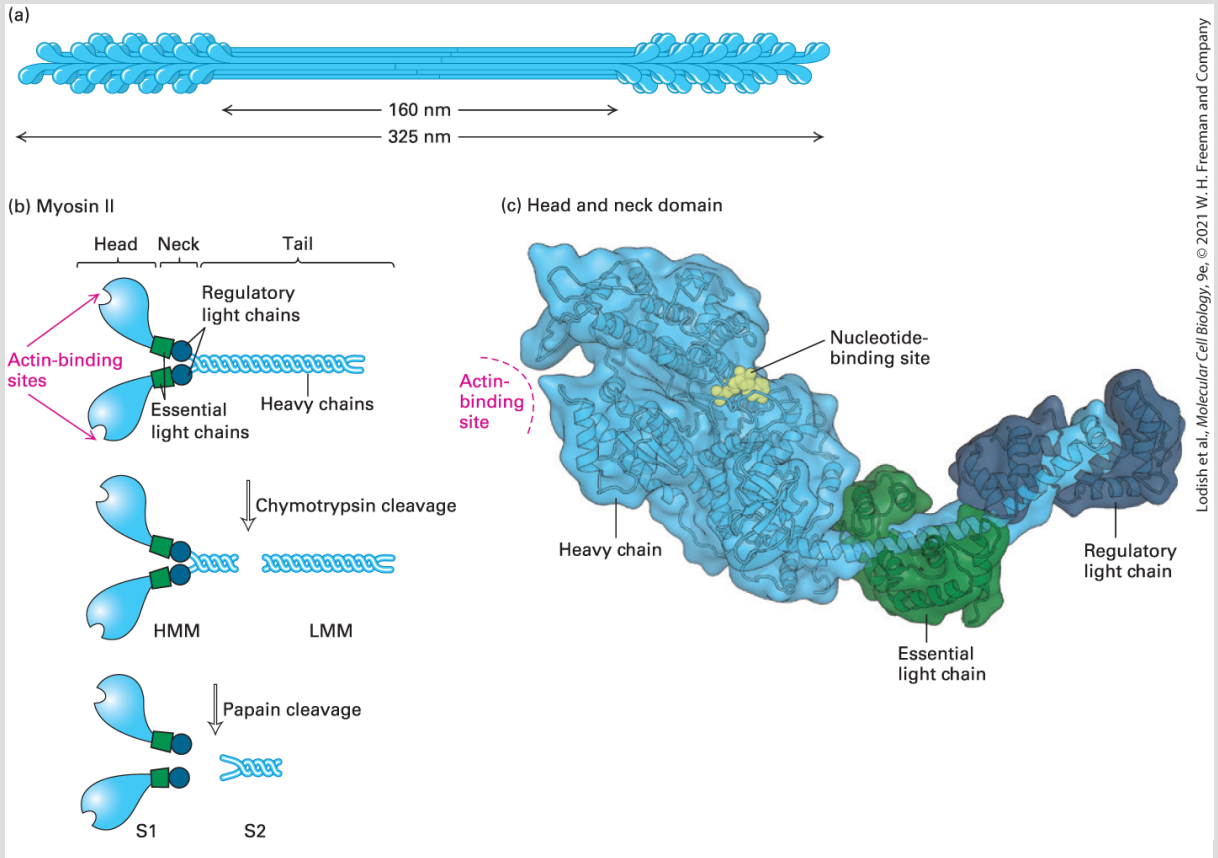


FIGURE 17-22 Structure of myosin II. (a) Organization of myosin II in filaments isolated from skeletal muscle. Myosin II assembles into bipolar filaments in which the tails form the shaft of the filament and the heads are exposed at the ends. Treatment of bipolar filaments with high salt concentrations and ATP disassembles the filament into individual myosin II molecules. (b) A myosin II molecule consists of two identical heavy chains (light blue) and four light chains (green and dark blue). The tails of the heavy chains form a coiled coil to dimerize; the neck region of each heavy chain has two light chains associated with it. Limited proteolytic cleavage of myosin II generates tail fragments — LMM and S2 — and the S1 motor domain. (c) Three-dimensional model of a single S1 head domain shows that it has a curved, elongated shape and is bisected by a cleft. The nucleotide-binding pocket lies on one side of this cleft, and the actin-binding site lies on the other side near the tip of the head. Wrapped around the shaft of the α-helical neck are two light chains. These chains

stiffen the neck so that it can act as a lever arm for the head. Shown here is the ADP-bound conformation.

[Part (c) Data from S. Gourinath et al., 2003, *Structure* **11**:1621–1627, PDB ID 1qvi.]

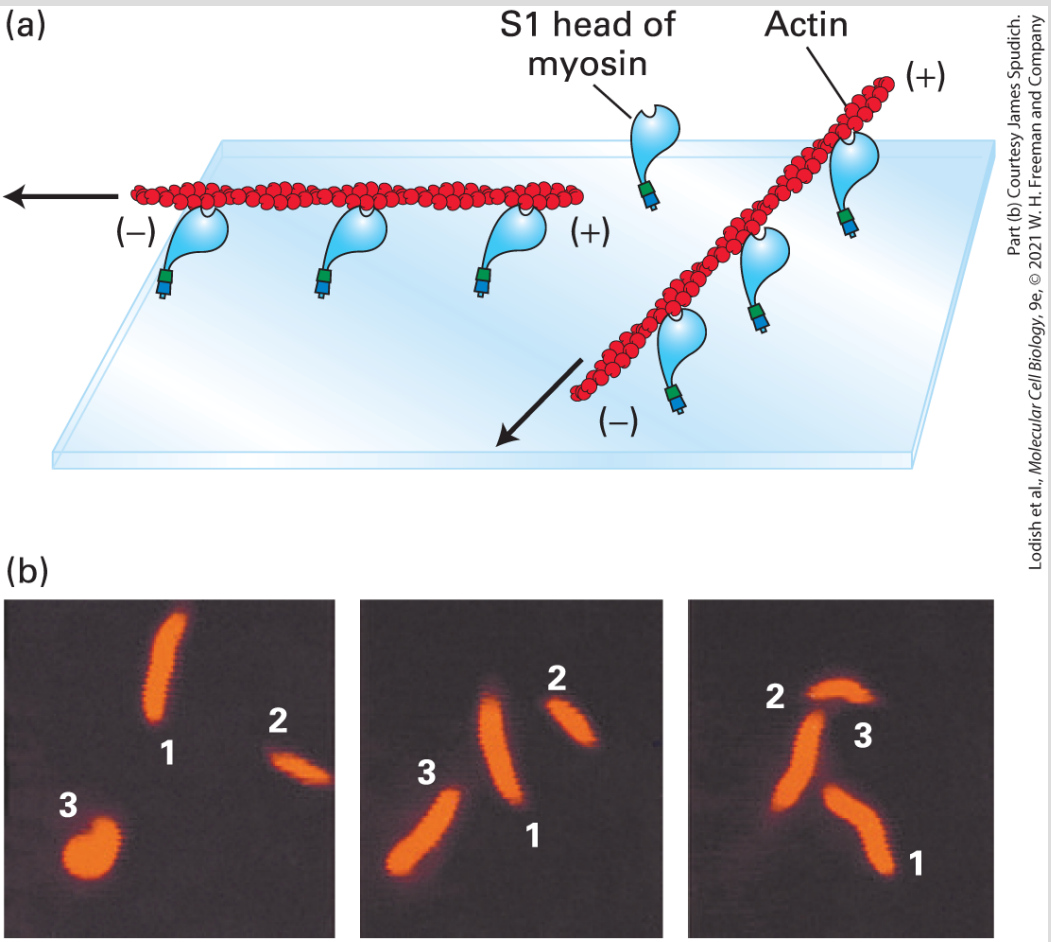
Dissolving the myosin thick filament in a solution of ATP and a high concentration of salt generates a pool of individual myosin II molecules. The soluble myosin II molecule is actually a protein complex consisting of six polypeptide subunits. Two of the subunits are identical high-molecular-weight polypeptides known as *myosin heavy chains*. Each consists of a globular *head* domain and a long *tail* domain, connected by a flexible *neck* domain. The tails of the two myosin heavy chains intertwine, so that the head regions are in close proximity. The remaining four subunits of the myosin complex are smaller in size and are known as the *light chains*. There are two types of light chains: the *essential light chain* and the *regulatory light chain*. One light chain of each type associates with the neck region of each heavy chain (Figure 17-22b, *top*). The myosin heavy chain and the two types of light chains are encoded by three different genes.

The soluble myosin II molecule has ATPase activity, reflecting its ability to power movements by hydrolysis of ATP, but which part of the myosin complex is responsible for this activity? A standard approach to identifying functional domains in a protein is to cleave the protein into fragments with specific proteases and then ask which fragments have the activity. Soluble myosin II can be cleaved

by gentle treatment with the protease chymotrypsin to yield two fragments, one called *heavy meromyosin* (*HMM*; *mero* means "part of") and the other, *light meromyosin (LMM)* (Figure 17-22b, *middle*). The heavy meromyosin can be further cleaved by the protease papain to yield subfragment 1 (S1) and subfragment 2 (S2) (Figure 17-22b, *bottom*). By analyzing the properties of the various fragments — S1, S2, and LMM — it was found that the intrinsic ATPase activity of myosin resides in the S1 fragment, as does its F-actin–binding site. Moreover, it was found that the ATPase activity of the S1 fragment is greatly enhanced by the presence of filamentous actin, so that fragment is said to have an *actin-activated ATPase activity*, which is a hallmark of all myosins. The S1 fragment of myosin II consists of the head and neck domains with associated light chains, whereas the S2 and LMM regions make up the tail domain.

X-ray crystallographic analysis of the head and neck domains has revealed the shapes of the subunits, the positions of the light chains, and the locations of the ATP-binding and actin-binding sites (Figure 17-22c). At the base of the myosin head is the α-helical neck, where two light-chain molecules wrap around the heavy chain like C-clamps. In this position, the light chains stiffen the neck region. The actin-binding site is an exposed region at the tip of the head domain. The ATP-binding site is also in the head domain, within a cleft opposite the actin-binding site.

How much of the myosin II molecule is necessary and sufficient for its motor activity? To answer this question, one needs a simple in vitro motility assay. In one such assay, the *sliding-filament assay*, myosin molecules are tethered to a coverslip to which is added stabilized, fluorescently labeled actin filaments and ATP. Because the myosin molecules are tethered, they cannot slide; thus any force generated by the interaction of myosin heads with actin filaments forces the filaments to move relative to the myosin. The tethered myosin molecules or fragments pull toward the (+) end of a filament, so the force that myosin generates causes the filaments to move with the (−) end leading (Figure 17-23). Using the sliding-filament assay, one can show that the S1 fragment of myosin II is sufficient to move actin filaments (Figure 17-23a). The rate at which myosin moves an actin filament can be determined from video recordings of sliding-filament assays (Figure 17-23b).



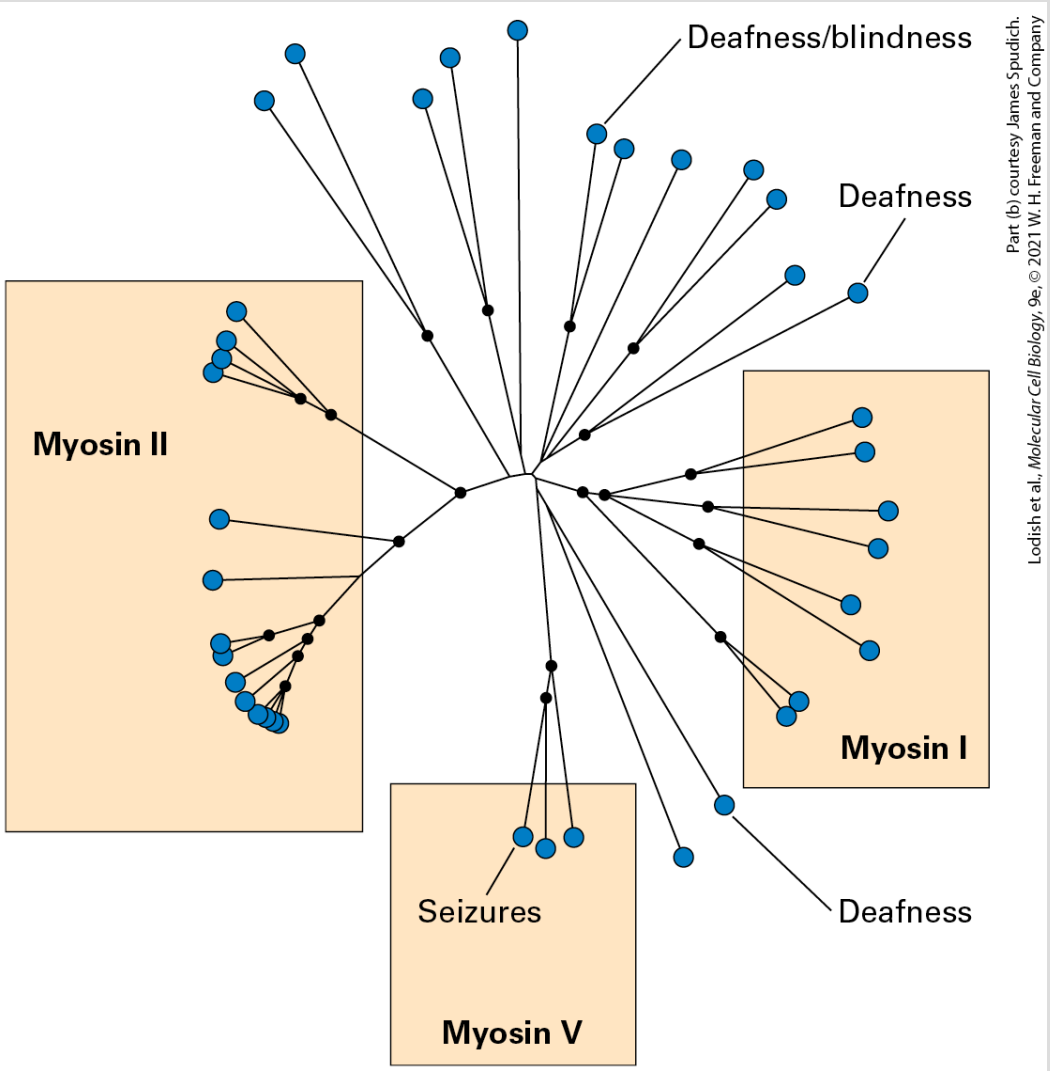
Part (b) Courtesy James Spudich.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 17-23 The sliding-filament assay is used to detect myosin-powered movement. (a) After myosin S1 fragments are adsorbed onto the surface of a glass coverslip, excess unbound myosin is removed; the coverslip is then placed myosin-side-down on a glass slide to form a chamber through which solutions can flow. A solution of actin filaments, made visible and stable by staining with rhodamine-labeled phalloidin, is allowed to flow into the chamber. In the presence of ATP, the myosin heads pulls toward the (+) ends of the actin filaments by the mechanism discussed later and illustrated in Figure 17-26. Because the myosin tails are immobilized, movement of the heads toward the (+) ends causes sliding of the filaments, which appear to be moving with their (–) ends leading the way. (b) Movement of individual filaments can be observed in a fluorescence light microscope. These photographs show the positions of three actin filaments (numbered 1, 2, 3) at 30-second intervals recorded by video microscopy. The rate of filament movement can be determined from such recordings.

All myosins have a domain related to the S1 domain of myosin II, which in all cases is responsible for the motor activity. As we will see in a later section, the length of the neck domain and the numbers and types of light chains associated with it vary among myosin classes. The tail domain does not contribute to motility, but rather defines what is moved by the S1-related domain. Thus as might be expected, tail domains differ among myosins and are tailored to bind specific cargoes.

Myosins Make Up a Large Family of Mechanochemical Motor Proteins

Myosin S1 domains share considerable similarity in their primary amino acid sequence, so it is possible to determine how many myosin genes and how many different classes of myosins exist in the sequenced genome of an organism. For example, there are about 40 myosin genes in the human genome (Figure 17-24), 9 in *Drosophila*, and 5 in budding yeast. Bioinformatic analysis of the sequence relationships among myosin head domains indicates that about 20 distinct classes of myosins have evolved in eukaryotes. As shown in Figure 17-24, the genetic bases for some conditions have been traced to mutations in genes encoding myosins.



Part (b) courtesy James Spudich.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 17-24 The myosin superfamily in humans. Results of a computer analysis of the relatedness of the S1 head domains of all of the approximately 40 myosins encoded by the human genome. Each myosin is indicated by a blue dot. The lengths of the black lines indicate phylogenetic distance relationships: myosins connected by short lines are closely related, whereas those separated by longer lines are more distantly related. Among these myosins, three classes — myosins I, II, and V — are widely represented among eukaryotes; others have more specialized functions. Indicated are examples in which loss of a specific myosin causes a disease.

[Data from R. E. Cheney, 2001, *Mol. Biol. Cell* **12**:780.]

All myosin head domains use the same mechanism to convert ATP hydrolysis into mechanical work. However, as we will see, subtle variations in this mechanism can have profound effects on the functional properties of different myosin classes. How do these different classes relate with respect to their tail domains? If one uses the protein sequences of myosin tail domains to place the myosins into classes, the myosins fall into the same groupings found when comparing the motor domains. This implies that head domains with specific properties have coevolved with specific classes of tail domains, which makes a lot of sense, and further suggests that each class of myosin has evolved to carry out a specific function.

Among the myosins are three especially well-studied families, commonly found in animals and fungi: the *myosin I*, *myosin II*, and *myosin V* families (Figure 17-25). In humans, 8 genes encode heavy chains for the myosin I family, 14 for the myosin II family, and 3 for the myosin V family (see Figure 17-24).

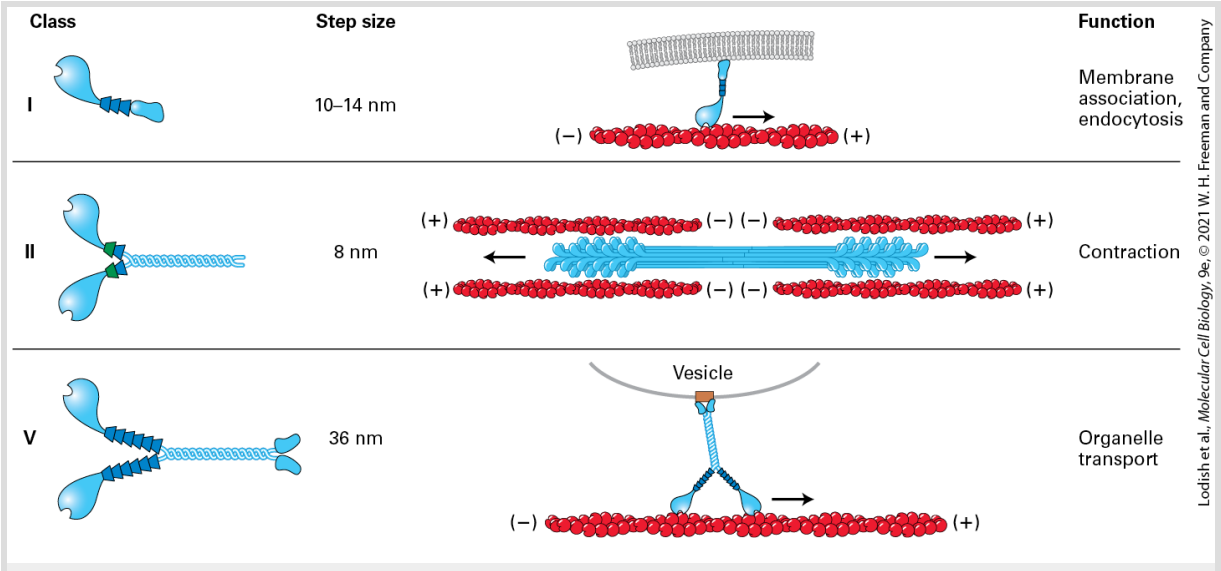


FIGURE 17-25 Three common classes of myosins. Myosin I molecules consist of a head domain and a neck domain; a variable number of light chains is associated with the neck domain. Members of the myosin I class are the only myosins to have a single head domain. Some of these myosins are believed to associate directly with membranes through lipid interactions. Myosin II molecules have two head domains and two light chains per neck and are the only class that can assemble into bipolar filaments. Myosin V molecules have two head domains and six light chains per neck. They bind specific receptors (brown box) on organelles, which they transport. All myosins in these three classes move toward the (+) end of actin filaments.

The large myosin I family has a variable number of light chains associated with the neck region and is the only type of myosin that has a single heavy chain and so functions with a single head. Only a few of the functions of myosin I molecules have been determined. These include connecting actin filaments to membranes and assisting in endocytosis. Since these myosin molecules have only one head each, several of them must work together to generate movement, and at least one must always be attached to an actin filament during the work cycle.

All myosin II members have a relatively short neck domain and have two light chains per heavy chain. The myosin II class is the only one that assembles into bipolar filaments. Myosin II molecules form bipolar filaments in which the tail domains assemble together into thick bundles with a cluster of head domains at each end. This organization is important for its involvement in contraction; indeed, this is the only class of myosins involved in contractile functions. The large number of members in this class reflects the need for

myosin II filaments with slightly different contractile properties seen in different muscles (e.g., skeletal, cardiac, and various types of smooth muscle) as well as in nonmuscle cells.

Members of the myosin V class have two heavy chains, resulting in a motor with two heads, long neck regions with six light chains each, and tail regions that dimerize and terminate in domains that bind to specific organelles to be transported.

In every case that has been tested so far, myosins move toward the (+) end of an actin filament — with one exception: *myosin VI* found in animals. This unusual myosin has an inserted amino acid sequence in its head domain that causes myosin and its cargo to move toward the (−) end of the actin filament. Myosin VI is believed to move endocytic vesicles along actin filaments away from the plasma membrane. Recall that membrane-associated actin filaments have their (+) ends toward the membrane, so a motor directed toward the (−) end would take carry vesicles away from the membrane toward the center of the cell.

Conformational Changes in the Myosin Head Couple ATP Hydrolysis to Movement

Studies of muscle contraction provided the first evidence that myosin heads slide or walk along actin filaments. Unraveling the mechanism of muscle contraction was greatly aided by the development of in vitro motility assays and single-molecule force measurements. On the basis of information obtained with these two techniques plus the three-dimensional structure of the myosin head (see Figure 17-22c), researchers have developed a model for how myosin harnesses the energy released by ATP hydrolysis to move along an actin filament (Figure 17-26). Because all myosins are thought to use the same basic mechanism to generate movement, we will ignore for the moment whether the myosin tail is bound to a vesicle or is part of a thick filament, as it is in muscle. The most important aspect of this model is that the hydrolysis of a single ATP molecule is coupled to each step taken by a myosin molecule along an actin filament.

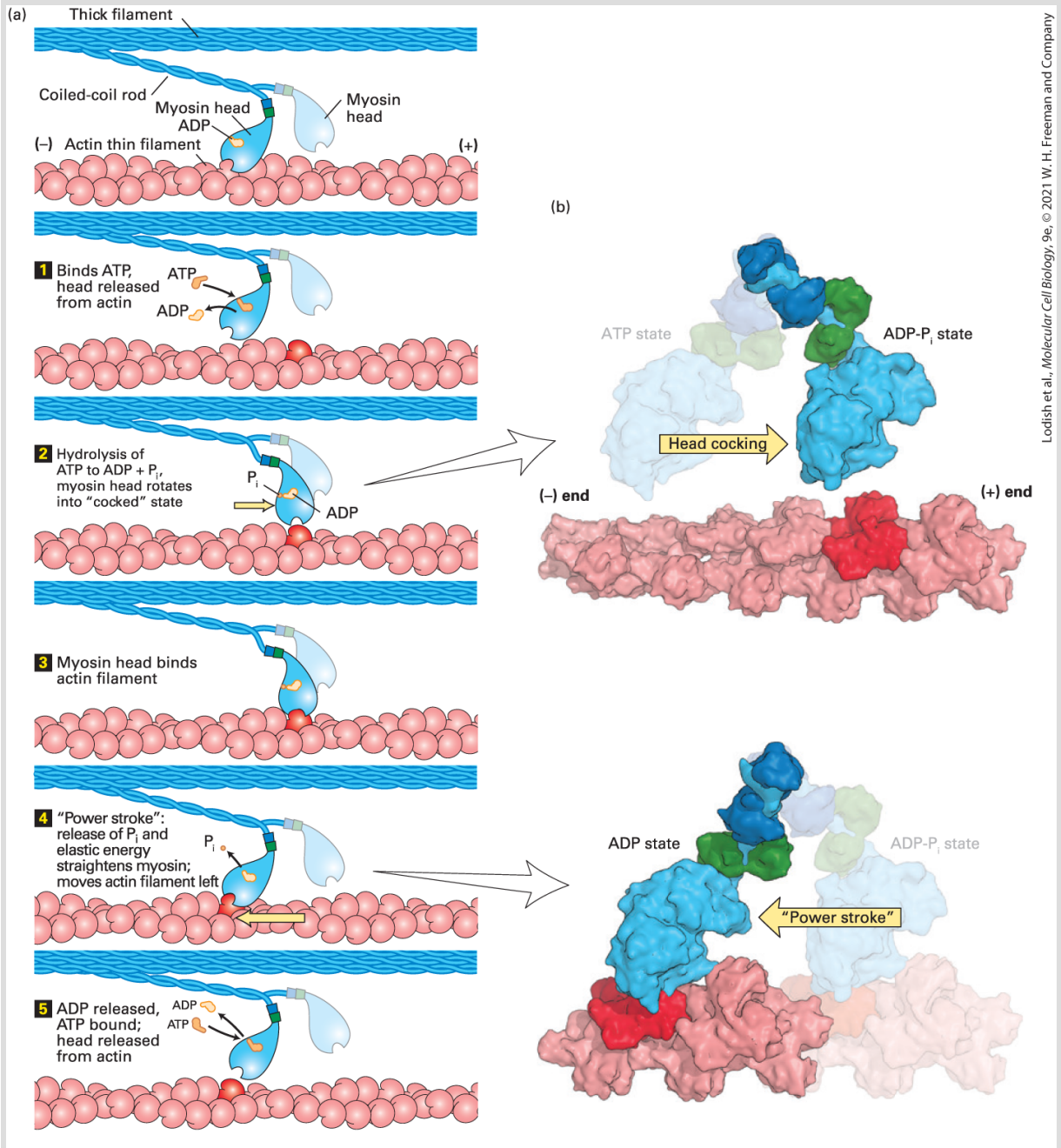


FIGURE 17-26 ATP-driven myosin movement along actin filaments. (a) When ADP is bound to the myosin head, it is firmly attached to the actin filament. Step **1**: On exchanging ATP for ADP, the myosin head releases from the actin filament. Step **2**: The head hydrolyzes the ATP to ADP and P_i, which induces a rotation in the head with respect to the neck. This "cocked" state stores the energy released by ATP hydrolysis as elastic energy, like a stretched spring. Step **3**: Myosin in the cocked state is stable until it binds actin. Step **4**: When it encounters actin, the myosin head couples release of P_i with release of the elastic energy to move the actin filament. This movement is known as the *power stroke*, as it

involves moving the actin filament with respect to the end of the myosin neck domain. In the presence of bound ADP, the head remains tightly bound to the actin filament in a state known as *rigor* and is the basis for the unmistakable sign of death, rigor mortis. Step **5**: When ADP is replaced by fresh ATP the head releases from the filament. See R. D. Vale and R. A. Milligan, 2002, *Science* **288**:88. (b) Molecular models of the conformational changes in the myosin head involved in "cocking" the head (*upper panel*) and after the power stroke (*lower panel*). The myosin light chains are shown in dark blue and green; the rest of the myosin head and neck are colored light blue, and actin is red. See S. Fischer et al., 2005, *Proc. Nat'l Acad. Sci. USA* **102**:6873–6876.

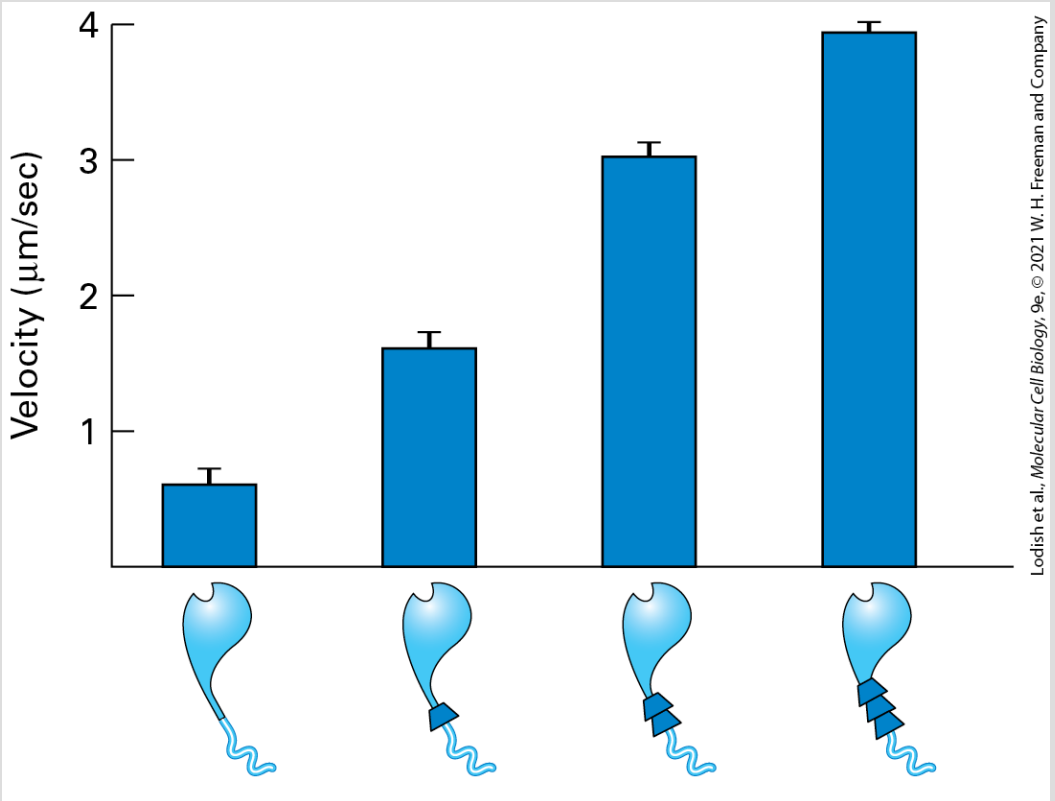
[Part (b) Data from Ken Holmes.]

How does myosin convert the chemical energy released by ATP hydrolysis into mechanical work? The S1 head of myosin is an ATPase. When it has ADP bound, the head of myosin binds very tightly to F-actin (Figure 17-26a). When the ADP in the head is exchanged for ATP, its affinity for F-actin is greatly reduced, and it releases from actin (Figure 17-26a, step **1**). Next the myosin head hydrolyzes the ATP, and the hydrolysis products, ADP and P_i, remain bound to it (step **2**). The energy from the hydrolysis of ATP induces a conformational change in the head that results in the head domain rotating with respect to the neck domain, assuming what is known as the "cocked" position (Figure 17-26b, *top*). Next the head binds F-actin tightly (step **3**), inducing both release of P_i and rotation of the head back to its original position, thus moving the actin filament relative to the neck domain (Figure 17-26a, step **4**, and 17-26b, *bottom*). In this way, binding to F-actin induces the movement of the head and the release of P_i, thereby coupling the two processes. This step is known as the **power stroke**. The head remains bound until

the ADP leaves and a fresh ATP binds the head, releasing it from the actin filament (step **5**). The cycle repeats.

How is hydrolysis of ATP in the nucleotide-binding pocket of the head converted into a mechanical force? Hydrolysis of ATP induces a small conformational change in the head domain. This small movement is amplified by a converter region at the base of the head, which acts like a fulcrum causing the rodlike neck (also known as the *lever arm*) to rotate. The rotation is amplified by the neck domain, so the actin filament moves by a few nanometers (see Figure 17-26b).

This model makes a strong prediction: the distance a myosin moves along actin during hydrolysis of one ATP — the myosin **step size** — should be proportional to the length of the neck domain. To test this prediction, mutant myosin molecules were constructed with neck domains of different lengths, and the rate they moved down an actin filament was determined. As the myosins with different neck domains step at the same rate, the rate of movement reflects the step size. As predicted, there was an excellent correlation between the length of the neck domain and the rate of movement (Figure 17-27).



EXPERIMENTAL FIGURE 17-27 The length of the myosin II neck domain determines the rate of movement. To test the lever-arm model of myosin movement, investigators used recombinant DNA techniques to make myosin heads attached to neck domains with increasing numbers of light-chain-binding sites and hence lengths. The rate at which these myosins moved on actin filaments was determined, which reflects the myosin step size. The longer the lever arm, the faster the myosin moved, supporting the proposed mechanism.

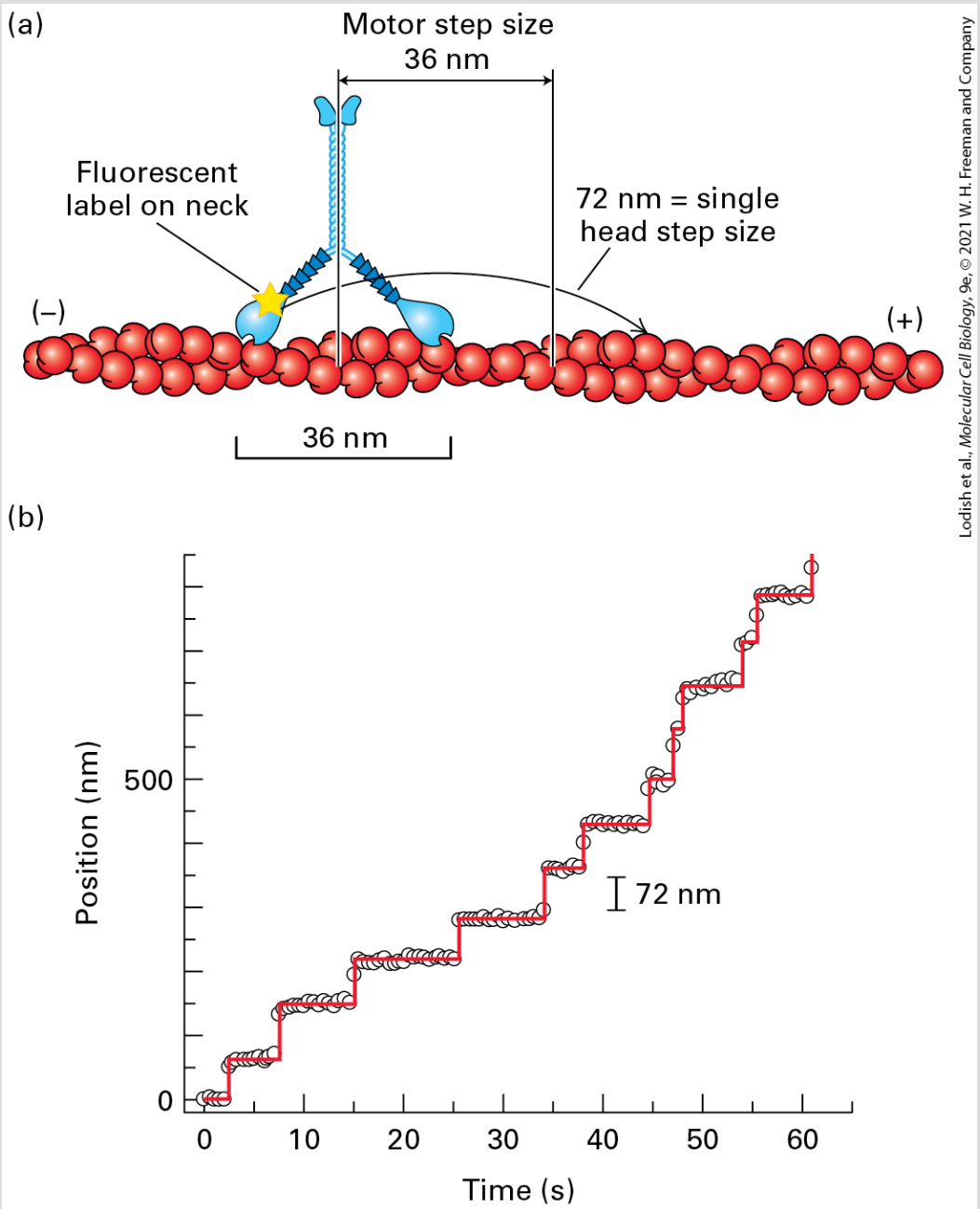
[Data from K. A. Ruppel and J. A. Spudich, 1996, *Ann. Rev. Cell Dev. Biol.* **12**:543–573.]

Myosin Heads Take Discrete Steps Along Actin Filaments

The defining property of a myosin is its ability to generate a force when it contacts an actin filament, but the details of this interaction are tailored to the function of the myosin. Researchers have measured the distance a single myosin II head moves a filament and the force generated. The results show that myosin II takes discrete steps, which average out to about 8 nm, and generates 3–5 piconewtons (pN) of force, approximately the same force as that exerted by gravity on a single bacterium. Moreover, myosin II does not interact with the actin filament continuously but rather binds, moves, and releases it. In fact, myosin II spends on average only about 10 percent of each ATPase cycle in contact with F-actin — it is said to have a **duty ratio** of 10 percent. This observation will be important later when we consider that in contracting muscle, hundreds of myosin heads pull on actin filaments, so that at any one time, 10 percent of the heads are engaged to provide a smooth contraction.

Now let us examine how myosin V moves. Recall that myosin V has a much longer neck region than does myosin II (see Figure 17-25). In one set of experiments, scientists added a fluorescent label to one of the two neck domains in myosin V molecules (see Section 4.2 for how this can be achieved). Mixing labeled myosin V molecules with F-actin and ATP, the movement of myosin V was observed using fluorescence microscopy (Figure 17-28a). The labeled head took many 72-nm steps without releasing from the actin — so it is said to move *processively* (Figure 17-28b). In a filament, actin subunits are

oriented to present a myosin-binding site every 36 nm (see Figures 17-5b and 17-28a), so each head binds to every other site along a filament. As the myosin has two heads, each alternately takes 72-nm steps — like someone walking on stepping stones across a river and placing each foot on every other stone. When the label was attached to the tail instead of the neck, the myosin V motor as a whole moved in 36-nm steps. Thus while the step size for each individual head is 72 nm, cargo attached to the myosin V tail region moves in 36-nm steps as a head moves from behind the cargo to in front of it (Figure 17-28a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 17-28 Myosin V has a step size of 36 nm, with each head stepping hand-over-hand in 72-nm steps. (a) Diagram of how the myosin V heads interact with binding sites spaced 36 nm along an actin filament, with a fluorescent label on the neck domains of one myosin head. While each myosin V head steps "hand-over-hand" down an actin filament in 72-nm steps, the cargo, and hence the motor as a whole, moves in 36-nm steps. (b) A trace of the label on one myosin V neck domain as the molecule walks down an actin

filament. The labeled myosin head takes successive 72-nm steps. As can be seen from the trace, myosin V takes many successive steps along a filament, so it is said to be processive. As shown in panel (a), the step size corresponds to equivalent sites on the helical structure of the actin filament.

[Data from A. Yildiz et al., 2003, *Science* **300**:2061.]

Presumably, myosin V has evolved a long neck domain to take steps that match the distance between myosin-binding sites. Moreover, its ATPase cycle has a high duty ratio of >70 percent. (Compare this with the 10 percent duty ratio of myosin II.) Myosin V has a high duty ratio because its rate of ADP release is quite slow; thus the head remains in contact with the actin filament for a much larger percentage of the cycle. Since a single myosin V molecule has two heads, a duty ratio greater than 50 percent ensures that at least one head is always in contact with actin as it moves down an actin filament. This keeps myosin V and its cargo from falling off. These are exactly the properties one would expect for a motor designed to transport cargo along an actin filament.

KEY CONCEPTS OF SECTION 17.5

Myosins: Actin-Based Motor Proteins

- Myosins are actin-based motors powered by ATP hydrolysis.
- Myosins have a motor head domain, a lever-arm neck domain, and a cargo-binding tail domain (see Figure 17-22).
- There are many classes of myosins, three of which are present in many eukaryotes: myosin I, which has a single head domain; myosin II, which has two heads and assembles into bipolar filaments; and myosin V, which has two heads but does not assemble into filaments (see Figure 17-25).

- Myosins convert ATP hydrolysis to mechanical work by amplifying a small conformational change in the head through the neck domain when the head is bound to F-actin (see Figure 17-26).
- Myosin heads take discrete steps along an actin filament, which can be small (8 nm) and nonprocessive in the case of myosin II or large (36 nm for the motor overall, 72 nm for each head) and processive for myosin V.

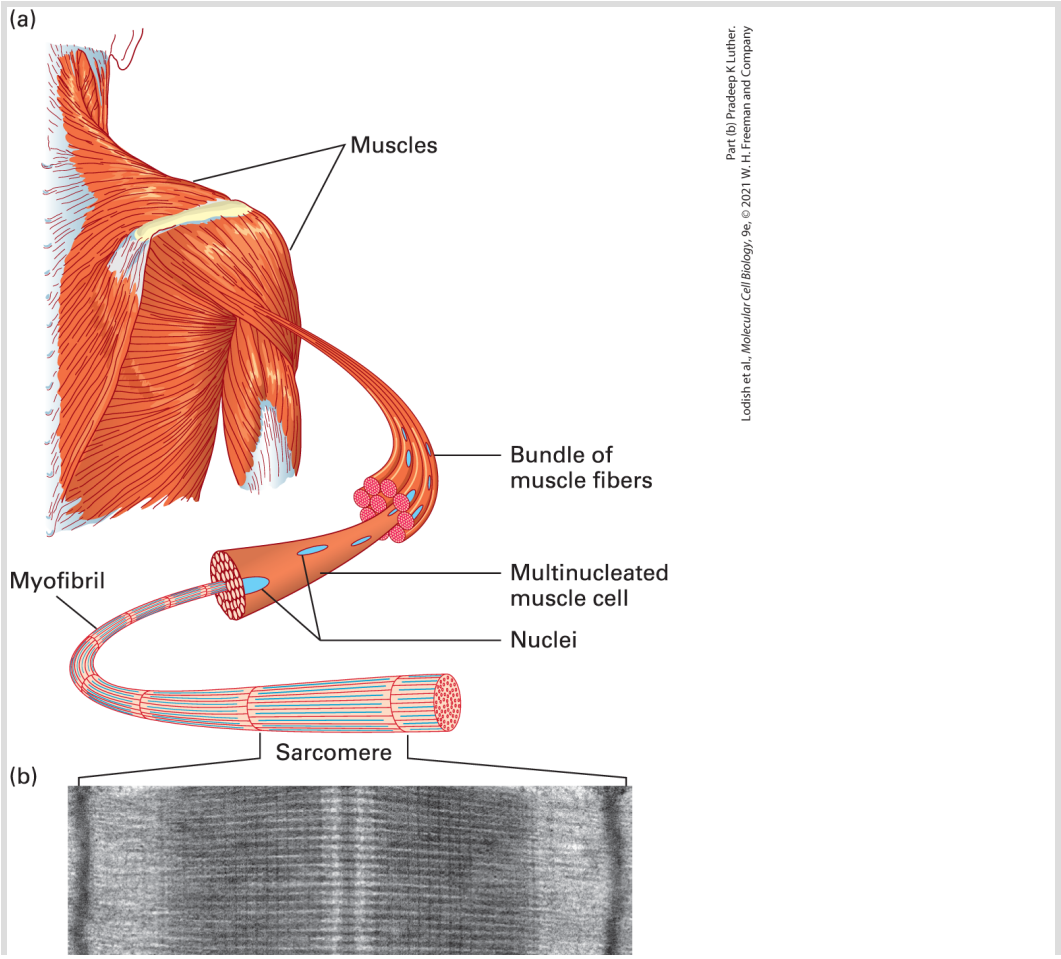
17.6 Myosin-Powered Movements

In the previous section, we discussed how the head and neck domains of myosins function as a motor that converts chemical energy into mechanical force. In this section, we discuss how the whole myosin motor molecules work and are regulated. Since the functions of many of the families of myosins found in metazoans are not yet known, we look at three examples for which we have a good idea of specific myosin function. Our first example is skeletal muscle. In muscle, many myosin II heads, each with a low duty ratio, are bundled into bipolar filaments that work together to bring about muscle contraction. Contractile machinery also functions in the contraction of smooth muscle, in stress fibers, and in the contractile ring during cytokinesis. In our second example, we will look at how contraction operates in nonskeletal muscle. In the third example, we turn to the myosin V class, whose high duty ratio allows cargo to be transported over relatively long distances.

Myosin Thick Filaments and Actin Thin Filaments in Skeletal Muscle Slide Past Each Other During Contraction

Muscle cells have evolved to carry out one highly specialized function: contraction. Muscle contractions must occur quickly and repetitively, and they must occur over long distances and with

enough force to move large loads. A typical skeletal muscle cell is cylindrical, large (1–40 mm in length and 10–50 μm in width), and multinucleated (containing as many as 100 nuclei) (Figure 17-29a). Within each muscle cell are many **myofibrils** consisting of a repeating array of a specialized structure called a **sarcomere** (Figure 17-29b). A sarcomere, which is about 2 μm long in resting muscle, shortens by about 70 percent of its length during contraction. Electron microscopy and biochemical analysis have shown that each sarcomere contains two major types of filaments: **thick filaments**, composed of myosin II, and **thin filaments**, containing actin and associated proteins (Figure 17-29c).



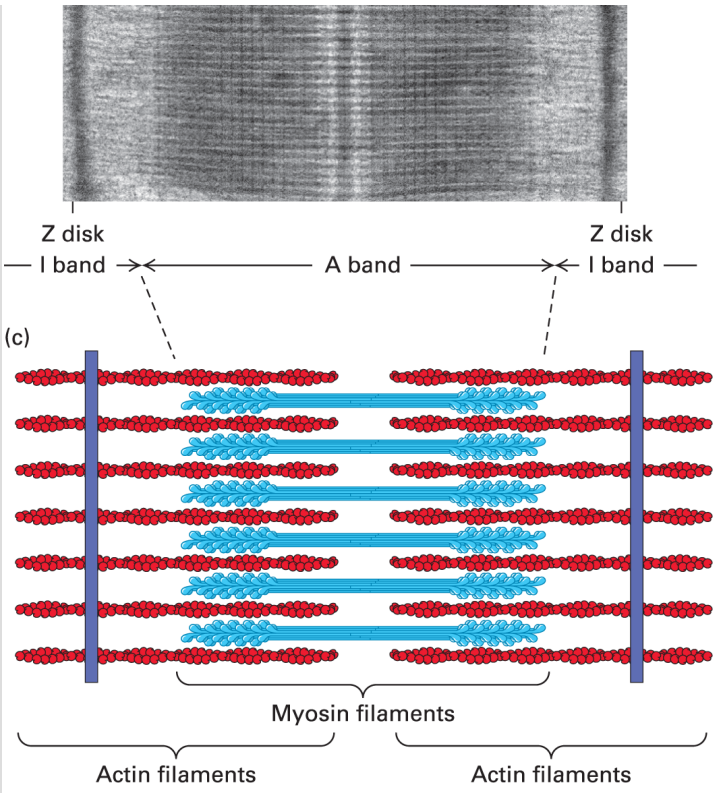


FIGURE 17-29 Structure of the skeletal muscle sarcomere. (a) Skeletal muscles consist of muscle fibers made of bundles of multinucleated cells. Each cell contains a bundle of myofibrils, which consist of thousands of repeating contractile structures called *sarcomeres*. (b) Electron micrograph of mouse skeletal muscle in longitudinal section, showing one sarcomere. On either side of the Z disks are the lightly stained I bands, composed entirely of actin thin filaments. These thin filaments extend from both sides of the Z disk to interdigitate with the dark-stained myosin thick filaments in the A band. (c) Diagram of the arrangement of myosin and actin filaments in a sarcomere.

The thick filaments are made of many myosin II bipolar filaments, in which the heads on each half of the filament are in opposite orientations (see Figure 17-22a). The thin actin filaments are assembled with their (+) ends embedded in a densely staining structure known as the *Z disk*, so that the two sets of actin filaments

in a sarcomere have opposite orientations (Figure 17-30). To understand how a muscle contracts, consider the interactions between one myosin head (among the hundreds in a thick filament) and one thin (actin) filament, as diagrammed in Figure 17-26. During these cyclical interactions, also called the *cross-bridge cycle*, the hydrolysis of ATP is coupled to the movement of a myosin head toward the Z disk, which corresponds to the (+) end of the actin thin filament. Because the thick filament is bipolar, the action of the myosin heads at opposite ends of the thick filament draws the thin filaments toward the center of the thick filament and therefore toward the center of the sarcomere (see Figure 17-30). This movement shortens the sarcomere until the ends of the thick filaments abut the Z disk. Contraction of an intact muscle results from the activity of hundreds of myosin heads on a single thick filament, amplified by the hundreds of thick and thin filaments in a sarcomere and thousands of sarcomeres in a muscle fiber. We can now see why myosin II is both nonprocessive and needs to have a low duty ratio: each head pulls a short distance on the actin filament and then lets go to allow other heads to pull, and so many heads working together allow the smooth contraction of the sarcomere. The first experimental basis for the sliding-filament model of muscle contraction is highlighted in Classic Experiment 17-1.

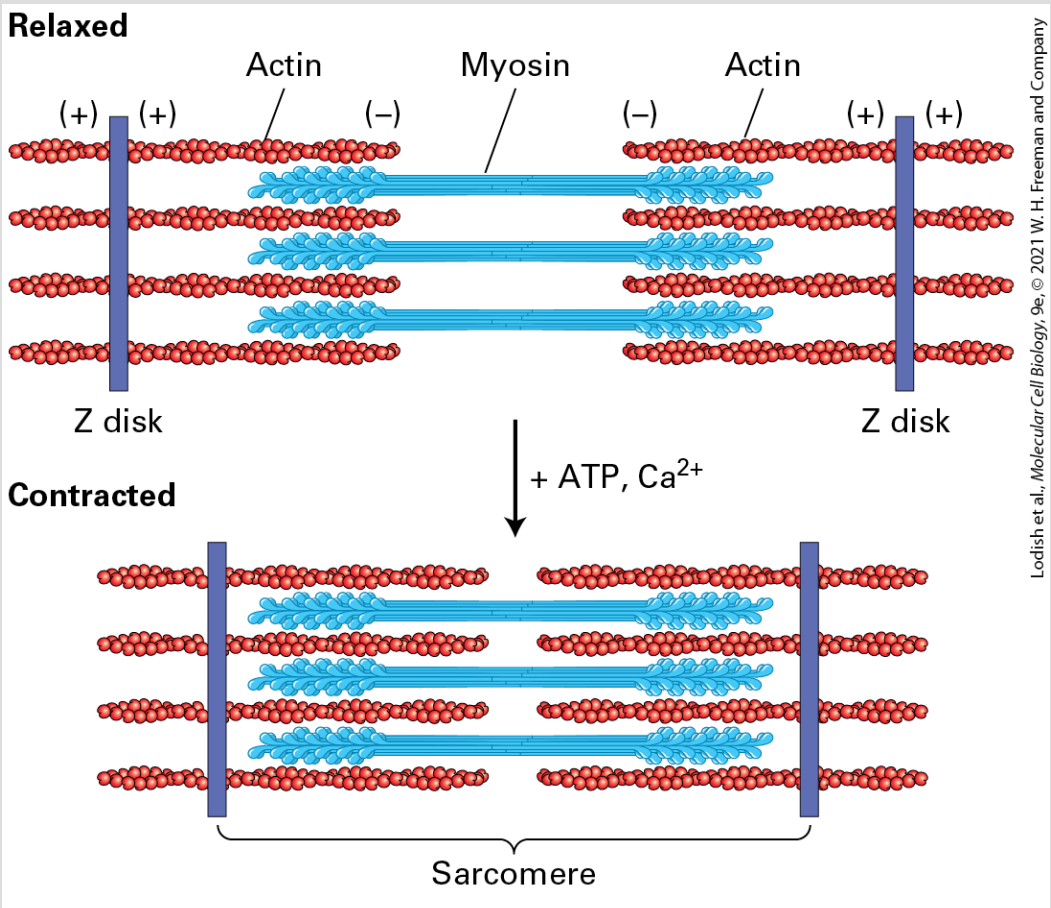


FIGURE 17-30 The sliding-filament model of contraction in skeletal muscle. The arrangement of thick myosin and thin actin filaments in the relaxed state is shown in the top diagram. In the presence of ATP and Ca^{2+}, the myosin heads extending from the thick filaments walk toward the $(+)$ ends of the thin filaments. Because the thin filaments are anchored at the Z disks (purple), movement of myosin pulls the actin filaments toward the center of the sarcomere, shortening its length in the contracted state, as shown in the bottom diagram.

The human heart is an amazing contractile organ — it contracts without interruption about 3 million times a year, or a fifth of a billion times in a lifetime. The muscle cells of the heart contain contractile machinery very similar to that of skeletal muscle, except

that they are mono- and bi-nucleated cells. In each cell, the end sarcomeres insert into structures at the plasma membrane called *intercalated disks*, which link the cells into a contractile chain. Heart muscle cells are generated only early in life, so they cannot be replaced in response to damage, such as occurs during a heart attack. Many different mutations in proteins of the heart contractile machinery give rise to *hypertrophic cardiomyopathies* — thickening of the heart wall muscle, which compromises its function. For example, many mutations have been documented in the cardiac myosin heavy-chain gene that compromise the protein's contractile function even in heterozygous individuals. In such individuals, the heart tries to compensate by hypertrophy (enlargement), which often results in fatal heart arrhythmia (irregular beating). In addition to myosin heavy-chain defects, defects that result in cardiomyopathies have been traced to mutations in other components of the contractile machinery, including actin, myosin light chains, tropomyosin and troponin, and structural components such as titin (discussed below). ■

Skeletal Muscle Is Structured by Stabilizing and Scaffolding Proteins

The structure of the sarcomere is maintained by a number of accessory proteins (Figure 17-31). The actin filaments are stabilized on their (+) ends by CapZ and on their (−) ends by tropomodulin

(see Section 17.2). A giant protein known as *nebulin* extends along the thin filament all the way from the Z disk to tropomodulin, to which it binds. Nebulin consists of repeating domains that bind to actin subunits in the filament, and it is believed that the number of actin-binding repeats, and therefore the length of nebulin, determines the length of the thin filaments. Another giant protein, called *titin*, has its head associated with the Z disk and extends to the M band in the middle of the thick filament, where another titin molecule extends to the subsequent Z disk. Titin is believed to be an elastic molecule that holds the thick filaments in the middle of the sarcomere and also prevents overstretching to ensure that the thick filaments remain interdigitated between the thin filaments.

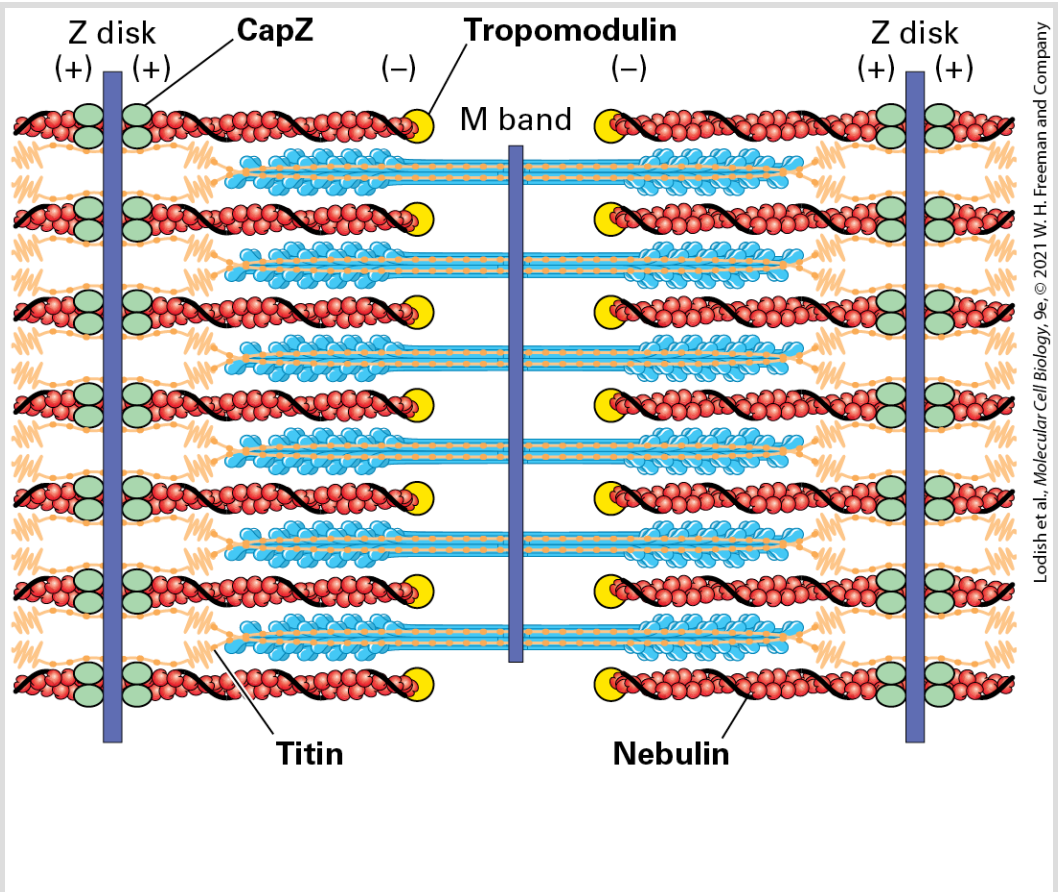


FIGURE 17-31 Accessory proteins found in skeletal muscle. To stabilize the actin filaments, CapZ caps the (+) end of the thin filaments at the Z disk, whereas tropomodulin caps the (−) end. The giant protein titin extends from the M band through the thick filaments and attaches to the Z disk. Nebulin binds actin subunits and determines the length of the thin filament.

Contraction of Skeletal Muscle Is Regulated by Ca^{2+} and Actin-Binding Proteins

Like many cellular processes, skeletal muscle contraction is initiated by an increase in the cytosolic Ca^{2+} concentration. As described in Chapter 11, the Ca^{2+} concentration of the cytosol is normally kept below 0.1 μM. In skeletal muscle cells, a low cytosolic Ca^{2+} level is maintained primarily by a unique Ca^{2+} ATPase that continually pumps Ca^{2+} ions from the cytosol containing the myofibrils into the **sarcoplasmic reticulum (SR)**, a specialized endoplasmic reticulum of muscle cells (Figure 17-32). This activity establishes a reservoir of Ca^{2+} in the SR.

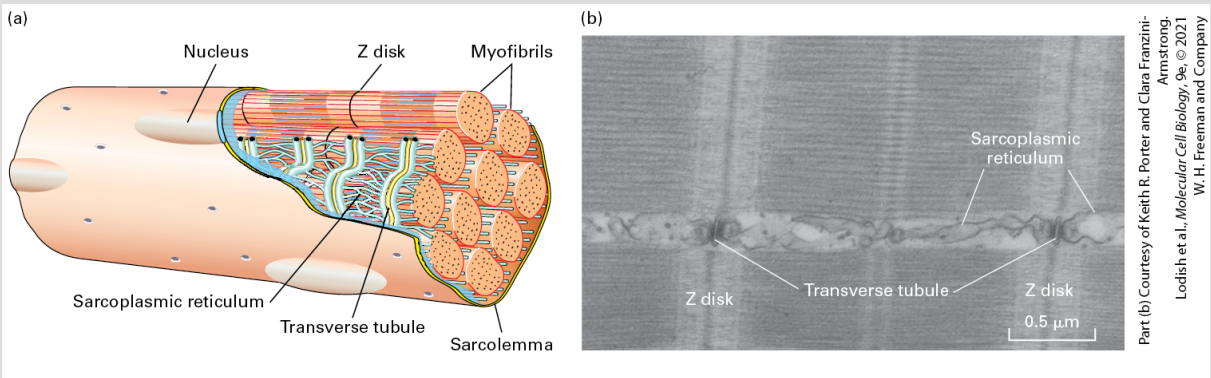


FIGURE 17-32 The sarcoplasmic reticulum regulates the level of free Ca^{2+} in myofibrils. (a) When a nerve impulse stimulates a muscle cell, the action potential is transmitted down a transverse tubule (yellow), which is continuous with the plasma membrane (sarcolemma), leading to release of Ca^{2+} from the adjacent sarcoplasmic reticulum (blue) into the myofibrils. (b) Thin-section electron micrograph of skeletal muscle, showing the intimate relationship of the sarcoplasmic reticulum to the muscle fibers.

The arrival of a nerve impulse (or *action potential*; see Chapter 23) at a neuromuscular junction triggers an action potential in the muscle-cell plasma membrane (also known as the *sarcolemma*). The action potential travels down invaginations of the plasma membrane known as *transverse tubules*, which pass into the cell and lie around each myofibril. The arrival of the action potential in the transverse tubules stimulates the opening of voltage-gated Ca^{2+} channels in the SR membrane, and the ensuing release of Ca^{2+} from the SR raises the cytosolic Ca^{2+} concentration in the myofibrils. This elevated Ca^{2+} concentration induces changes in two accessory proteins, tropomyosin and troponin, which are bound to the actin thin filaments and normally block myosin binding. Changes in the positions of these proteins on the actin thin filaments in turn permit

the myosin-actin interactions and hence contraction. This type of regulation is very rapid and is known as *thin-filament regulation*.

Tropomyosin (TM) is a ropelike molecule, about 40 nm in length, that binds to seven actin subunits in an actin filament. TM molecules are strung together head to tail, forming a continuous chain along each side of the actin thin filament (Figure 17-33a, b). Associated with each TM molecule is *troponin* (TN), a complex of three subunits, TN-T, TN-I, and TN-C. TN-C is the calcium-binding subunit of troponin; it controls the position of TM on the surface of an actin filament through the TN-I and TN-T subunits.

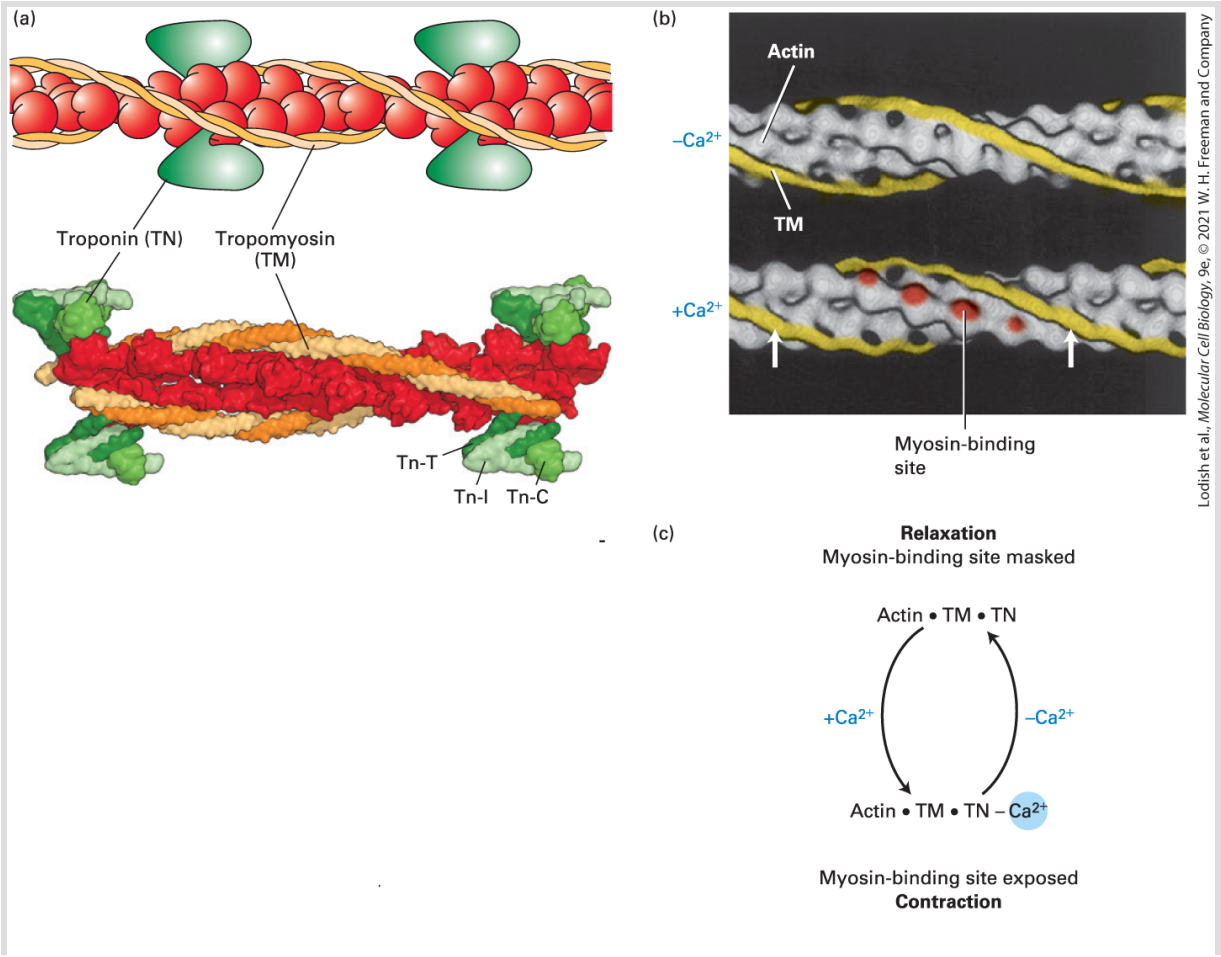


FIGURE 17-33 Ca^{2+}-dependent thin-filament regulation of skeletal muscle contraction. (a) Model and the corresponding structure of the tropomyosin-troponin regulatory complex on a thin filament. Troponin is a protein complex that is bound to the long α-helical tropomyosin molecule. (b) Three-dimensional electron-microscope reconstructions of the tropomyosin helix (yellow) on a muscle thin filament. Tropomyosin in the relaxed state (*top*) shifts to a new position (arrows) in the state inducing contraction (*bottom*) when the Ca^{2+} concentration in the sarcoplasm increases. This movement exposes myosin-binding sites (red) on actin. (Troponin is not shown in this representation, but it remains bound to tropomyosin in both states.) (c) Summary of the regulation of skeletal muscle contraction by Ca^{2+} binding to troponin.

[Part (a) Data from S. Wu et al., 2012, *PLOS One* **7**:39422, PDB ID 2w4u. Part (b) P. Vibert, R. Craig, and W. Lehman, 1993, *J. Cell Biol.* **123**(2):313–321; https://doi.org/10.1083/jcb.123.2.313.]

Under the control of Ca^{2+} and TN, TM can occupy either of two positions on a thin filament. In the absence of Ca^{2+}, TM blocks myosin's interaction with F-actin, and the muscle is relaxed. Binding of Ca^{2+} ions to TN-C triggers movement of TM to a new position on the filament, thereby exposing the myosin-binding sites on actin (Figure 17-33b). Thus at Ca^{2+} concentrations greater than 1 μM the inhibition exerted by the TM-TN complex is relieved, and contraction occurs. The Ca^{2+}-dependent cycling between relaxation and contraction states in skeletal muscle is summarized in Figure 17-33c.

Heart muscle, like skeletal muscle, is subject to thin-filament regulation, using cardiac-specific tropomyosin and troponin. During a heart attack (myocardial infarction), heart cells are deprived of

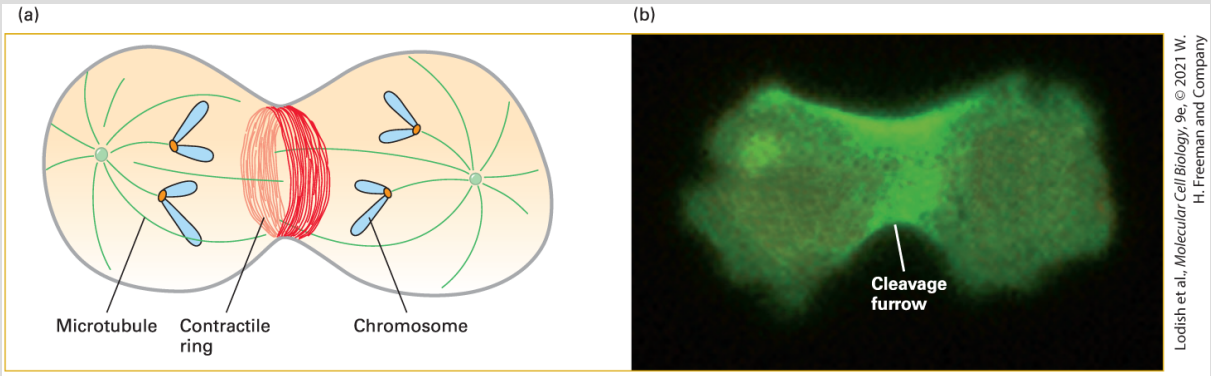
sufficient oxygen (cardiac ischemia) and may subsequently die. The plasma membrane of the dead cells ruptures and releases cellular components into the bloodstream. Blood tests that specifically measure the level of cardiac-specific troponins are used by physicians to determine the severity of a heart attack. ■

Actin and Myosin II Form Contractile Bundles in Nonmuscle Cells

In skeletal muscle, as we have seen, actin thin filaments and myosin II thick filaments assemble into highly organized contractile structures. Nonmuscle cells contain several types of **contractile bundles** composed of actin and myosin II filaments, which are similar to skeletal muscle fibers but much less organized. Moreover, these contractile bundles lack the troponin regulatory system and are instead regulated by myosin phosphorylation, as we discuss below. Let's look at a few examples.

In epithelial cells, contractile bundles are most commonly found as an adherens belt, also known as the *circumferential belt*, that encircles the inner surface of the cell at the level of the adherens junction (see Figure 17-4a). These bundles are important in maintaining the integrity of the epithelium (discussed in Chapter 20). A second type of contractile bundle are stress fibers, which are seen along the lower surfaces of cells cultured on artificial (glass or

plastic) surfaces or on extracellular matrices (see Figure 17-4a, b) and are important in cell adhesion. The ends of stress fibers terminate at integrin-containing focal adhesions, special structures that attach a cell to the underlying substratum (see Figure 17-39b, c and Chapter 20). Circumferential belts and stress fibers contain several proteins found in the contractile apparatus of smooth muscle and exhibit some organizational features resembling those of muscle sarcomeres. A third type of contractile bundle is the contractile ring, a transient structure in animal cells that assembles at the equator of a dividing cell, encircling the cell midway between the poles of the mitotic spindle (Figure 17-34a). As the ring contracts, pulling the plasma membrane in, the cytoplasm is divided and eventually pinched into two parts in a process known as *cytokinesis*, giving rise to two daughter cells. In a dividing cell expressing GFP-myosin-II, myosin II is localized to the contractile ring (Figure 17-34b). As evidence for the role of myosin II in forming the cleavage furrow, when the gene for the heavy chain of myosin II has been deleted, cells are unable to undergo cytokinesis. Instead, these cells form a multinucleated syncytium because cytokinesis, but not nuclear division, is inhibited.



EXPERIMENTAL FIGURE 17-34 Localization of myosin II during cytokinesis. (a) Diagram of a cell going through cytokinesis, showing the mitotic spindle (microtubules green, chromosomes blue) and the contractile ring with actin filaments (red). (b) Fluorescence micrograph of a *Dictyostelium* amoeba expressing GFP-myosin-II reveals myosin-II enrichment in the cleavage furrow cortex, as the cell pinches into two during cytokinesis.

[Part (b) Republished with permission from Elsevier, from J. C. Effler et al., 2006, *Curr. Biol.* **16**(9):1962–1967, Fig. 1A (149); permission conveyed through Copyright Clearance Center, Inc.; photo courtesy Douglas Robinson.]

Myosin-Dependent Mechanisms Regulate Contraction in Smooth Muscle and Nonmuscle Cells

Like skeletal muscle, smooth muscle is composed of contractile cells. Smooth muscle is found in the walls of hollow organs and the walls of passageways. For example, smooth muscle surrounds blood vessels to regulate blood pressure, surrounds the intestine to move food through the gut, and restricts airway passages in the lung. The contractile apparatus of smooth muscle and its regulation are a valuable model for studying these processes because contraction in

smooth muscle is regulated in a manner similar to that in nonmuscle cells. As we have just seen, in skeletal muscle the activity of the myosin II is not directly regulated, but its accessibility to F-actin is regulated by the tropomyosin-troponin complex bound to the actin thin filament that switches between the contraction-inducing state in the presence of Ca^{2+} and the relaxed state in its absence. In contrast, smooth muscle contraction is regulated by the cycling of myosin II between on and off states. Myosin II cycling, and thus contraction of smooth muscle and nonmuscle cells, is regulated in response to many extracellular signaling molecules.

Contraction of vertebrate smooth muscle is regulated primarily by a pathway in which the *myosin regulatory light chain (LC)* associated with the myosin II neck domain (see Figure 17-22b) undergoes phosphorylation and dephosphorylation. When the regulatory LC is not phosphorylated, smooth muscle myosin II folds into a self-inhibited state where its ATPase cycle is inactive. When the regulatory LC is phosphorylated by the enzyme **myosin light-chain kinase (MLC kinase)**, myosin II opens up by unfolding, assembles into filaments, and becomes able to induce contraction (Figure 17-35). The activity of MLC kinase is regulated by the level of free Ca^{2+} in the cytoplasm. The Ca^{2+}-dependent regulation of MLC kinase activity is mediated by the Ca^{2+}-binding protein calmodulin (see Figure 3-34). When an extracellular signal leads to a rise in intracellular free Ca^{2+} concentration, Ca^{2+} binds to calmodulin inducing a conformational change in the protein. The

Ca^{2+}/calmodulin complex then binds to MLC kinase and activates it. This enables MLC kinase to phosphorylate the myosin regulatory LC. When the cytosolic Ca^{2+} concentration returns to its resting level, MLC kinase becomes inactive, and myosin light-chain phosphatase removes the phosphates from myosin, which returns the system to its relaxed state. This mode of regulation relies on the action of a protein kinase and on the diffusion of Ca^{2+} over greater distances than in sarcomeres, so contraction is slower in smooth muscle than in skeletal muscle. Because this regulation involves myosin, it is known as *thick-filament regulation*.

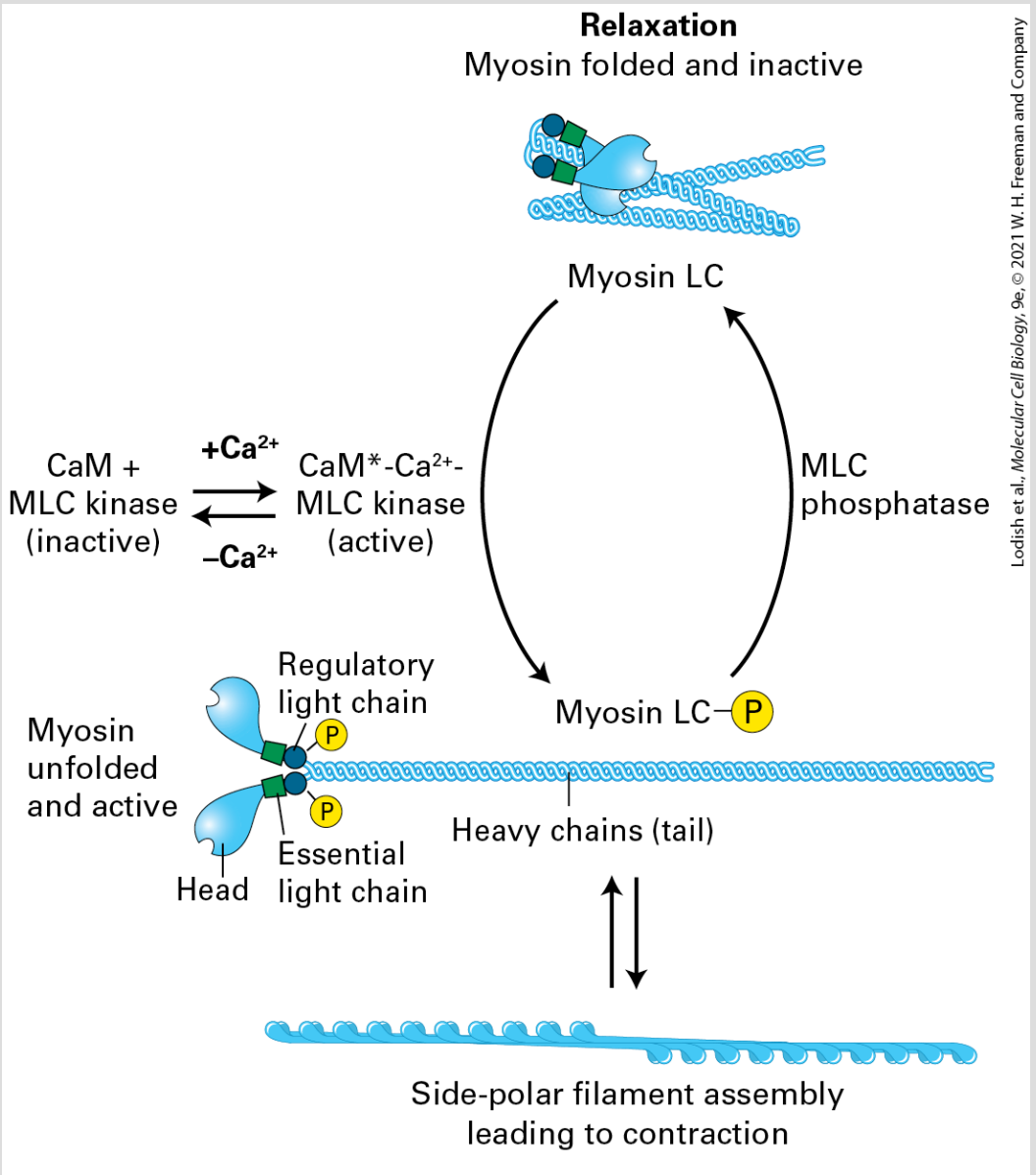


FIGURE 17-35 Myosin light-chain phosphorylation regulates smooth muscle contraction. In vertebrate smooth muscle, phosphorylation of the myosin regulatory light chain (LC) activates contraction. At Ca^{2+} concentrations of less than 10^{-6} M, the regulatory light chain is not phosphorylated, and the myosin adopts a folded conformation. When the Ca^{2+} level rises, Ca^{2+} binds calmodulin (CaM), which undergoes a conformational change (CaM*). The CaM*-Ca^{2+} complex binds and activates myosin light-chain kinase (MLC kinase), which then phosphorylates the myosin LC. This phosphorylation event opens myosin II, which assembles into side-polar filaments that participate in contraction. Note that the organization of side-polar filaments is distinct from

that of the bipolar filaments seen in skeletal muscle. When Ca^{2+} levels drop, the myosin LC is dephosphorylated by myosin light-chain (MLC) phosphatase, causing muscle relaxation.

Unlike skeletal muscle, which is stimulated to contract solely by nerve impulses, smooth muscle cells and nonmuscle cells are regulated by many types of external signals. For example, norepinephrine, angiotensin, endothelin, histamine, and other signaling molecules modulate or induce the contraction of smooth muscle or elicit changes in the shape and adhesion of nonmuscle cells by triggering various signal transduction pathways. Some of these pathways lead to an increase in the cytosolic Ca^{2+} concentration, which can stimulate myosin activity by activating MLC kinase (see Figure 17-35). As we will see in Section 17.7, other signaling pathways activate *Rho kinase*, which also activates myosin activity by phosphorylating the regulatory light chain, but in a Ca^{2+}-independent manner.

Myosin V Carries Vesicles Along Actin Filaments

The myosin V family of proteins transport cargo along actin filaments and are the most processive myosin motors known. In the next chapter we discuss how they work in conjunction with microtubule motors to transport organelles. Here we look at how these motors carry endocytic and secretory vesicles around the cell.

Much is known about myosin V motors in experimentally accessible systems such as the budding yeast *Saccharomyces cerevisiae*. This well-studied organism grows by budding, which requires its secretory machinery to target newly synthesized materials to the growing bud (Figure 17-36a). Myosin V transports secretory vesicles along actin filaments at 3 μm/s to grow the bud (Figure 17-36b). Myosin V proteins are also important at a later stage of the yeast cell cycle, when the organelles have to be distributed between the mother and daughter cells. Myosin V molecules in yeast transport some of the mother cell's peroxisomes, mitochondria, vacuoles, endoplasmic reticulum, the *trans*-Golgi network, and even the ends of microtubules and some specific messenger RNAs into the bud (Figure 17-36c). These organelles each have a distinct receptor that the tail of myosin V binds to. The motor makes many delivery cycles, so it has to have a way to pick up, transport, and deliver its organelle cargo. Recent work has shown that myosin V can exist in two states: an inactive folded state in which the tail domain binds and inhibits the activity of the motor head domain, and an active state in which binding to cargo unfolds the tail, thus releasing the head from the tail, which allows the myosin motor to operate (Figure 17-36d, e). How the motor releases its cargo upon delivery is poorly understood — although in one case, it is known that the receptor on the organelle that binds myosin becomes degraded upon delivery of the organelle to its destination. Whereas budding yeast use myosin V and polarized actin filaments to transport many organelles, animal cells employ microtubules and their motors to transport many of the

same organelles over the longer distances inside an animal cell. We discuss those transport mechanisms in the next chapter.

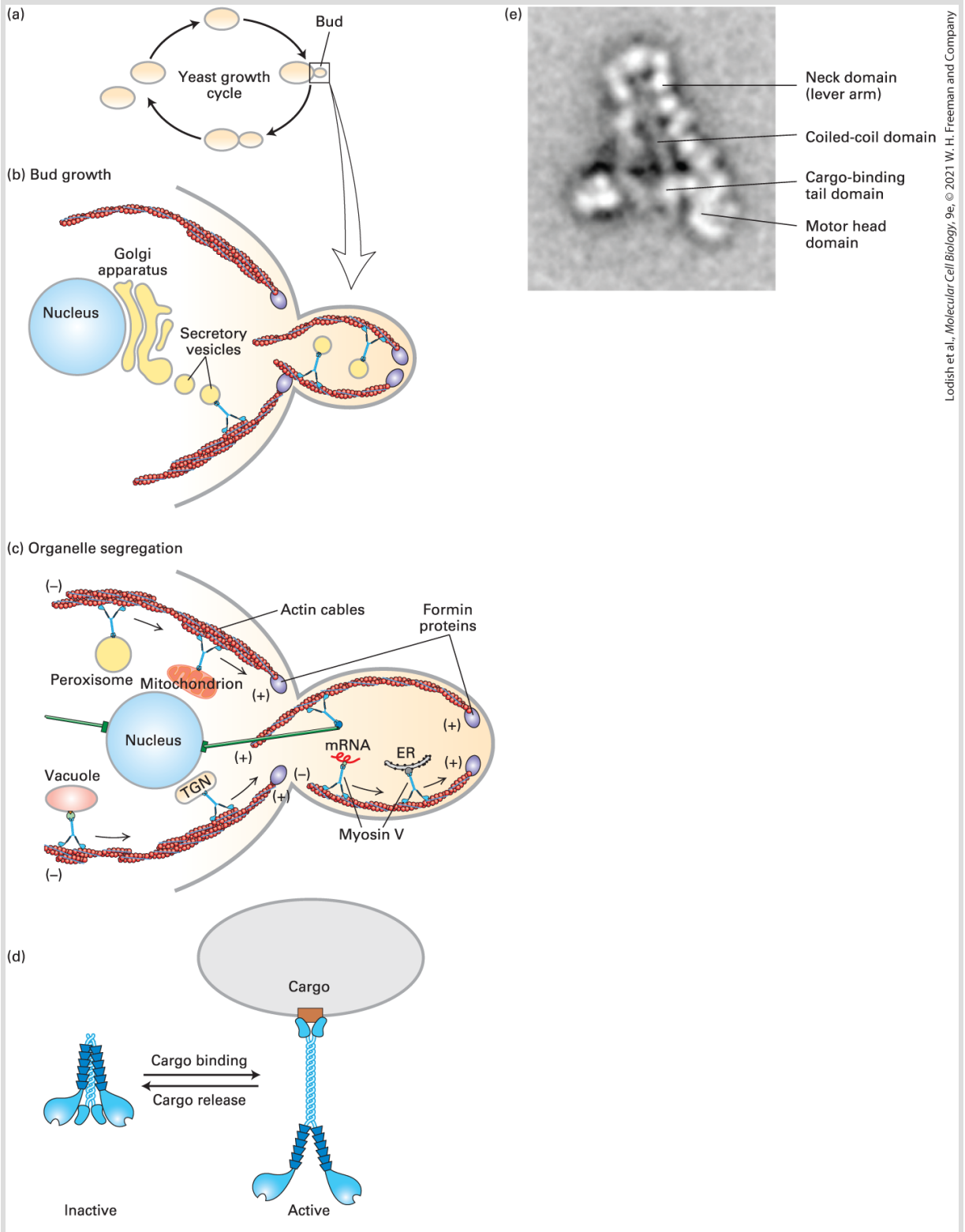


FIGURE 17-36 Cargo movement by myosin V. (a) The yeast *Saccharomyces cerevisiae* (used in making bread, beer, and wine) grows by budding. (b) Secretory vesicles are transported by myosin V along actin cables, nucleated by formins, into the bud, which grows to swell to

about the size of the mother cell. (c) Before cell division can occur, all the organelles must be segregated between the mother and the bud, including the vacuole (the yeast equivalent of a lysosome), peroxisomes, mitochondria, endoplasmic reticulum (ER), *trans*-Golgi network (TGN), and even selected mRNAs into the bud, which are all mediated by myosin V. Myosin V also binds the ends of cytoplasmic microtubules (green) to orient the nucleus in preparation for mitosis. See D. Pruyne et al., 2004, *Ann. Rev. Cell Dev. Biol.* **20**:559. (d) Myosin V is regulated by binding cargo. In the inactive folded state, the tails of myosin V bind and inactivate the motor head domains. Upon cargo binding, the head-to-tail interaction is alleviated, and the motor is now active to transport cargo. (e) Electron micrograph of inactive myosin V. Myosin V molecules were negatively stained and viewed in a transmission electron microscope. The image shown is the composite average from several individual images.

[Part (e) Reprinted by permission from Macmillan Publishers Ltd, from K. Thirumurugan et al., 2006, "The Cargo-Binding Domain Regulates Structure and Activity of Myosin 5," *Nature* **442**:212–215; permission conveyed through Copyright Clearance Center, Inc.]

Perhaps the most dramatic example of movement by myosin V is seen in giant green algae, such as *Nitella* and *Chara*. These algae can be found in ponds during the summer and their movement easily observed using a simple microscope. In their large cells, which can be as much as 2 cm in length, cytosol flows rapidly, at a rate approaching 4.5 mm per minute, in an endless loop around the inner circumference of the cell (Figure 17-37). This *cytoplasmic streaming* is a principal mechanism for distributing cellular metabolites and is also seen in other large cells such as plant cells and amoebae. The algal cells have bundles of actin filaments aligned along the length of the cell, lying just above the stationary chloroplasts located adjacent to the membrane. The bulk cytosol is propelled by myosin V (also known as *myosin XI* in plants) attached

to parts of the ER adjacent to the actin filaments. The flow rate of the cytosol in *Nitella* is about 15 times faster than movement produced by any other myosins. ■

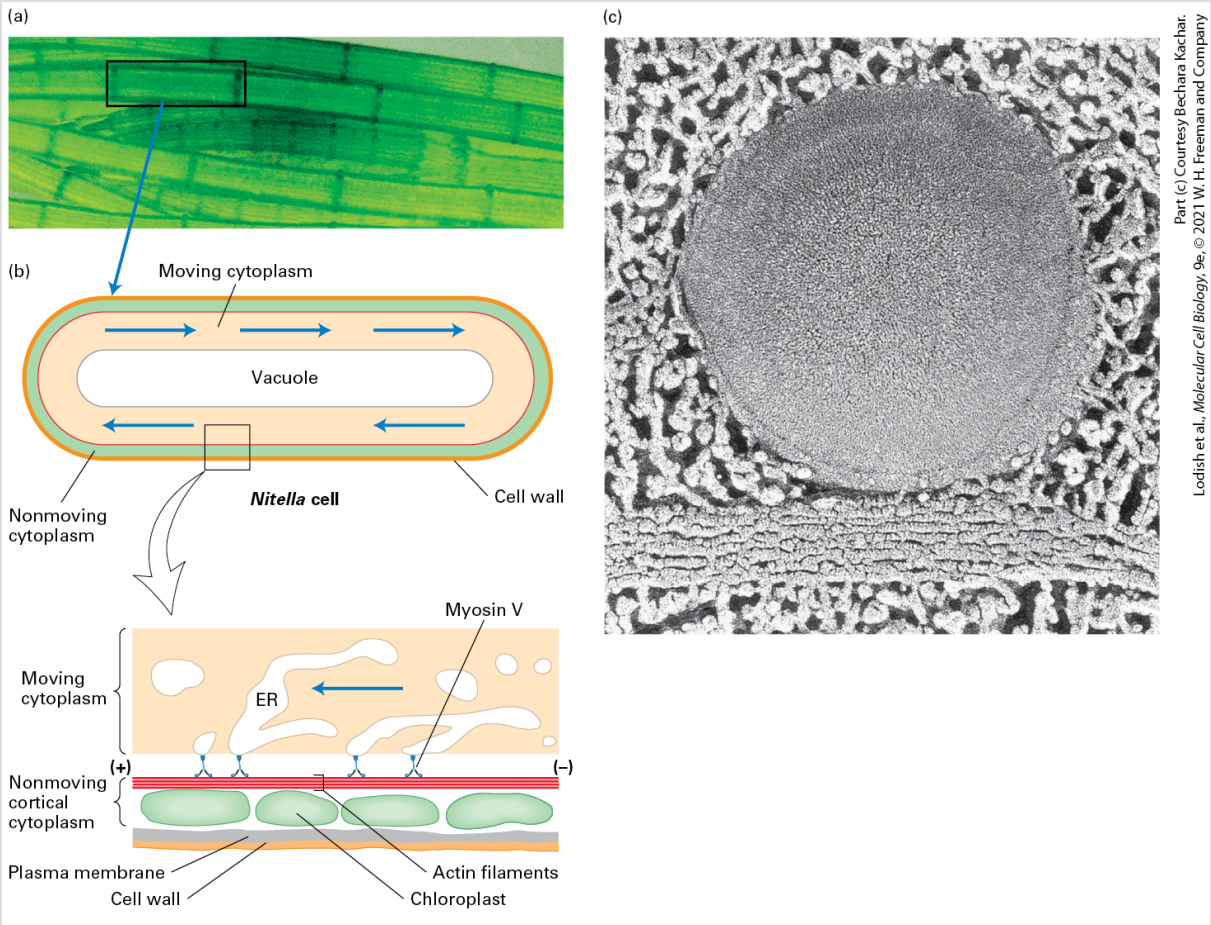


FIGURE 17-37 Cytoplasmic streaming in cylindrical giant algae. (a) Cells of *Nitella*, a freshwater alga commonly found in ponds in the summer. The cytoplasmic movement, described below, can be readily observed with a simple microscope, so go find some *Nitella* (or related algae) and watch this amazing phenomenon! (b) The center of a *Nitella* cell is filled with a single large water-filled vacuole, which is surrounded by a layer of moving cytoplasm (blue arrows). A nonmoving layer of cortical cytoplasm filled with chloroplasts lies just under the plasma membrane (enlarged in bottom figure). On the inner side of this layer are bundles of stationary actin filaments (red), all oriented with the same polarity. A motor protein (blue), a plant myosin V, carries parts of the endoplasmic reticulum (ER) along the actin filaments. The movement of the ER network propels the entire viscous cytoplasm,

including organelles that are enmeshed in the ER network. (c) Electron micrograph of the cortical cytoplasm showing a large vesicle connected to an underlying bundle of actin filaments.

KEY CONCEPTS OF SECTION 17.6

Myosin-Powered Movements

- In skeletal muscle, contractile myofibrils are composed of thousands of repeating units called *sarcomeres*. Each sarcomere consists of two interdigitating filament types: myosin thick filaments and actin thin filaments (see Figure 17-29).
- Skeletal muscle contraction involves the ATP-dependent sliding of myosin thick filaments along actin thin filaments to shorten the sarcomere and hence the myofibril (see Figure 17-30).
- The ends of the actin thin filaments in skeletal muscle are stabilized by CapZ at the $(+)$ end and by tropomodulin at the $(-)$ end. Two large proteins, nebulin associated with the thin filaments and titin with the thick filaments, also contribute to skeletal muscle organization.
- Skeletal muscle contraction is subject to thin-filament regulation. At low levels of free Ca^{2+}, tropomyosin blocks the interaction of myosin and F-actin, and the muscle is relaxed. At elevated levels of free Ca^{2+}, the troponin complex associated with tropomyosin binds Ca^{2+} and moves the tropomyosin to uncover the myosin-binding sites on actin, allowing contraction (see Figure 17-33).
- Smooth muscle and nonmuscle cells have contractile bundles of actin and myosin filaments with an organization similar to that in skeletal muscle but less well ordered.
- Contractile bundles are subject to thick-filament regulation. A myosin light chain is phosphorylated by myosin light-chain kinase, which activates myosin and hence induces contraction. The MLC kinase is activated by binding to Ca^{2+}-calmodulin when the free intracellular Ca^{2+} concentration rises (see Figure 17-35).
- Inactive myosin V is activated by binding cargo, which it transports processively along actin filaments.

17.7 Cell Migration: Mechanism, Signaling, and Chemotaxis

So far in this chapter, we have examined some of the mechanisms cells use to create movement using actin, myosin, and other actin-associated proteins. Some of these mechanisms generate the forces used by cells to migrate. **Cell migration** results from the coordination of actin-based motions generated in different parts of a cell, which are integrated with an endocytic cycle.

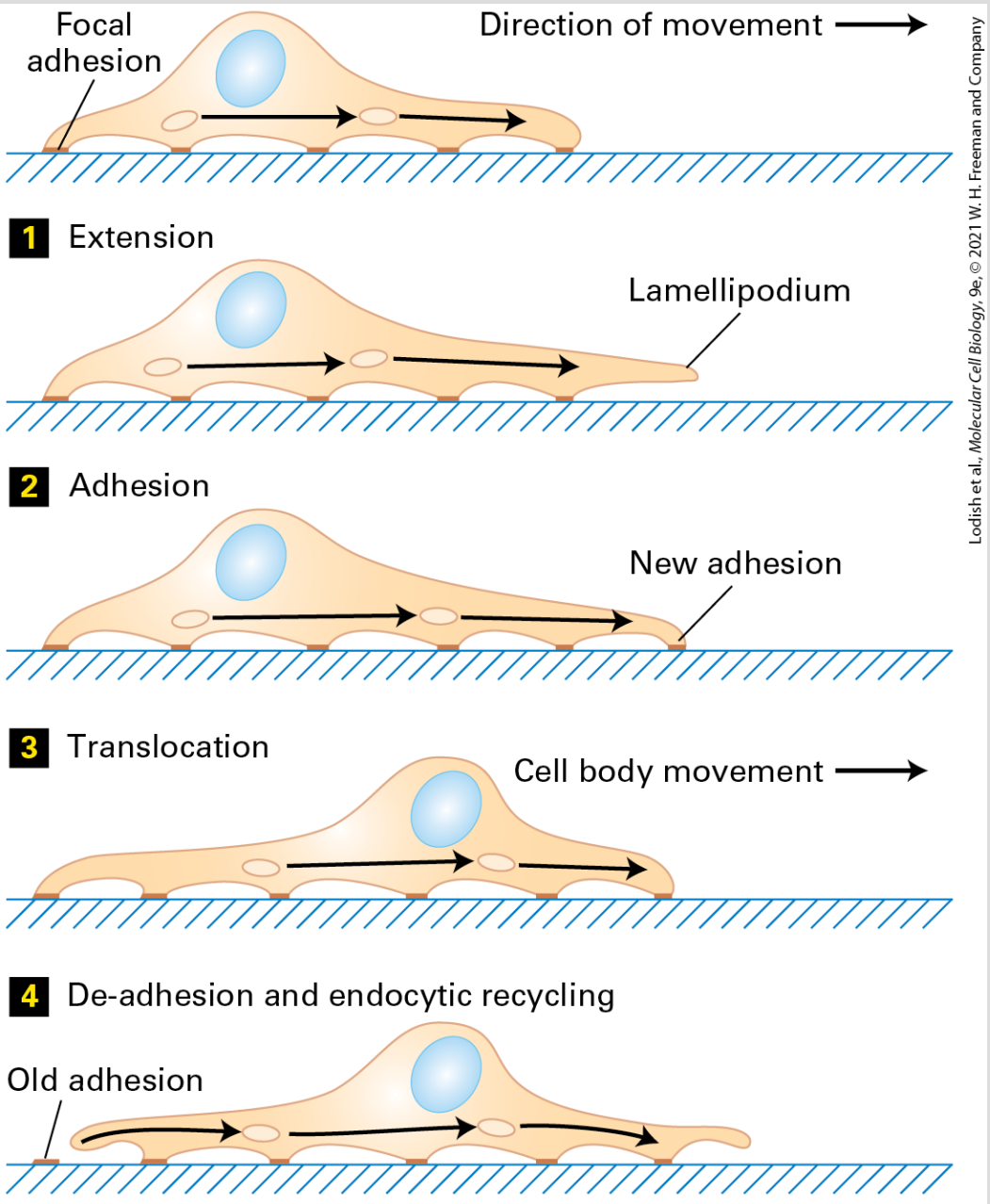
The study of cell migration is important to many fields of biology and medicine. For example, an essential feature of animal development is the migration of specific cells along predetermined paths. Epithelial cells in an adult animal migrate to heal a wound, and white blood cells migrate to sites of infection. Less obvious are the continuous slow migration of intestinal epithelial cells along the villi in the intestine and of endothelial cells that line the blood vessels. The inappropriate migration of cancer cells after breaking away from their normal tissue results in metastasis.

Cell migration is initiated by the formation of a large, broad membrane protrusion at the leading edge of a cell. Video microscopy reveals that a major feature of this movement is the polymerization of actin at the membrane. Actin filaments at the leading edge are rapidly cross-linked into bundles and networks in a protruding region, called a *lamellipodium* in vertebrate cells. In

some cases, slender, fingerlike membrane projections, called *filopodia*, also extend from the leading edge. These structures form stable contacts with the underlying surface (e.g., the extracellular matrix) that the cell moves across. In this section, we take a closer look at how cells coordinate various microfilament-based processes with endocytosis to move across a surface. We also consider the role signaling pathways play in coordinating and integrating the actions of the cytoskeleton, a major focus of current research.

Cell Migration Coordinates Force Generation with Cell Adhesion and Membrane Recycling

If you watch a moving fibroblast (connective tissue cell), it displays a characteristic sequence of events as it migrates over a surface: initial extension of a membrane protrusion, attachment to the substratum, forward flow of cytosol, and retraction of the rear of the cell (Figure 17-38). These events occur in an ordered pattern in a slowly moving cell such as a fibroblast, but in rapidly moving cells, such as macrophages, all four steps occur simultaneously in a coordinated manner. In this section, we first consider the role the actin cytoskeleton plays in assembly at the leading edge as well as attachment to (and detachment from) the substratum (Figure 17-39a, b). Second, we discuss the role of the endocytic cycle in cell migration.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 17-38 Steps in cell locomotion. Movement begins with the extension of one or more lamellipodia from the leading edge of the cell (step 1); some lamellipodia adhere to the substratum by focal adhesions (step 2). Then the bulk of the cytoplasm in the cell body flows forward due to contraction at the rear of the cell (step 3). The trailing edge of the cell remains attached to the substratum until the tail eventually detaches and retracts into the cell body. During this cytoskeleton-based cycle, the endocytic cycle internalizes

membrane and integrins at the rear of the cell and transports them to the front of the cell (arrows) for reuse in making new adhesions (step **4**).

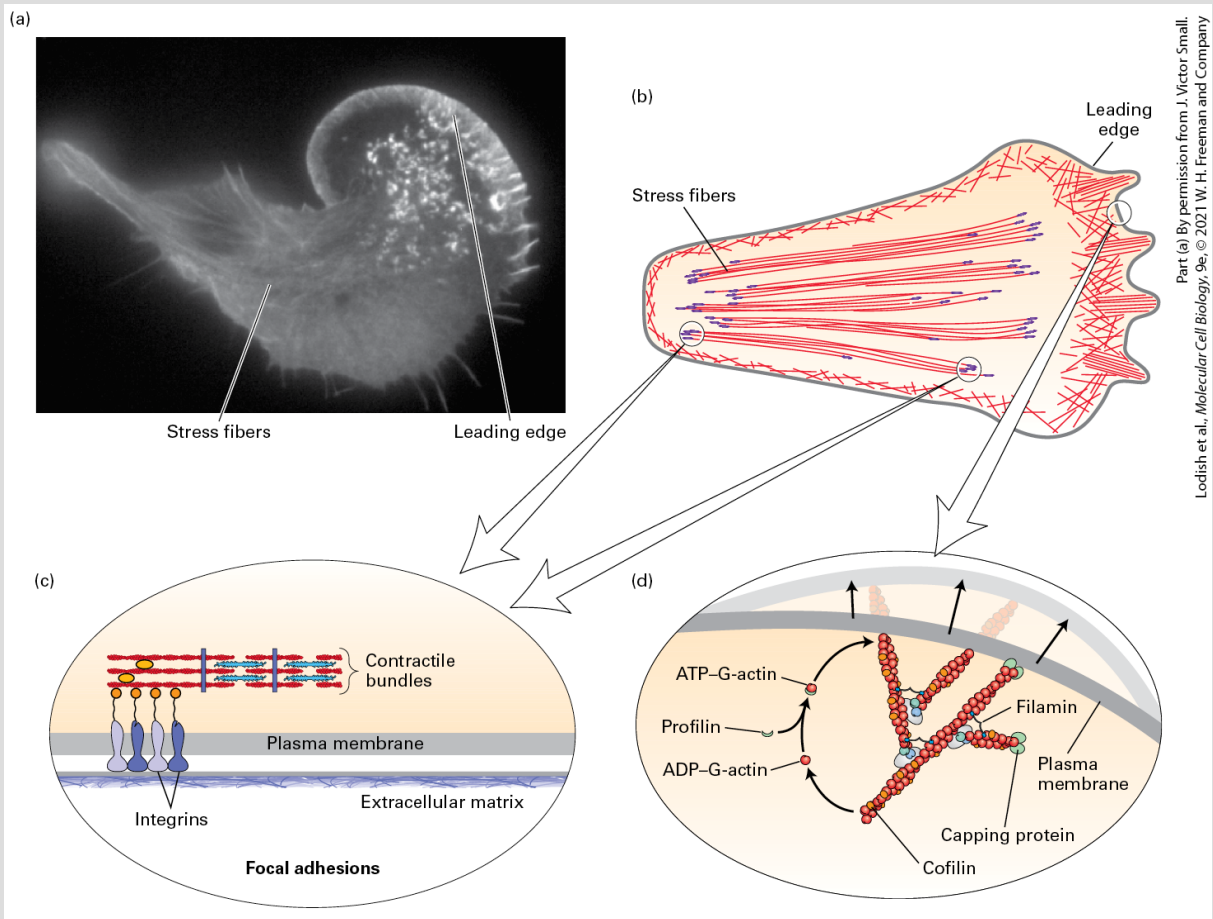


FIGURE 17-39 Actin-based structures involved in cell locomotion. (a) Localization of actin in a fibroblast expressing GFP-actin. (b) Diagram of the classes of microfilaments involved in cell migration. The network of actin filaments in the leading edge advances the cell forward. Contractile fibers in the cell cortex squeeze the cell body forward, and stress fibers terminating in focal adhesions also pull the bulk of the cell body up as the rear adhesions are released. (c) The structure of focal adhesions involves the attachment of the ends of stress fibers through integrins to the underlying extracellular matrix. Focal adhesions also contain many signaling molecules important for cell locomotion. (d) The dynamic actin network in the leading edge is nucleated by the $\mathrm{Arp2/3}$ complex and employs the same set of factors that control assembly and disassembly of actin filaments in the *Listeria* tail (see Figure 17-17).

Step 1: Membrane Extension

The network of actin filaments at the leading edge is a type of cellular engine that pushes the membrane forward in a manner very similar to the propulsion of *Listeria* by actin polymerization (Figure 17-39d; for *Listeria*, see Figure 17-17c). At the membrane of the leading edge, actin is nucleated by the activated Arp2/3 complex (see Figure 17-15), and filaments are elongated by assembly onto (+) ends adjacent to the plasma membrane. Because the actin network is fixed with respect to the substratum, the front membrane is pushed out as the filaments elongate. Actin turnover, and thus treadmilling, is mediated, as it is in the comet tails of *Listeria*, by the action of profilin and cofilin (Figure 17-39d).

Step 2: Cell-Substratum Adhesions

When the membrane has been extended and the actin network has been assembled, the plasma membrane becomes firmly attached to the substratum. Time-lapse microscopy shows that actin bundles in the leading edge become anchored to structures known as *focal adhesions* (Figure 17-39c). The membrane proteins within a focal adhesion that connect actin bundles to the extracellular matrix are called *integrins*. Integrins have an external domain that binds to specific components of the extracellular matrix, such as fibronectin and collagen, and a cytoplasmic domain that links them to the actin cytoskeleton (see Figure 17-39c and Chapter 20). The attachment of

cell to substrate serves two purposes: it prevents the lamellipodium from retracting, and it allows the cell to move forward. Given the importance of focal adhesions and their regulation during cell locomotion, it is not surprising that they are very rich in molecules involved in signal transduction pathways. Focal adhesions are discussed in Chapter 20, where we discuss in detail the interactions that occur between a cell and the extracellular matrix.

Step 3: Cell-Body Translocation

After the forward adhesions at the leading edge have been made, the bulk contents of the cell body are translocated forward (see Figure 17-38, step **3**). The nucleus and the other organelles embedded in the cytoskeleton are moved forward by myosin II–dependent cortical contraction in the rear part of the cell, like toothpaste when the lower half of the tube is squeezed. Consistent with this model, myosin II is localized to the rear cell cortex. As the cell migrates forward, the adhesions effectively move toward the rear of the cell.

Step 4a: Breaking Cell Attachments

Finally, in the last step of movement, the focal adhesions at the rear of the cell are broken (de-adhesion), the integrins are recycled, and the freed tail of the cell is brought forward. In the light microscope, the tail is often seen to snap loose from its connections — perhaps by the contraction of stress fibers in the tail or by elastic tension —

and it sometimes leaves a little bit of its membrane behind, still firmly attached to the substratum.

Cells cannot move as just described if they are either too strongly attached or too weakly attached to a surface. The ability of a cell to move corresponds to a balance between the mechanical forces generated by the cytoskeleton and the resisting forces generated by cell adhesions. This relationship can be demonstrated by measuring the rate of movement in cells that express varying levels of integrins. Such measurements show that the fastest migration occurs at an intermediate level of adhesion, with the rate of movement falling off at high and low levels of adhesion. Cell locomotion thus results from traction forces exerted by the cell on the underlying substratum.

Step 4b: Recycling Membrane and Integrins by Endocytosis

The dynamic changes in the actin cytoskeleton alone are not sufficient to drive cell migration. It is also dependent on endocytic recycling of plasma membrane. The membrane needed during lamellipodium extension is provided by internal endosomes as the membrane recycles back to the cell surface. Adhesion molecules in focal adhesions at the rear of the cell are internalized as those adhesions are disassembled and are then transported by an endocytic cycle to the front of the cell to make new substratum attachments at the leading edge (see Figure 17-38, step **4**). This

cycling of adhesion molecules in a migrating cell resembles the way a tank uses its treads to move forward. The movement of membrane internally through the cell also generates a rearward membrane flow across the surface of the cell. Indeed, this type of flow may contribute to the mechanics of cell locomotion, as it has been found that white blood cells can move in a liquid (i.e., "swim") in the absence of attachment to a substratum, presumably as surface structures operating like paddles move backward across the cell surface.

The Small GTP-Binding Proteins Cdc42, Rac, and Rho Control Actin Organization

A striking feature of a moving cell is its polarity: it has a front and a back. In addition, when a cell makes a turn, a new leading edge forms in the new direction of movement. To sustain movement in a particular direction, a cell requires signals to coordinate events at the front of the cell with events at the back and, indeed, signals to tell the cell where its front is. Our understanding of how this coordination occurs has emerged from studies of growth factors.

Growth factors, such as epidermal growth factor (EGF) and platelet-derived growth factor (PDGF), bind to specific cell-surface receptors (see Chapter 16) and consequently stimulate cells to move or to

divide. For example, in a wound, activated blood platelets secrete PDGF to attract fibroblasts and epithelial cells to enter the wound and repair it. It is possible to watch part of this process in vitro. If you grow cells in a culture dish and, after starving them of growth factors, you add some fresh growth factor, within a minute or two the cells respond by forming membrane ruffles. The growth factor activates cellular components that control recycling of membrane from endosomes coupled with actin assembly. In this regard, membrane ruffles are very similar to the lamellipodia of migrating cells.

Growth factors bind to specific receptors on the cell surface and induce a signal transduction pathway on the inner surface of the plasma membrane (see Chapter 15). The signal transduction pathway activates **Rac**, a member of the small GTPase superfamily of Ras-related proteins. Rac is one member of a GTPase superfamily that regulates microfilament organization; two others are **Cdc42** and **Rho**. (Due to the history of their discovery, the protein family of which Cdc42, Rac, and Rho are members has been collectively named "Rho proteins.") To understand how these proteins work, we first have to recall how small GTP-binding proteins function (see Chapter 15).

Like all small GTPases of the Ras superfamily, Cdc42, Rac, and Rho act as molecular switches: inactive in the GDP-bound state and active in the GTP-bound state (Figure 17-40). In their GDP-bound state, they exist free in the cytoplasm in an inactive form bound to a

protein known as *guanine nucleotide dissociation inhibitor (GDI)*. Growth factors can bind and activate their receptors to turn on specific membrane-bound regulatory proteins, called *guanine nucleotide exchange factors (GEFs)*, which activate Rho proteins at the membrane by releasing them from GDI and catalyzing the exchange of GDP for GTP. The GTP-bound active Rho protein associates with the plasma membrane, where it binds *effector proteins* to transmit the biological response. The small GTPase remains active until the GTP is hydrolyzed to GDP, a process that is stimulated by specific GTPase-activating proteins (GAPs).

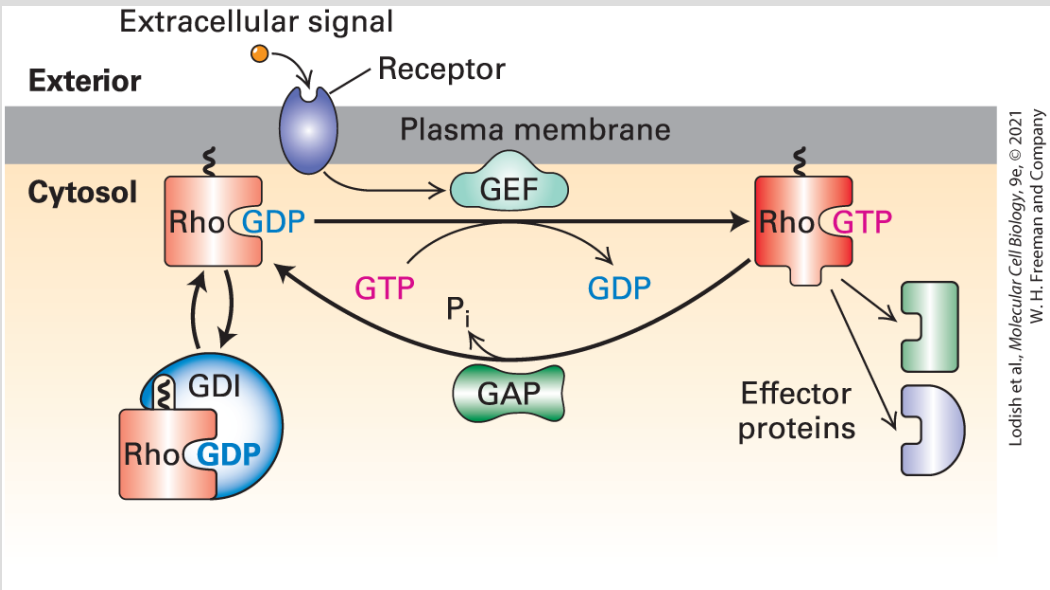


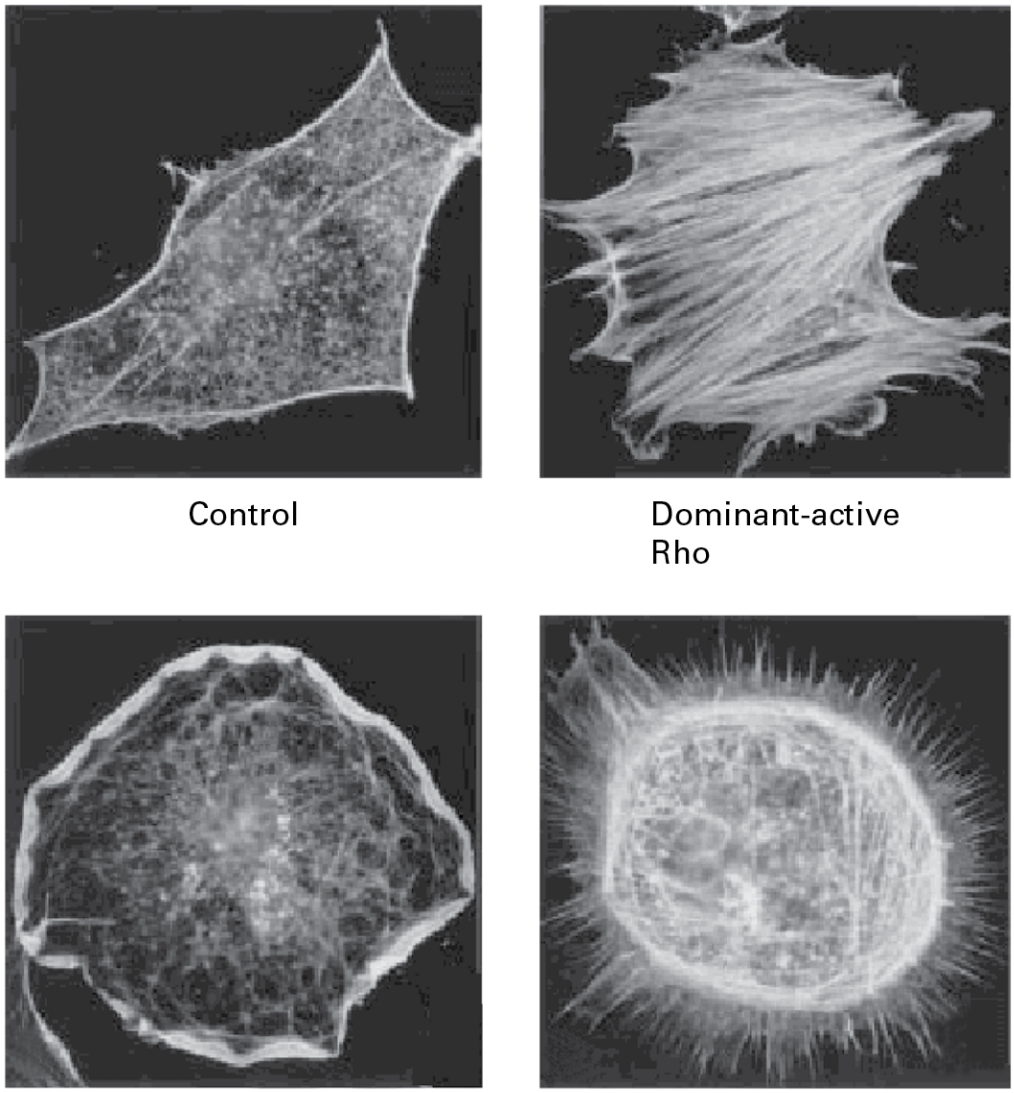
FIGURE 17-40 Regulation of the Rho family of small GTPases. The small GTPases of the Rho family are molecular switches regulated by accessory proteins. Rho proteins exist in the Rho-GDP bound form complexed with a protein known as *GDI (guanine nucleotide dissociation inhibitor)*, which keeps them in an inactive state in the cytosol. Membrane-bound signaling pathways bring Rho proteins to the membrane and, through the action of a GEF (guanine nucleotide exchange factor), exchange the GDP for GTP, thus activating them. Membrane-bound activated Rho-GTP can then bind effector proteins that

cause changes in the actin cytoskeleton. The Rho protein remains in the active Rho-GTP state until acted on by a GAP (GTPase-activating protein), which allows it to interact with the GDI and be returned to the cytoplasm. See S. Etienne-Manneville and A. Hall, 2002, *Nature* **420**:629.

To unravel the functions of the Rho proteins in cell migration, cells were modified to express mutant proteins that were locked either in the active Rho-GTP state or in the inactive Rho-GDP state. In general, a small GTPase that is locked in the active state is called a *dominant-active* protein. This always active GTPase will constitutively bind effector proteins (Figure 17-40), and one can then assess the biological outcomes. Alternatively, a *dominant-negative protein* can be introduced into a cell, such as a protein that binds to and inhibits the GEF protein. Introduction of a dominant-negative protein interferes with the signal transduction pathway, enabling one to assess what processes are blocked.

When experiments like these were performed in fibroblasts, Cdc42, Rac, and Rho were implicated in the regulation of microfilament organization because introduction of dominant-active mutant proteins had dramatic effects on the actin cytoskeleton, even in the absence of growth factors. Dominant-active Cdc42 resulted in the appearance of filopodia, dominant-active Rac resulted in the appearance of membrane ruffles, and dominant-active Rho resulted in the formation of stress fibers, which were seen to contract (Figure 17-41). How can one tell if dominant-active Rac and growth factor stimulation, both of which stimulate membrane ruffle formation,

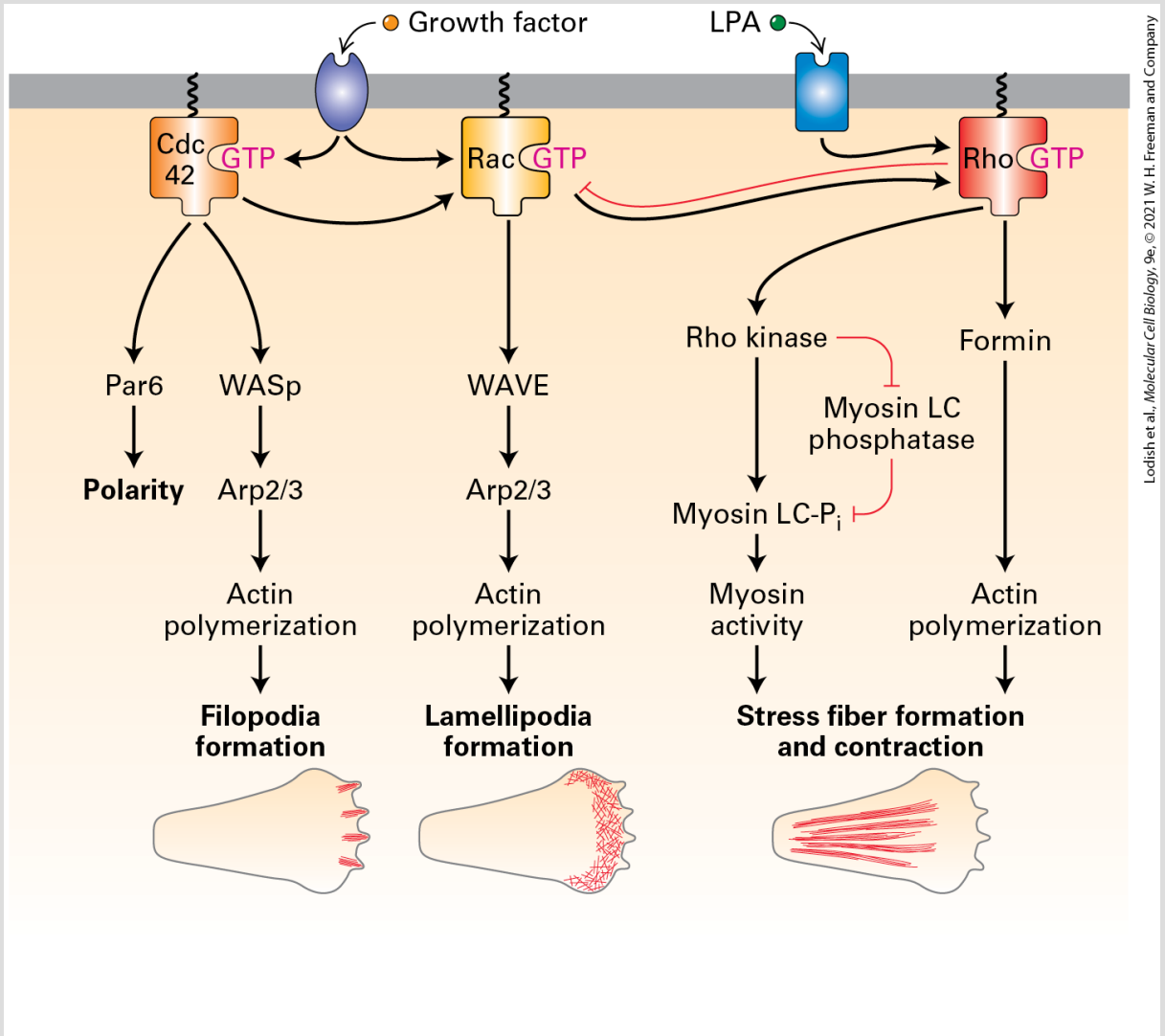
use the same signal transduction pathway? If growth factor stimulation leads to Rac activation, then introduction of a dominant-negative Rac protein into a cell should block the ability of a growth factor to induce membrane ruffling. This is precisely what was found. Using this and many other biochemical strategies, scientists have identified the signaling pathways involving Cdc42, Rac, and Rho (Figure 17-42).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 17-41 Dominant-active Rac, Rho, and Cdc42 induce different actin-containing structures. To look at the effects of constitutively active Rac, Rho, and Cdc42, growth-factor-starved fibroblasts were microinjected with plasmids to express dominant-active versions of the three proteins. The cells were then treated with fluorescent phalloidin, which stains filamentous actin. Dominant-active Rac induces the formation of peripheral membrane ruffles, whereas dominant-active Rho induces abundant contractile stress fibers, and dominant-active Cdc42 induces filopodia.

[Republished with permission from AAAS, from A. Hall, 1998, "Rho GTPases and the Actin Cytoskeleton," *Science* **279**(5350):509–514; permission conveyed through Copyright Clearance Center, Inc.]



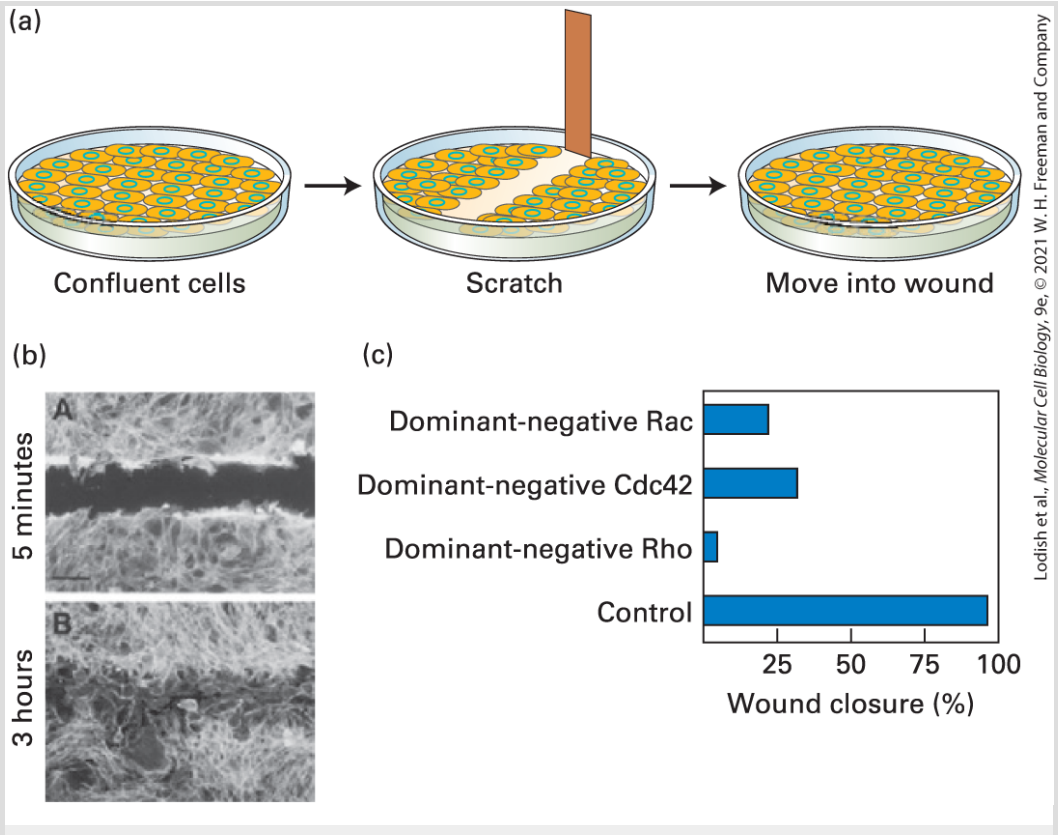
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 17-42 Summary of signal-induced changes in the actin cytoskeleton. Specific signals, such as growth factors and lysophosphatidic acid (LPA), are detected by cell-surface receptors. Detection leads to the activation of the small GTP-binding proteins, which then interact with effectors to bring about cytoskeletal changes as indicated.

Some of the pathways that these proteins regulate contain proteins we are familiar with. Activation of Cdc42 stimulates actin assembly by Arp2/3 through activation of WASp, a nucleation promoting factor (NPF) (see Figure 17-16), resulting in the formation of filopodia. Activation of Rac also induces Arp2/3, mediated by the WAVE complex, leading to the assembly of branched actin filaments in the leading edge. Activation of Rho has at least two effects. First, it can activate a formin for unbranched actin filament assembly. Second, through activation of Rho kinase, it can phosphorylate the myosin light chain to activate nonmuscle myosin II and can also inhibit light-chain dephosphorylation by phosphorylating myosin light-chain phosphatase to inhibit its activity. Both actions of Rho kinase lead to a higher level of myosin light-chain phosphorylation and therefore higher myosin activity and contraction. Cdc42, Rac, and Rho are also linked by activation and inhibition pathways, as shown in Figure 17-42.

Cell Migration Involves the Coordinate Regulation of Cdc42, Rac, and Rho

How does each of these small GTP-binding proteins contribute to cell migration? To answer this question, researchers developed the wounded-cell monolayer assay (Figure 17-43a). Cells in culture are grown in a petri dish with growth factors until they are confluent and form a tight monolayer, at which point they stop dividing. The cell monolayer is then scratched with a needle to remove a swath of cells, generating a wound containing a free edge of cells. The cells on the edge sense the loss of their neighbors and, in response to components of the extracellular matrix now exposed on the dish surface, move to fill up the empty wound area (Figure 17-43b). To do this, they orient themselves toward the empty area, first extending a lamellipodium and then moving in that direction. In this way, one can study the induction of directed cell migration in vitro.



EXPERIMENTAL FIGURE 17-43 The wounded-cell monolayer assay can be used to dissect signaling pathways in directed cell movement. (a) A confluent layer of cells is scratched to remove a swath about three cells wide to generate a free cell border. The remaining cells detect the free space and newly exposed extracellular matrix and, over a period of hours, fill in the area. (b) Localization of actin in a wounded monolayer 5 minutes and 3 hours after scratching; after 3 hours, cells have migrated into the wounded area. (c) Effect of introducing dominant-negative Cdc42, Rac, and Rho into cells at the wound edge; all affect wound closure.

[Part (b, c) C. D. Nobes and A. Hall, 1999, *J. Cell Biol.* **144**(6):1235–1244; https://doi.org/10.1083/jcb.144.6.1235.]

Using this wounded-cell monolayer assay, researchers have introduced dominant-negative Rac into cells on the wound edge to see how it affects the ability of the cells to migrate and fill the wound. Since Rac is needed for activation of the Arp2/3 complex to form the lamellipodium, it is not surprising that cells containing dominant-negative Rac fail to form this structure and do not migrate, and so the wound does not close (Figure 17-43c). A very interesting result is obtained when dominant-negative Cdc42 is introduced into the cells at the wound edge: they can form a lamellipodium, but they do not orient in the correct direction — in fact, they try to migrate in random directions. This observation suggests that Cdc42 is critical for regulating the overall polarity of the cell. Studies from yeast (in which Cdc42 was first described), wounded-cell monolayers, epithelial cells, and neurons reveal that Cdc42 is a master regulator of polarity in many different systems. Part of this regulation in animals involves the binding of Cdc42 to its

effector, Par6, a polarity protein that functions in nematodes (in which it was first discovered), neurons, and epithelial cells. We explore these polarity pathways in more detail in Chapter 22.

Studies such as these suggest a general model of how cell migration is controlled (Figure 17-44). Signals from the environment are transmitted to Cdc42, which orients the cell. The oriented cell has high Rac activity at the front, which induces the formation of the leading edge; Rho activity is high in the rear, inducing the assembly of contractile structures and activating the myosin-II-based contractile machinery. It is important to notice that different regions of the cell have different levels of active Cdc42, Rac, or Rho, so the amount of these regulators are controlled locally within the cell. Part of this spatial regulation occurs because some small G proteins can work antagonistically. For example, active Rho can stimulate pathways that lead to the inactivation of Rac. This process might help ensure that no leading-edge structures form at the rear of the cell.

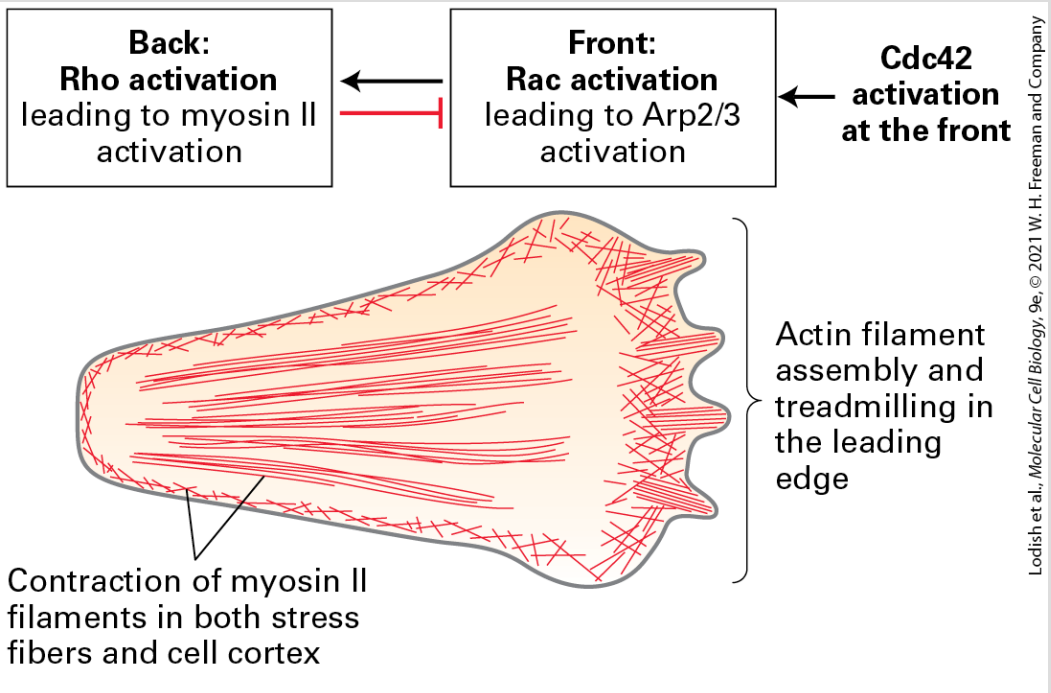
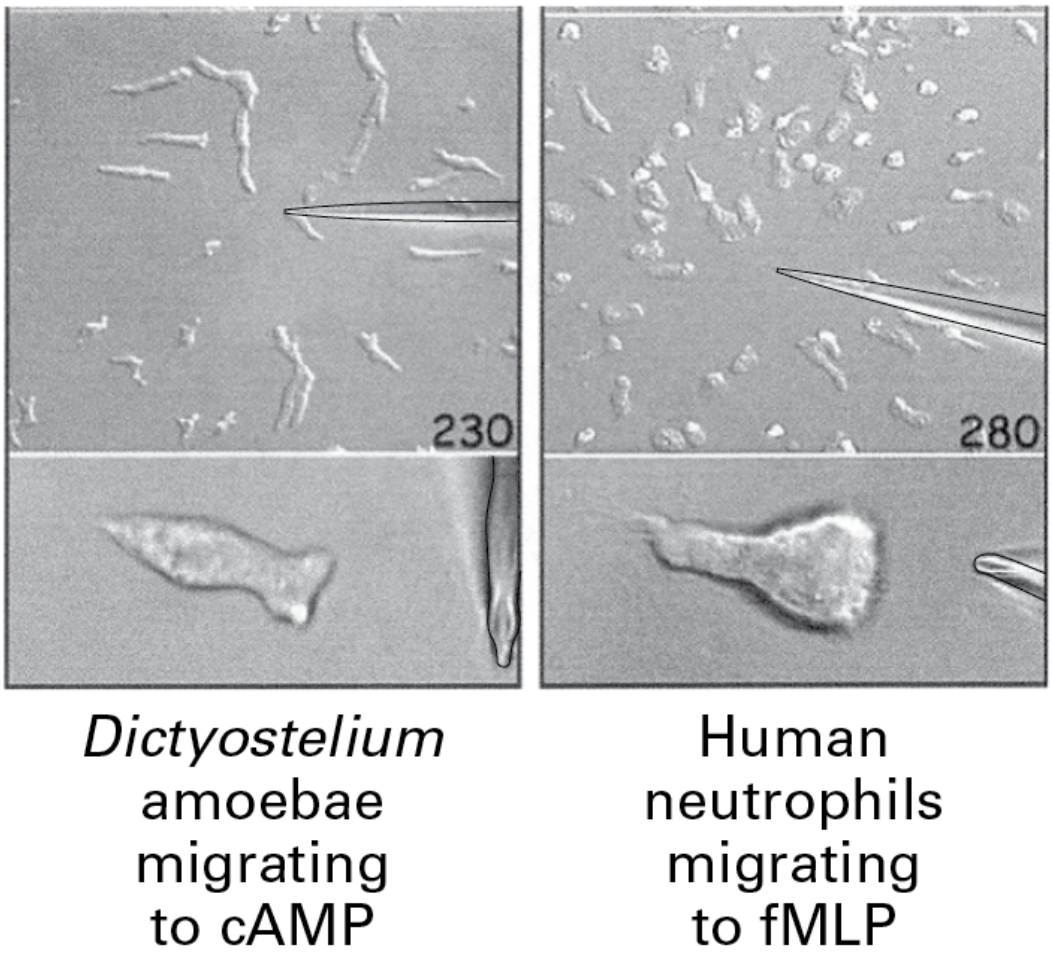


FIGURE 17-44 Contributions of Cdc42, Rac, and Rho to cell movement. The overall polarity of a migrating cell is controlled by Cdc42, which is activated at the front of a cell. Cdc42 activation leads to active Rac in the front of the cell, which generates the leading edge, and active Rho at the back of the cell, which causes myosin II activation and contraction. Active Rho inhibits Rac activation, ensuring the asymmetry of the two active G-proteins.

Migrating Cells Are Steered by Chemotactic Molecules

Under certain conditions, extracellular chemical cues guide the migration of a cell in a particular direction. In some cases, the movement is guided by insoluble molecules in the underlying substratum, as in the wounded-cell monolayer assay described above. In other cases, the cell senses soluble molecules and follows

them, along a concentration gradient, to their source — a process known as **chemotaxis**. For example, leukocytes are guided toward an infection by a tripeptide that is released by many bacterial cells (Figure 17-45). In another example, during the development of skeletal muscle, a secreted protein signal called *scatter factor* guides the migration of myoblasts to the proper locations in limb buds. One of the best studied examples of chemotaxis is the migration of *Dictyostelium* amoebae during their starvation response. When these soil amoebae are stressed, they begin to secrete cAMP, which is an extracellular chemotactic agent in this organism. Other *Dictyostelium* cells move up the cAMP concentration gradient toward its source (see Figure 17-45). Thus the amoebae move toward one another, aggregate into a migratory slug, and then differentiate into a fruiting body in which starvation-resistant spores are formed.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 17-45 Chemotactic molecules guide migrating cells by signaling the actin cytoskeleton. *Dictyostelium* cells migrate toward a pipette of cAMP (left), and human neutrophils (a type of leukocyte) migrate toward a pipette of fMLP (formylated Met-Leu-Phe), a chemotactic peptide produced by bacteria (*right*). In the lower two panels are individual chemotaxing *Dictyostelium* and neutrophil cells, which look remarkably similar, despite about 800 million years of evolution separating them.

[Republished with permission from Elsevier, from C. A. Parent, 2004, "Making All The Right Moves: Chemotaxis in Neutrophils and *Dictyostelium*," *Curr. Opin. Cell Biol.* **16**(1):4–13; permission conveyed through Copyright Clearance Center, Inc.]

Despite the variety of different chemotactic molecules — sugars, peptides, cell metabolites, cell-wall or membrane lipids — they all work through a common and familiar mechanism: binding to cell-surface receptors, activation of intracellular signaling pathways, and remodeling of the cytoskeleton through the activation or inhibition of various actin-binding proteins. What is quite amazing is that just a 2 percent difference in the concentration of chemotactic molecules between the front and back of the cell is sufficient to induce directed cell migration (see Classic Experiment 17-2). Equally amazing is the finding that the internal signal transduction pathways used in chemotaxis have been conserved between *Dictyostelium* amoebae and human leukocytes despite almost a billion years of evolution (Figure 17-45).

KEY CONCEPTS OF SECTION 17.7

Cell Migration: Mechanism, Signaling, and Chemotaxis

- Cell migration involves the extension of an actin-rich leading edge at the front of the cell, the formation of focal adhesions that move backward with respect to the cell, and their subsequent release, combined with rear contraction to push the cell forward (see Figure 17-38).
- Cell migration also involves a directed endocytic cycle, taking membrane and adhesion molecules from the rear of the cell and inserting them at the front.
- The assembly and function of actin filaments is controlled by signaling pathways through small GTP-binding proteins of the Rho family. Cdc42 regulates overall polarity and the formation of filopodia, Rac regulates actin network formation through the $\mathrm{Arp2/3}$ complex, and Rho regulates actin filament formation by formins as well as contraction through regulation of myosin II (see Figure 17-42).
- Chemotaxis, the directed movement of cells toward extracellular chemical cues, involves signaling pathways that contribute to the regulation of the actin cytoskeleton

and direction of cell migration.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Classic Experiment 17-1: Looking at Muscle Contraction
- Classic Experiment 17-2: Sensing Chemotactic Gradients
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

actin
actin cross-linking proteins
Arp2/3 complex
CapZ protein
Cdc42 protein
cell migration
cell polarity
chemotaxis
cofilin
contractile bundles
critical concentration (C_c)

cytoskeleton
duty ratio
filopodia
formin
intermediate filament
lamellipodium
leading edge
microfilament
microtubule
microvilli
motor protein
myofibrils
myosin
myosin light-chain kinase
nucleation promoting factor (NPF)
power stroke
profilin
Rac protein
Rho protein
sarcomere
sarcoplasmic reticulum (SR)
step size
stress fibers
thick filaments
thin filaments
thymosin-β_4
treadmilling
tropomodulin

tropomyosin
WASp

Review the Concepts

1. Three systems of cytoskeletal filaments exist in most eukaryotic cells. Compare them in terms of composition, function, and structure.
2. Actin filaments have a defined polarity. What is filament polarity? How is it generated at the subunit level? How is filament polarity detectable?
3. In cells, actin filaments form bundles or networks. How do cells form such structures, and what specifically determines whether actin filaments will form a bundle or a network?
4. Much of our understanding of actin assembly in the cell is derived from experiments using purified actin in vitro. What techniques can be used to study actin assembly in vitro? Explain how each of these techniques works. Which of these techniques would tell you whether the mass of actin filaments is made up of many short actin filaments or fewer longer filaments?
5. The predominant forms of actin inside a cell are ATP–G-actin and ADP–F-actin. Explain how the interconversion of the nucleotide state is coupled to the assembly and disassembly of actin subunits. What would be the consequence for actin filament assembly/disassembly if a

mutation prevented actin's ability to bind ATP? What would be the consequence if a mutation prevented actin's ability to hydrolyze ATP?

6. Actin filaments at the leading edge of a crawling cell are believed to undergo treadmilling. What is treadmilling, and what accounts for this assembly behavior?
7. Although purified actin can assemble reversibly in vitro, various actin-binding proteins regulate the assembly of actin filaments in the cell. Predict the effect on a cell's actin cytoskeleton if function-blocking antibodies against each of the following were independently microinjected into cells: profilin, thymosin-β_4, CapZ, and the Arp2/3 complex.
8. Predict how actin would polymerize on a myosin-decorated short actin filament in the presence of CapZ, tropomodulin, or profilin-actin.
9. Compare and contrast the ways in which formin and WASp are activated and explain how each stimulates actin filament formation.
10. There are at least 20 different types of myosin. What properties do all types share, and what makes them different? Why is myosin II the only myosin capable of producing contractile force?
11. The ability of myosin to walk along an actin filament may be observed with the aid of an appropriately equipped microscope. Describe how such assays are typically performed. Why is ATP required in these assays? How can such assays be used to determine the direction of myosin movement or the force produced by myosin?

12. Contractile bundles occur in nonmuscle cells; these structures are less organized than the sarcomeres of muscle cells. What is the purpose of nonmuscle contractile bundles? Which type of myosin is found in contractile bundles?
13. How does myosin convert the chemical energy released by ATP hydrolysis into mechanical work?
14. Myosin II has a duty ratio of 10 percent, and its step size is 8 nm. In contrast, myosin V has a much higher duty ratio (about 70 percent) and takes 36-nm steps as it walks down an actin filament. What differences between myosin II and myosin V account for their different properties? How do the different structures and properties of myosin II and myosin V reflect their different functions in cells?
15. Contraction of both skeletal and smooth muscle is triggered by an increase in cytosolic Ca^{2+}. Compare the mechanisms by which each type of muscle converts a rise in Ca^{2+} concentration into contraction.
16. Phosphorylation of myosin light-chain kinase (MLC kinase) by protein kinase A (PKA) inhibits MLC kinase activation by Ca^{2+}/calmodulin. Drugs such as albuterol bind to the β-adrenergic receptor, which causes a rise in cAMP in cells and activation of PKA. Explain why albuterol is useful for treating the severe contraction of the smooth muscle cells surrounding airway passages involved in an asthma attack.
17. Several types of cells use the actin cytoskeleton to power their locomotion across surfaces. How are different assemblies of actin filaments involved in locomotion?

18. To move in a specific direction, a migrating cell must use extracellular cues to establish which portion of the cell will act as the front and which will act as the back. Describe how small GTPase proteins appear to be involved in the signaling pathways used by migrating cells to determine direction of movement.
19. Cell motility has been described as being like the motion of tank treads. At the leading edge, actin filaments form rapidly into bundles and networks that make protrusions and move the cell forward. At the rear, cell adhesions are broken and the tail end of the cell is brought forward. What provides the traction for moving cells? How does cell-body translocation happen? How are cell adhesions released as cells move forward?

CHAPTER 18

Cell Organization and Movement II: Microtubules and Intermediate Filaments

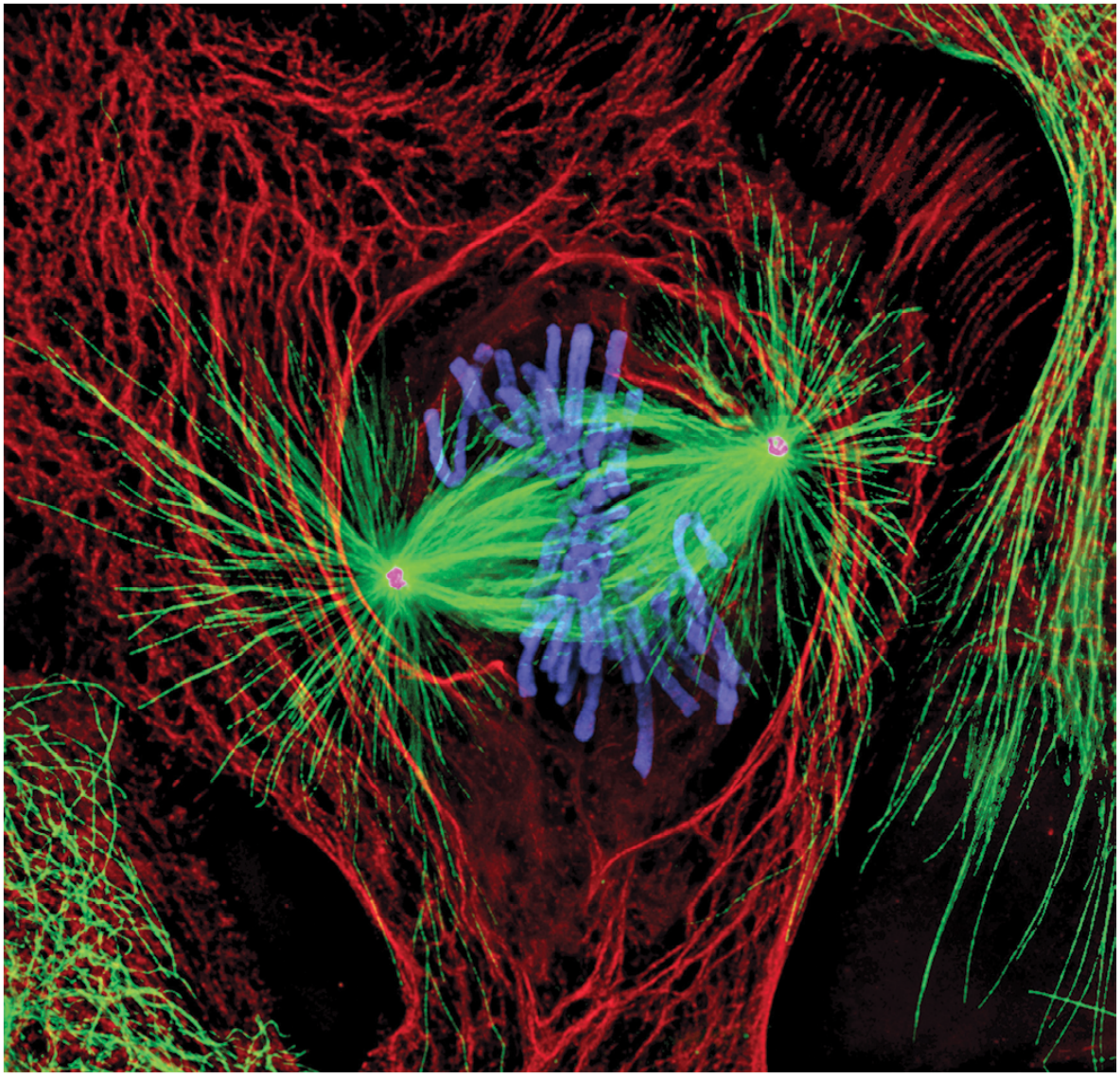


photo courtesy Alexey Khodjakov.

Newt lung cell in mitosis stained for centrosomes (magenta), microtubules (green), chromosomes (blue), and keratin intermediate filaments (red).

[Reprinted by permission of Nature Publishing Group, from A. Khodjakov, "Olympus/Nature competition: A 1, 2, 3 in Light Microscopy," *Nature*, 2000, **408**:423–424; permission conveyed through Copyright Clearance Center, Inc.]

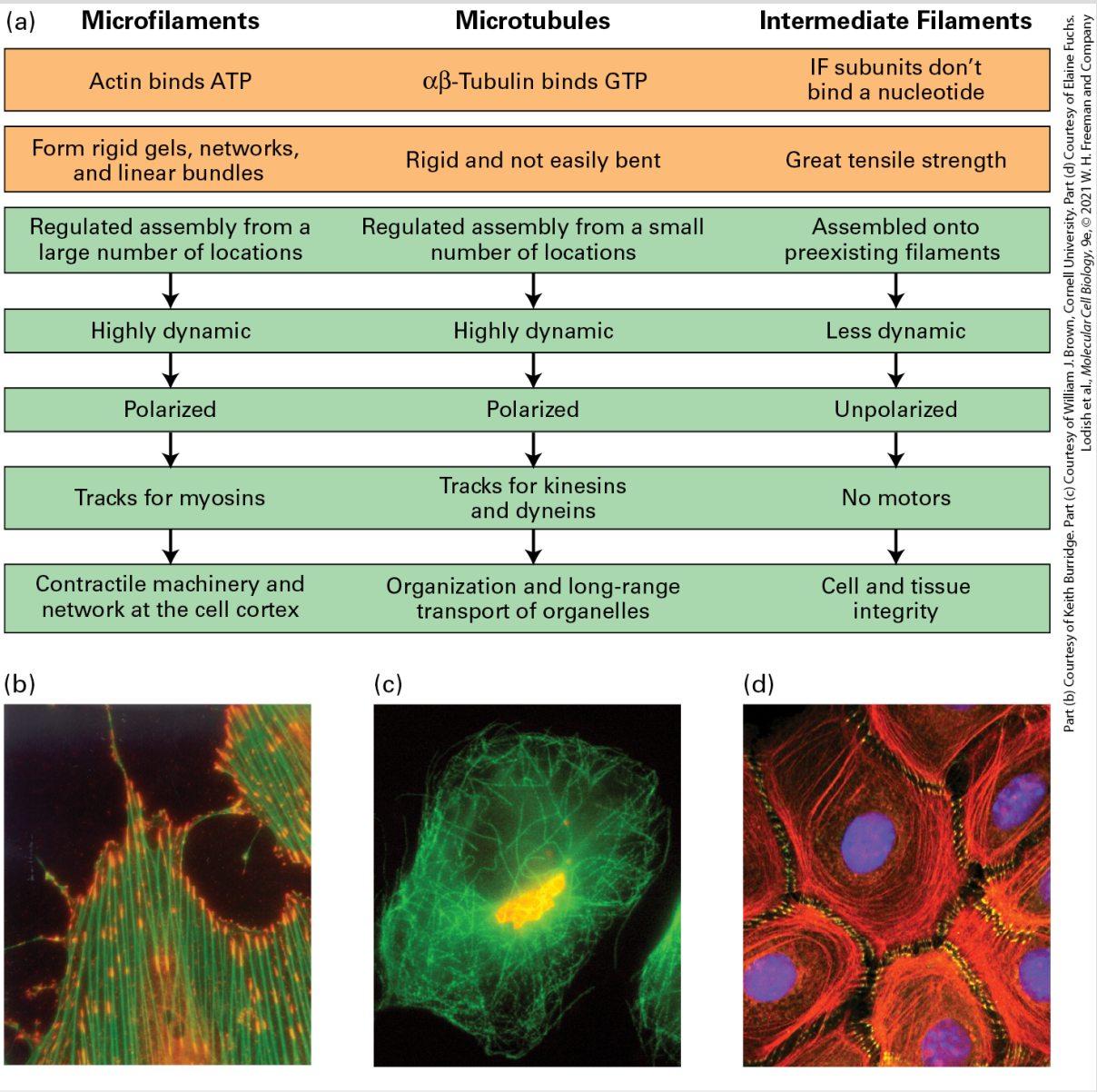
OUTLINE

18.1 Microtubule Structure and Organization
18.2 Microtubule Dynamics
18.3 Regulation of Microtubule Structure and Dynamics
18.4 Kinesins and Dyneins: Microtubule-Based Motor Proteins
18.5 Cilia and Flagella: Microtubule-Based Surface Structures
18.6 Mitosis
18.7 Intermediate Filaments
18.8 Coordination and Cooperation Between Cytoskeletal Elements

As we learned in the previous chapter, three types of filaments make up the cytoskeleton in an animal cell: microfilaments, microtubules, and intermediate filaments. Why have these three distinct types of filaments evolved? It seems likely that their physical properties are suited to different functions. In Chapter 17, we described how actin microfilaments are often cross-linked into networks or bundles to form mechanically strong yet flexible and dynamic structures and to serve as tracks for the many different classes of myosin motors. Similarly, **microtubules** are stiff tubes that can exist as a single structure, extending up to 20 μm in cells, or in bundled arrangements such as those seen in specialized cell-surface structures like cilia and flagella. As a consequence of their

tubular design, microtubules are able to generate pulling and pushing forces without buckling, a property that allows single tubules to extend long distances within a cell and bundles to slide past each other, as occurs in flagella and in the mitotic spindle. Microtubules' ability to extend long distances in the cell, together with their intrinsic polarity due to the same orientation of subunits along a microtubule, is exploited by microtubule-dependent motors, which use them as tracks for long-range transport of organelles. Microtubules can be highly dynamic — assembling and disassembling rapidly from their ends — and can thus provide the cell with the flexibility to alter microtubule organization as needed.

In contrast to microfilaments and microtubules, **intermediate filaments** have great tensile strength and have evolved to withstand much larger stresses and strains. With properties akin to strong molecular ropes, they are ideally suited to endow both cells and tissues with structural integrity and to contribute to cellular organization. Intermediate filaments do not have an intrinsic polarity as microfilaments and microtubules do, so it is not surprising that there are no motor proteins that use intermediate filaments as tracks. Although we discuss microtubules and intermediate filaments together in this chapter — and although their localization in the cytosol can look superficially quite similar — we will see that their dynamics and functions are very different. A summary of the similarities and differences among these three filament systems is shown in Figure 18-1.



Part (b) Courtesy of Keith Burridge. Part (c) Courtesy of William J. Brown, Cornell University. Part (d) Courtesy of Elaine Fuchs.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-1 Overview of the physical properties and functions of the three filament systems in animal cells. (a) Biophysical and biochemical properties (orange) and biological properties (green) are shown for each filament type. The micrographs (b–d) show examples of each filament type in a particular cellular context, but note that microtubules also make up other structures, and that intermediate filaments also line the inner surface of the nucleus. (b) Cultured cells stained for actin (green) and sites of actin attachment to the substratum (orange). (c) Localization of microtubules (green) and the Golgi complex (yellow). Notice the central location of the Golgi complex, which is collected there by transport along microtubules. (d) Localization of cytokeratins (red), a type of intermediate

filament, and a component of desmosomes (yellow) in epithelial cells. Cytokeratins from individual cells are attached to each other through the desmosomes.

This chapter covers five main topics. First, we discuss the structure and dynamics of microtubules and their motor proteins. Second, we examine how microtubules and their motors contribute to the movement of cilia and flagella. Third, we discuss the role of microtubules in the mitotic spindle — a molecular machine that has evolved to accurately segregate duplicated chromosomes. Fourth, we explore the roles of the different classes of intermediate filaments that provide structure to the nuclear envelope as well as strength and organization to cells and tissues. Although we consider microtubules, microfilaments, and intermediate filaments individually, the three filament systems do not act independently of one another, and we consider some examples of this interdependence in the last section of the chapter.

18.1 Microtubule Structure and Organization

In the early days of electron microscopy, cell biologists noted long tubules in the cytoplasm that they called *microtubules*. Morphologically similar microtubules were seen making up the fibers of the mitotic spindle, as components of axons, and as the structural elements in cilia and flagella (Figure 18-2a, b). Careful examination of single microtubules from various sources in transverse section indicated that they were all made up of 13 longitudinal repeating units (Figure 18-2c), now called *protofilaments*, suggesting that these various microtubules shared a common structure. Microtubules purified from brain were found to consist of a major protein, **tubulin**, together with associated proteins. Purified tubulin alone can assemble into a microtubule under favorable conditions, proving that tubulin is the structural component of the microtubule. Associated proteins help mediate the assembly, dynamics, and function of microtubules. In this section, we consider the general structure and organization of microtubules before turning to a more detailed discussion of their dynamics and regulation in Sections 18.2 and 18.3.

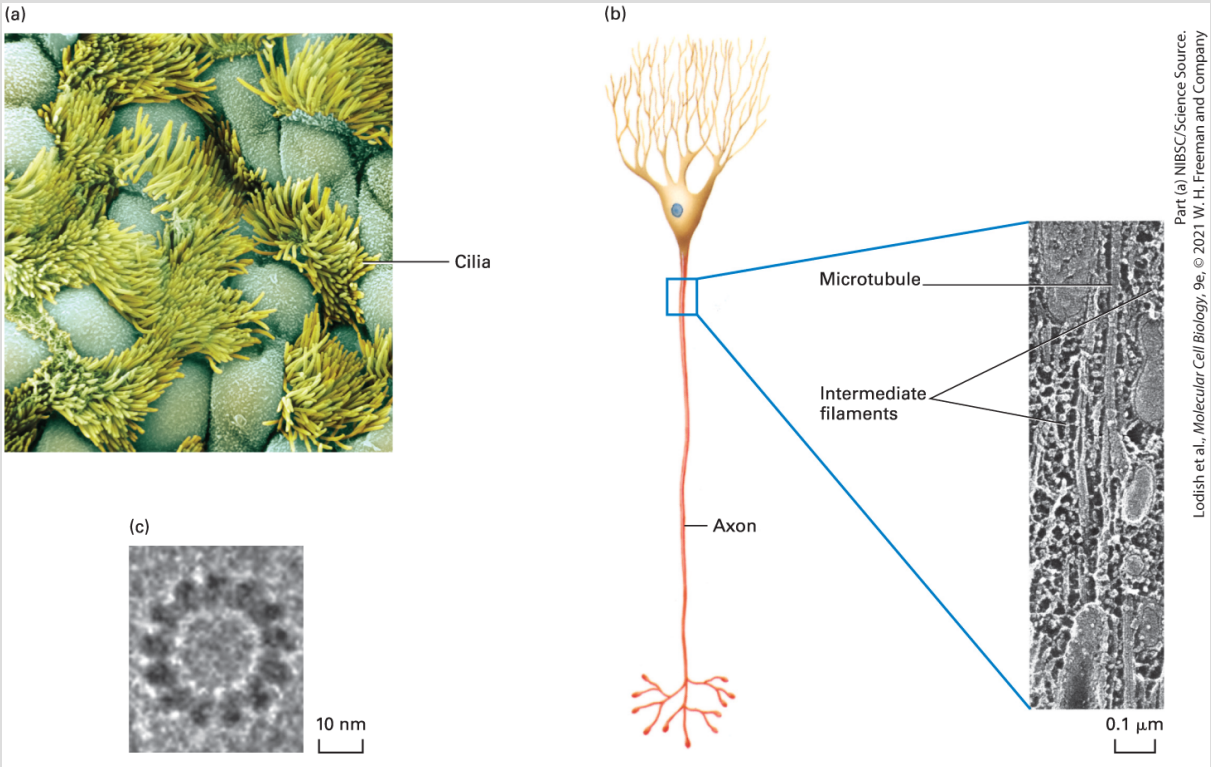


FIGURE 18-2 Microtubules are found in many different locations, and all have similar structures. (a) Surface of the ciliated epithelium lining a rabbit oviduct viewed in a scanning electron microscope. Beating cilia, each of which has a core of microtubules, propel eggs down the oviduct. (b) Microtubules and intermediate filaments in a quick-frozen and deep-etched frog axon visualized in a transmission electron microscope. (c) High-magnification view of a single microtubule showing the 13 repeating units known as *protofilaments*.

[Part (b) N. Hirokawa, 1982, *J. Cell Biol.* **94**(1): 129–142; https://doi.org/10.1083/jcb.94.1.129. Part (c) Republished with permission from John Wiley and Sons, Inc., from H. Sosa and D. Chrétien, 1998, "Relationship Between Moiré Patterns, Tubulin Shape, and Microtubule Polarity," *Cytoskeleton* **40**(1):38–43; permission conveyed through Copyright Clearance Center, Inc.]

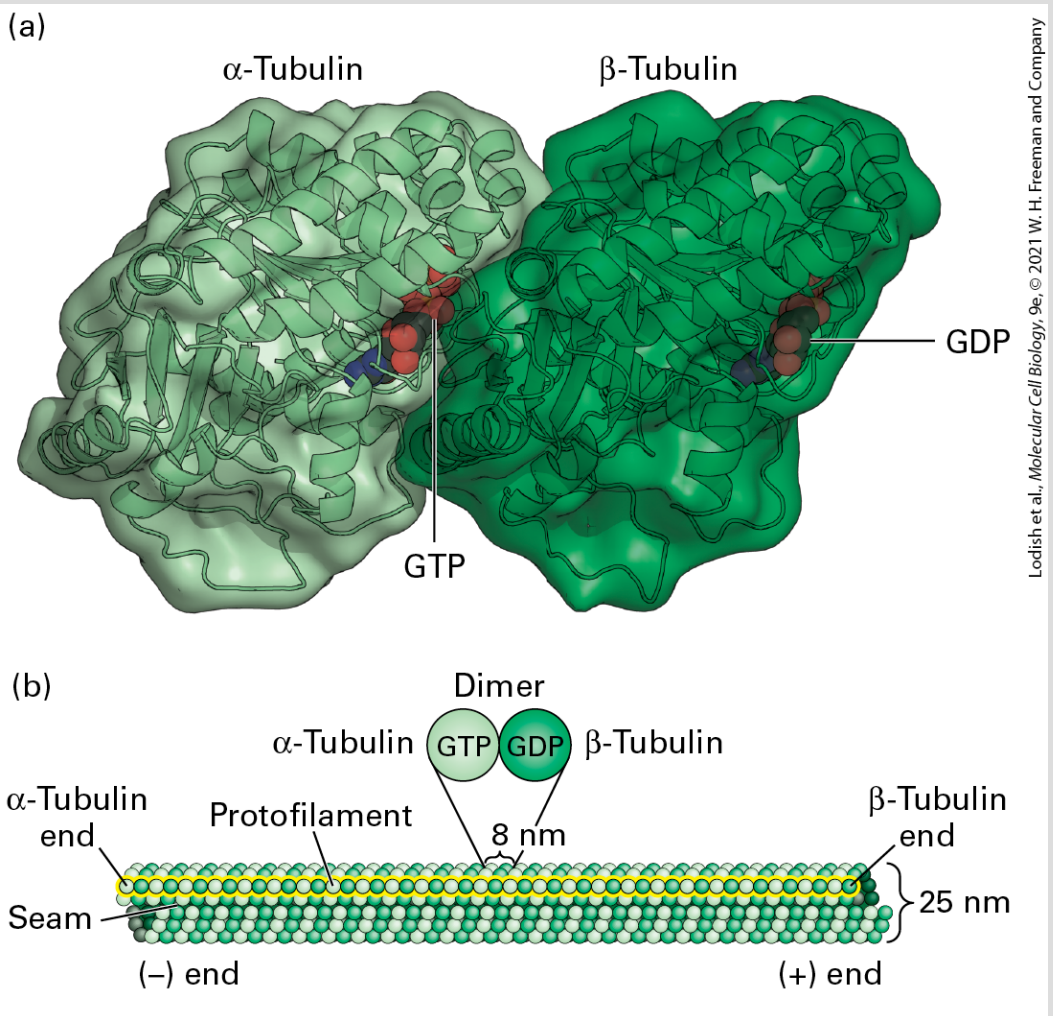
Microtubule Walls Are Polarized Structures Built from αβ-Tubulin

Dimers

Tubulin isolated in a pure and soluble form consists of two closely related subunits called *α-* and *β-tubulin*, each with a molecular weight of about 55,000 Da. The building block of a microtubule is an αβ-tubulin dimer. Genomic analyses reveal that genes encoding α- and β-tubulins are present in all eukaryotes and that the number of genes has expanded considerably in multicellular organisms. For example, budding yeast has two genes specifying α-tubulin and one for β-tubulin, whereas the soil nematode *Caenorhabditis elegans* has nine genes encoding α-tubulin and six for β-tubulin. In addition to α- and β-tubulin, all eukaryotes also have genes specifying a third tubulin, γ-tubulin, which is involved in microtubule assembly, as we will see shortly. Additional isoforms of tubulin have also been discovered that are present only in organisms that possess cellular structures called centrioles and basal bodies, suggesting that these tubulin isoforms are important for those structures. As we'll learn in this chapter, centrioles and basal bodies are specialized structures that nucleate and organize microtubule assembly.

The α- and β-subunits of the tubulin dimer can each bind one molecule of GTP (Figure 18-3a). The GTP in the α-tubulin subunit is never hydrolyzed and is trapped by the interface between the α- and β-subunits. By contrast, the GTP-binding site on the β-subunit is at the surface of the dimer. GTP bound by the β-subunit can be hydrolyzed, and the resulting GDP can be exchanged for free GTP.

Under appropriate conditions, soluble tubulin dimers can assemble into microtubules (Figure 18-3b). As we saw in Chapter 17 for the polymerization of actin, ATP–G-actin is preferentially added to one end of the filament, designated the (+) end because it is the end favored for assembly. Once incorporated into the filament, the bound ATP is hydrolyzed to ADP and P_i. In a similar manner, tubulin dimers in which the β-subunit has bound GTP are added preferentially to one end of the microtubule, also designated the (+) end. As we will see, the GTP is hydrolyzed once tubulin is incorporated into the microtubule and this has dramatic effects on the behavior of the microtubule (+) end.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-3 Structure of tubulin dimers and their organization into microtubules. (a) Ribbon diagram of the tubulin dimer. The GTP bound to the α-tubulin monomer is nonexchangeable, whereas the GDP bound to the β-tubulin monomer is exchangeable with free GTP. (b) The organization of tubulin subunits in a microtubule. The dimers are aligned end to end into protofilaments, which pack side by side to form the wall of the microtubule. The protofilaments are slightly staggered so that α-tubulin in one protofilament is in contact with α-tubulin in the neighboring protofilaments, except at the seam, where an α-subunit contacts a β-subunit. The microtubule displays a structural polarity in that subunits are added preferentially at the end where β-tubulin monomers are exposed. This end of the microtubule is known as the (+) *end*.

[Part (a) Data from E. Nogales et al., 1998, *Nature* **391**:199, PDB 1D 1tub.]

A microtubule is composed of 13 laterally associated protofilaments, which form a tubule whose external diameter is about 25 nm (see Figure 18-3b). Each of the 13 protofilaments is a string of αβ-tubulin dimers, longitudinally arranged so that the subunits alternate along a protofilament, with each subunit type repeating every 8 nm. Because the αβ-tubulin dimers in a protofilament are all oriented in the same way, each protofilament has an α-subunit at one end and a β-subunit at the other — thus the protofilaments have an intrinsic **polarity**. In a microtubule, all the laterally associated protofilaments have the same polarity, so the microtubule also has an overall polarity. The end with exposed β-subunits is the (+) end, while the end with exposed α-subunits is the (−) end. In microtubules, the heterodimers in adjacent protofilaments are staggered slightly, forming tilted rows of α- and β-tubulin monomers in the microtubule wall. If you follow a row of β-subunits, for example, spiraling around a microtubule for one full turn, you will end up precisely three subunits up the protofilament, abutting an α-subunit. Thus all microtubules have a single longitudinal *seam*, where an α-subunit in one protofilament meets a β-subunit in the adjacent protofilament.

Most microtubules in a cell consist of a simple tube, a *singlet* microtubule, built from 13 protofilaments. In rare cases, singlet microtubules contain more or fewer protofilaments. For example, certain microtubules in the neurons of nematode worms contain 11 or 15 protofilaments. In addition to this simple singlet structure,

doublet or *triplet* microtubules are found in specialized structures such as cilia and flagella (doublet microtubules) and centrioles and basal bodies (triplet microtubules), structures we will explore later in the chapter. Each doublet or triplet contains one complete 13-protofilament microtubule (called the *A tubule*) and one or two additional tubules (B and C) consisting of 10 protofilaments each (Figure 18-4).

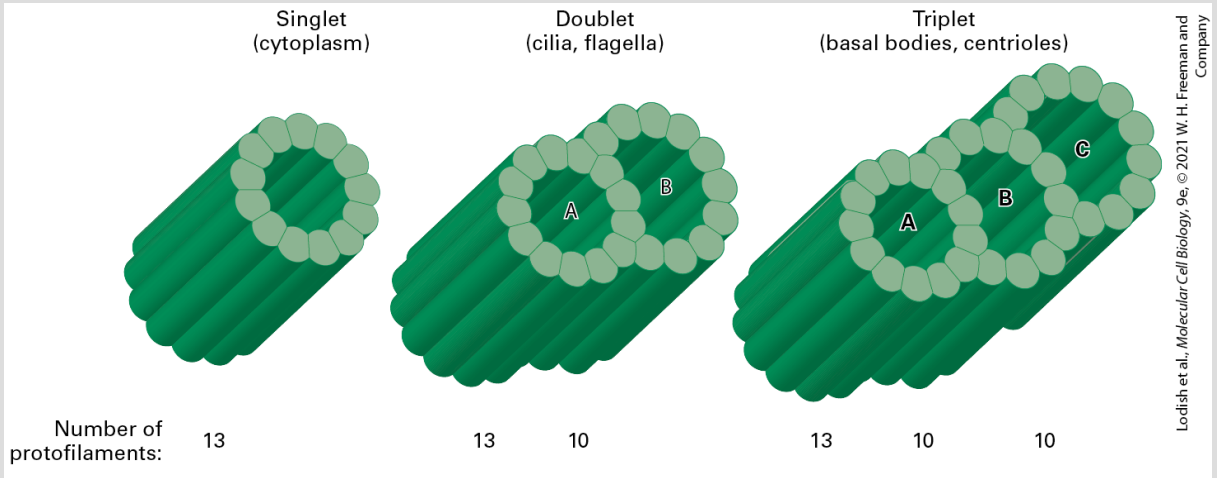


FIGURE 18-4 Singlet, doublet, and triplet microtubules. In cross section, a typical microtubule, a singlet, is a simple tube built from 13 protofilaments. In a doublet microtubule, an additional set of 10 protofilaments forms a second tubule (B) by fusing to the wall of a singlet (A) microtubule. Attaching another 10 protofilaments to the (B) tubule of a doublet microtubule creates a (C) tubule and a triplet structure.

Microtubules Are Assembled from MTOCs to Generate Diverse Configurations

Once tubulin had been identified as the major structural component of microtubules, antibodies to tubulin could be generated and used in immunofluorescence microscopy to localize microtubules in cells (Figure 18-5a, b). This approach, coupled with electron microscopy, revealed that microtubules are assembled from specific sites in the cell.

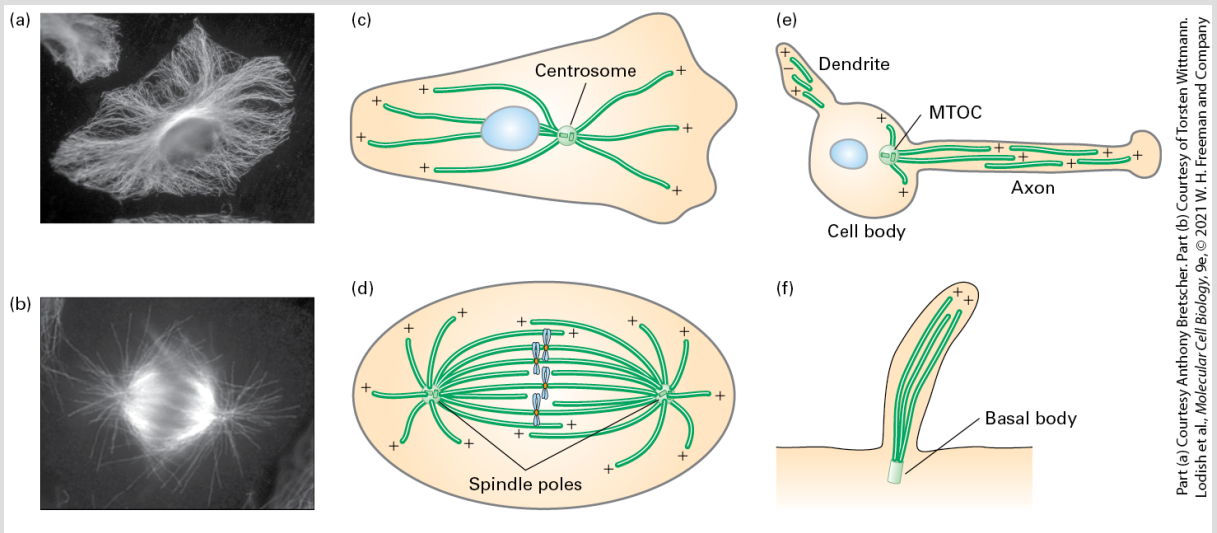


FIGURE 18-5 Microtubules are assembled from microtubule organizing centers (MTOCs). (a–b) The distribution of microtubules in cultured cells, as seen by immunofluorescence microscopy using antibodies to tubulin, in an interphase cell (a) and a cell in mitosis (b). (c–f) Diagrams of the distributions of microtubules in various cells and structures. All of these microtubules are assembled from distinct MTOCs. (c) In an interphase cell, the MTOC is called a centrosome (the nucleus is indicated by a blue oval). (d) In a mitotic cell, the two MTOCs are called *spindle poles* (the chromosomes are shown in blue). (e) In a neuron, microtubules in both axons and dendrites are assembled from an MTOC in the cell body and then released from it. (f) The microtubules that make up the shaft of a cilium or flagellum are assembled from an MTOC known as a basal body. The polarity of microtubules is indicted by (+) and (−).

Spontaneous nucleation of microtubules is such an energetically unfavorable reaction that it does not play a significant role in microtubule assembly in vivo. Rather, all microtubules are nucleated from structures known as **microtubule-organizing centers, or MTOCs**.

The **centrosome** is the main MTOC in animal cells. During interphase, the centrosome is generally located near the nucleus, producing an array of microtubules with their (+) ends radiating toward the cell periphery (Figure 18-5c). In most cases, the (−) end of the microtubule stays anchored in the centrosome. This radial array provides tracks for microtubule-based motor proteins to organize and transport membrane-bounded compartments, such as those constituting the secretory and endocytic pathways. During mitosis, cells reorganize their microtubules to form a bipolar spindle extending from two centrosomes, also known as *spindle poles*, that can accurately segregate copies of the duplicated chromosomes (Figure 18-5d). Neurons, in another example, have long processes called *axons*, in which organelles are transported in both directions along microtubules (Figure 18-5e). The microtubules in axons, which can be as long as 1 m, are not continuous and have been released from the MTOC, but nevertheless all have the same polarity. The microtubules in shorter processes, called *dendrites*, have mixed polarity, although the functional significance of this difference is not clear. In cilia and flagella (Figure 18-5f), microtubules are assembled from an MTOC called a **basal body**. As

we mention later, plants do not have centrosomes and basal bodies; they use other mechanisms to nucleate the assembly of microtubules.

Electron microscopy shows that each centrosome in an animal cell consists of a pair of orthogonally arranged cylindrical **centrioles** surrounded by amorphous material called **pericentriolar material** (Figure 18-6a). Centrioles, which are about 0.5 μm long and 0.2 μm in diameter, are highly organized and stable structures that consist of nine triplet microtubules (Figure 18-6b, c). The centrioles themselves do not nucleate the cytoplasmic microtubule array. Instead, factors in the pericentriolar material do. A critical component is the *γ-tubulin ring complex (γ-TuRC)* (Figure 18-6b and Figure 18-7), which is located in the pericentriolar material and consists of many copies of γ-tubulin associated with several other proteins. The γ-TuRC acts as a helical template to bind αβ-tubulin dimers for the formation of a new microtubule, whose (−) end is associated with γ-TuRC and whose (+) end is free for further assembly.

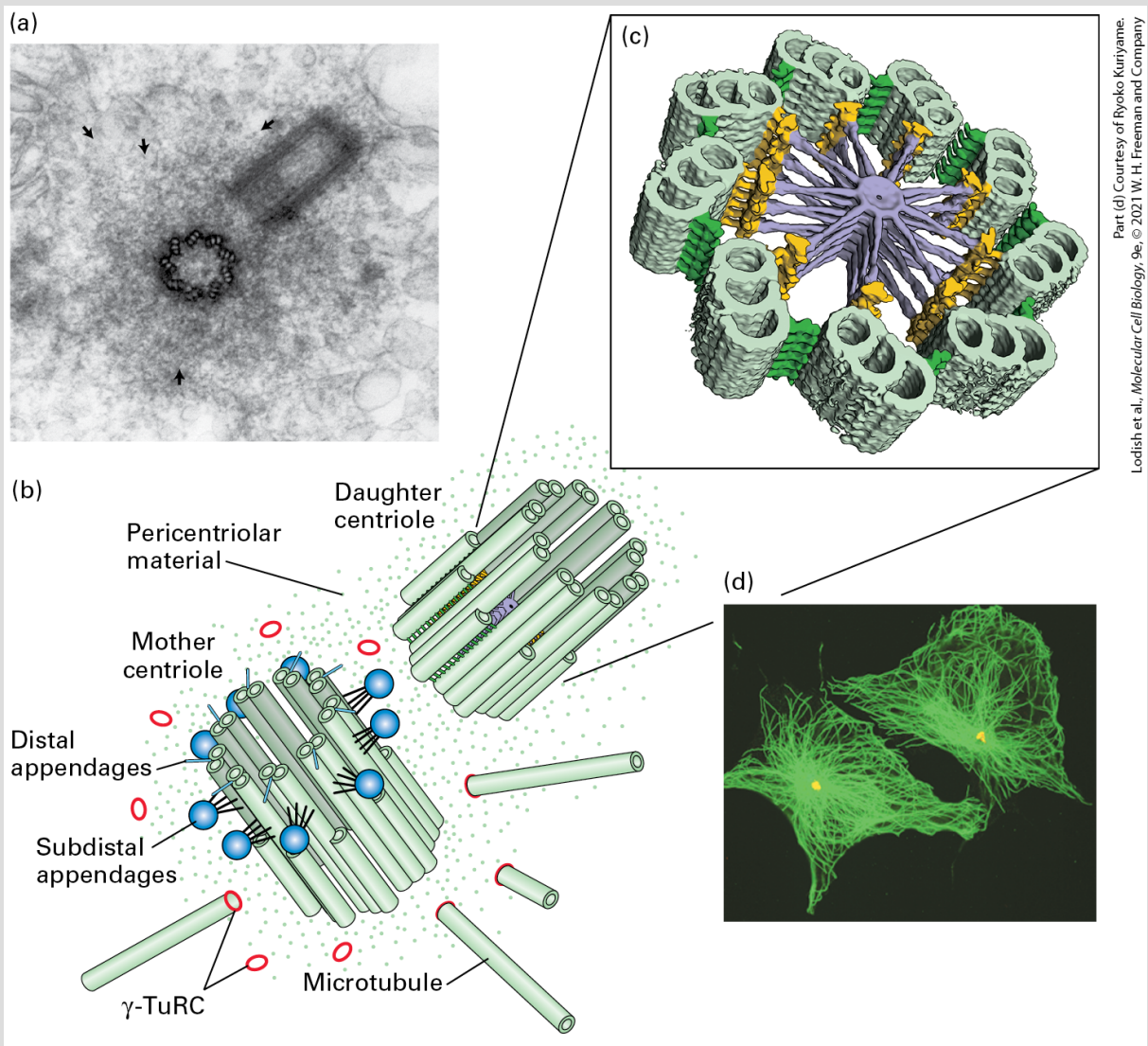


FIGURE 18-6 Structure of centrosomes. (a) Thin section of an animal-cell centrosome showing the two centrioles at right angles to each other, surrounded by pericentriolar material (arrows). (b) Diagram of a centrosome showing the mother and daughter centrioles, each of which consists of nine linked triplet microtubules, embedded in pericentriolar material that contains γ-TuRC nucleating structures. The mother centriole is distinct from the daughter as it has distal appendages (blue spheres). (c) A tomographic image of a section through a daughter centriole of the alga *Chlamydomonas*. A protein called *Sas6* can self-assemble into a ninefold symmetrical cartwheel structure to serve as a template for the daughter centriole (purple structure). It is later removed as the centriole matures. (d) Immunofluorescence micrograph showing the microtubule array (green) in a cultured animal cell and the location of the MTOC, using an antibody to a centrosomal protein (yellow).

[Part (a) Reprinted by permission from Nature Publishing Group, from G. Sluder, 2005, "Two-Way Traffic: Centrosomes and the Cell Cycle," *Nat. Rev. Mol. Cell Biol.* **6**:743–748; permission conveyed through Copyright Clearance Center, Inc. Part (c) Data from P. Guichard et al., 2013, *Curr. Biol.* **23**:1620, EMD-2329 and EMD-2330.]

Basal bodies have a structure similar to centrioles and are the MTOCs found at the base of cilia and flagella. The A and B tubules of their triplet microtubules provide a template for the assembly of the microtubules making up the core structure of cilia and flagella. Later, we discuss the structure of basal bodies in the context of cilia and flagella.

Recent work has uncovered an additional mechanism for the nucleation of microtubules in animal cells, also involving γ-TuRC. A protein complex called the *augmin complex*, consisting of eight polypeptides, can bind to the sides of existing microtubules, then recruit γ-TuRC and nucleate the assembly of new ones (Figure 18-7c). As we discuss in a later section, the augmin complex contributes to microtubule assembly in the mitotic spindle.

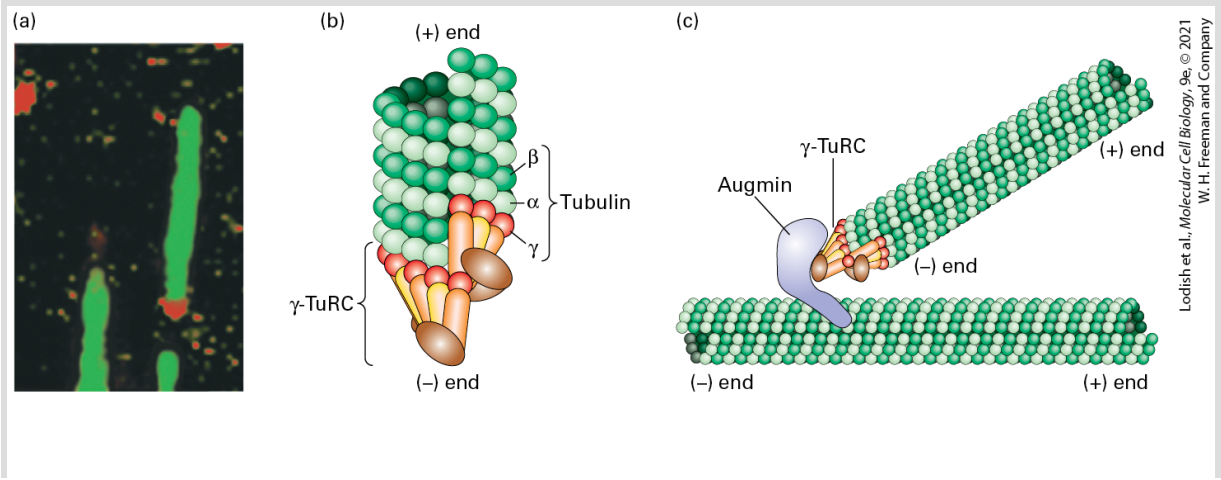


FIGURE 18-7 The γ-tubulin ring complex (*γ*-TuRC) that nucleates microtubule assembly. (a) Immunofluorescence micrograph in which microtubules assembled in vitro are labeled green and a component of the γ-TuRC is labeled red, showing that it is located specifically at one end of the microtubule. (b) Model of how γ-TuRC nucleates assembly of a microtubule by forming a template corresponding to the (−) end. (c) Diagram of how augmin can bind both γ-TuRC and the side of an existing microtubule to nucleate assembly of a branched microtubule.

[Part (a) Reprinted with permission from Nature Publishing Group, from T. J. Keating and G. G. Borisy, 2000, "Immunostructural Evidence for the Template Mechanism of Microtubule Nucleation," *Nat. Cell Biol.* **2**:352–357; permission conveyed through Copyright Clearance Center, Inc.]

KEY CONCEPTS OF SECTION 18.1

Microtubule Structure and Organization

- Tubulin is the major structural component of microtubules (see Figure 18-3). Microtubule-associated proteins (MAPs) associate with tubulin and help mediate the assembly, dynamics, and function of microtubules.
- Free tubulin exists as a dimer in which the α-subunit contains a trapped and nonhydrolyzable GTP and the β-subunit binds an exchangeable and hydrolyzable GTP.
- αβ-Tubulin assembles into microtubules, each of which is made up of 13 laterally associated protofilaments, with an α-subunit exposed at the (−) end and a β-subunit at the (+) end of each protofilament.
- In cilia, flagella, centrioles, and basal bodies, doublet and triplet microtubules exist in which the additional microtubules have 10 protofilaments (see Figure 18-4).
- All microtubules are nucleated from microtubule-organizing centers (MTOCs), and many remain anchored with their (−) ends there. Thus the end away from the MTOC is always the (+) end.
- The centrosome is the MTOC that nucleates the radial array of microtubules in interphase animal cells. Two centrosomes, or spindle poles, are the MTOCs that nucleate the microtubules of the mitotic spindle in animal cells. Basal bodies are the MTOCs that assemble the microtubules of cilia and flagella (see Figure 18-5).
- Centrosomes consist of two centrioles and the surrounding pericentriolar material, which contains the γ-TuRC microtubule-nucleating complex (see Figures 18-6 and 18-

7).

18.2 Microtubule Dynamics

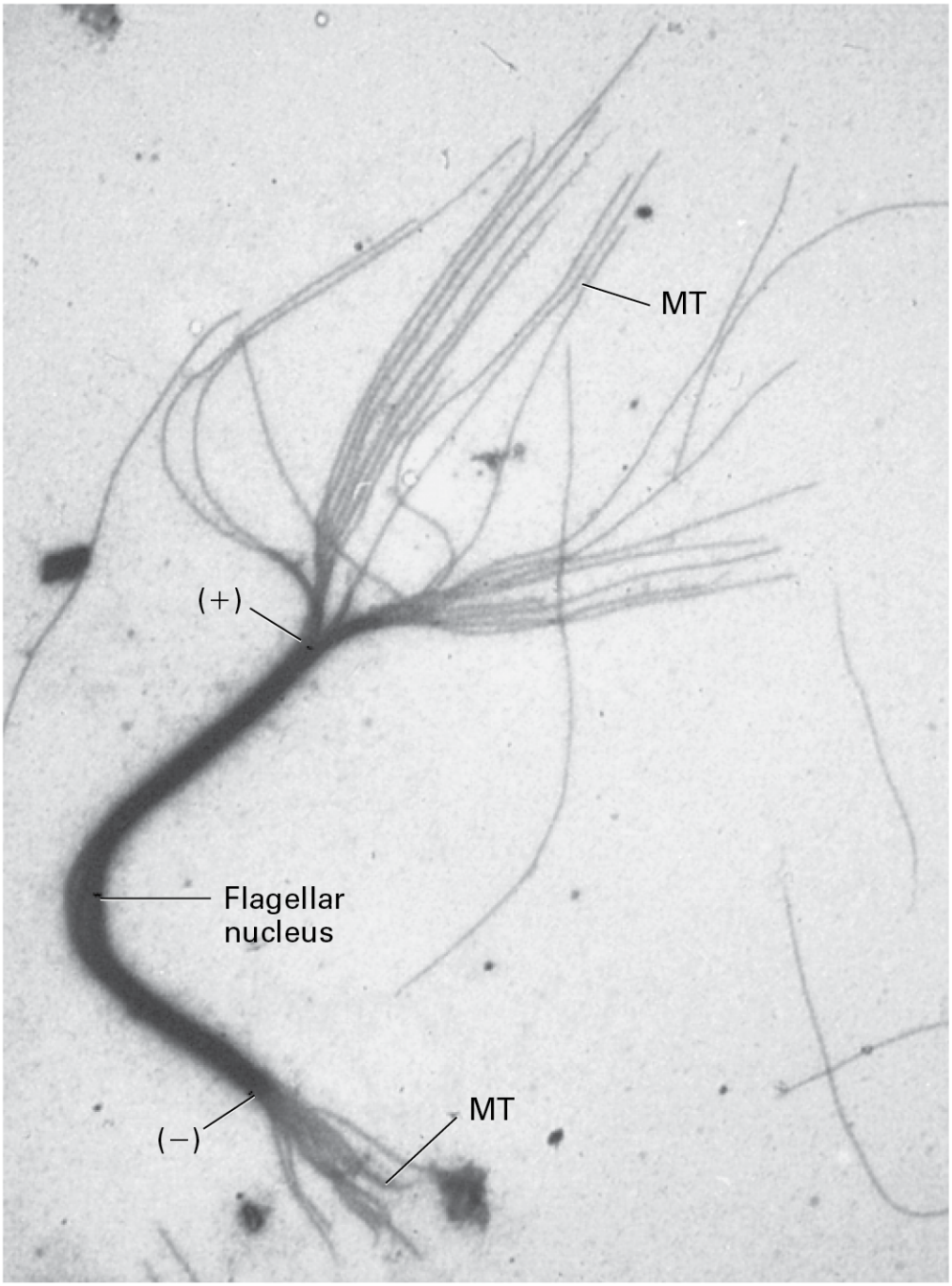
Microtubules are dynamic structures that can assemble or disassemble rapidly at their ends. Their lifetimes can vary enormously, averaging less than 1 minute for cells in mitosis and about 5–10 minutes for the microtubules that make up the radial array seen in interphase animal cells. Microtubule lifetime is longer in axons and much longer in cilia and flagella. To see how these differences occur, we will look at the dynamic properties of microtubules and how these properties contribute to their cellular organization.

Individual Microtubules Exhibit Dynamic Instability

Early experiments revealed that most microtubules in animal cells disassemble when the cells are cooled to 4 °C and reassemble when the cells are warmed to 37 °C. Researchers realized that this intrinsic property of microtubules could be exploited to purify their components. Since brain tissue is rich in microtubules, soluble extracts of pig brains were prepared at 4 °C; these clarified extracts were then warmed to 37 °C to induce microtubule assembly. The assembled microtubules were collected into a pellet by centrifugation, separated from the supernatant, and then disassembled by adding buffer at 4 °C. After another cycle of

assembly by warming, collection, and disassembly by cooling, researchers recovered *microtubular protein*, a collective term for αβ-tubulin and **microtubule-associated proteins (MAPs)**. They were then able to fractionate the microtubular protein into pure αβ-tubulin and MAPs to study their behaviors separately. Investigators found that polymerization of dimeric αβ-tubulin into microtubules can be catalyzed by the presence of the MAPs.

Research on the bulk polymerization properties of microtubular proteins in solution revealed two important things. First, for assembly to occur, the αβ-tubulin concentration must be above the *critical concentration* (C_c), just as we saw for actin polymerization (see Figure 17-8). Second, at αβ-tubulin concentrations higher than C_c, tubulin dimers are added faster to one end of the microtubule than to the other (Figure 18-8). As with F-actin assembly, the preferred end for assembly, which is the end with β-tubulin exposed, is designated the (+) end. The (−) end has α-tubulin exposed (see Figure 18-3b).



Courtesy of Gary G. Borisy.

EXPERIMENTAL FIGURE 18-8 Microtubules grow preferentially at the (+) end. A fragment of a microtubule bundle from a flagellum was used as a nucleus for the in vitro addition of αβ-tubulin. The nucleating flagellar fragment is the thick bundle seen in this electron micrograph, with the newly formed microtubules (MT) radiating from its ends. The greater length of the microtubules at one end,

the (+) end, indicates that tubulin subunits are added preferentially to this end.

When studying the bulk properties of microtubule assembly by raising the temperature of a tubulin solution from 4 °C to 37 °C, a typical curve reflecting nucleation, elongation, and steady state is observed (Figure 18-9a). Although one might assume that all microtubules would behave similarly, when researchers explored the behavior of individual microtubules within a population, they found that this was not the case. Individual microtubule behavior was examined in a simple experiment: microtubules were assembled in vitro and then sheared to break them into shorter pieces whose individual lengths could be analyzed by microscopy. Under these conditions, one might expect all the short microtubules either to grow or to shrink, depending on the free tubulin concentration. However, the investigators found that some of the microtubules grew in length, whereas others shortened very rapidly — thus indicating the existence of two distinct populations of microtubules. Further studies showed that individual microtubules could grow and then suddenly experience a *catastrophe*: an abrupt transition to a shrinking phase during which the microtubule would undergo rapid depolymerization. Moreover, sometimes a depolymerizing microtubule end could go through a *rescue* and begin growing again (Figure 18-9b). Although this phenomenon was first seen in vitro, analysis of fluorescently labeled tubulin microinjected into live cells showed that microtubules in cells also

undergo periods of growth and shrinkage (Figure 18-10). This alternation between growing and shrinking states is known as *dynamic instability*. Thus the dynamic life of a microtubule end is determined by the rate of growth, the frequency of catastrophes, the rate of depolymerization, and the frequency of rescues. As we will see, these four features of microtubule dynamics are controlled in vivo. Since the (−) ends of microtubules in animal cells are generally anchored to an MTOC, this dynamism is most relevant to the (+) end of the microtubule.

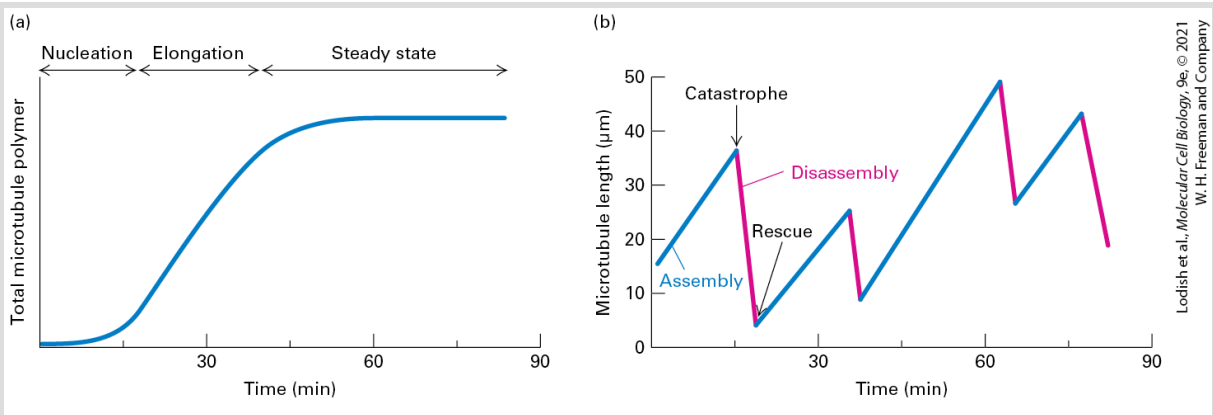
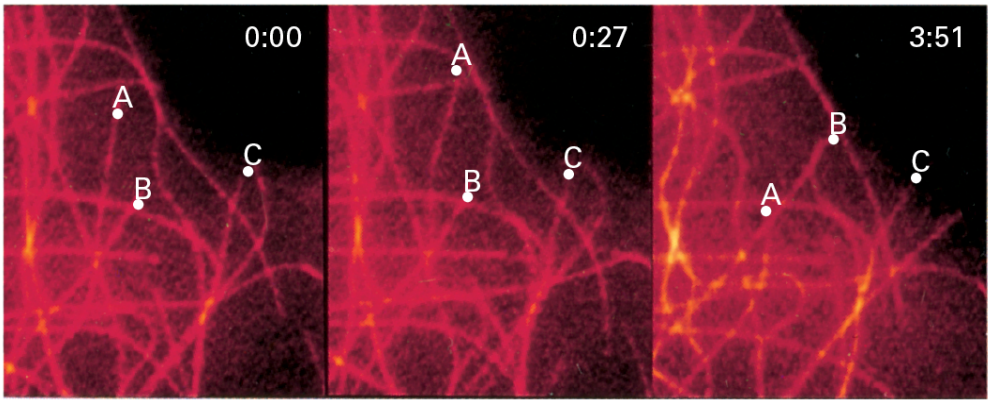


FIGURE 18-9 Comparison of the assembly properties of a preparation of bulk tubulin compared with the dynamics of an individual microtubule. (a) When pure tubulin above its C_c is warmed from 4 °C to 37 °C to induce polymerization, typical nucleation, elongation, and steady-state phases are seen. (b) However, when individual microtubules in an assembling population are observed in the light microscope and their lengths plotted at different times, they can be seen to grow and shrink in a process known as *dynamic instability*. Assembly and disassembly each proceed at uniform rates, but there is a big difference between the rate of assembly and that of disassembly, as seen in the different slopes of the lines. Shortening of a microtubule is much more rapid (7 μm/min) than growth (1 μm/min). Notice the abrupt transitions to the shrinkage stage (catastrophe) and to the elongation stage (rescue).

[Data from P. M. Bayley, K. K. Sharma, and S. R. Martin, 1994, in J. S. Hyams and C. W. Lloyd, eds., *Microtubules*, p. 118. New York: Wiley-Liss.]



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-10 Fluorescence microscopy reveals growth and shrinkage of individual microtubules in vivo. Fluorescently labeled tubulin was microinjected into cultured human fibroblasts. The cells were chilled to depolymerize preexisting microtubules into tubulin dimers and were then incubated at 37 °C to allow repolymerization, which incorporated the fluorescent tubulin into all the cells' microtubules. A region of a cell periphery was viewed in the fluorescence microscope at 0 seconds, 27 seconds later, and 3 minutes 51 seconds later (*left to right panels*). During this period, several microtubules can be seen to have elongated and shortened. The dots labeled *A*, *B*, and *C* mark the positions of the ends of three microtubules.

[Reprinted with permission from Nature Publishing Group, from P. J. Sammak and G. G. Borisy, 1988, "Direct Observation of Microtubule Dynamics in Living Cells," *Nature*, **332**:724–726; permission conveyed through Copyright Clearance Center, Inc.]

What is the molecular basis of dynamic instability? If we look carefully at the ends of growing and shrinking microtubules in the electron microscope, we can see that they are quite different. A growing microtubule has a relatively blunt end, whereas a

depolymerizing end has highly curved protofilaments peeling off like rams' horns (Figure 18-11). In fact, the growing microtubule end is not simply a blunt end, but a ragged collection of gently curving protofilaments that then associate laterally to form the cylindrical microtubule with straight protofilaments.

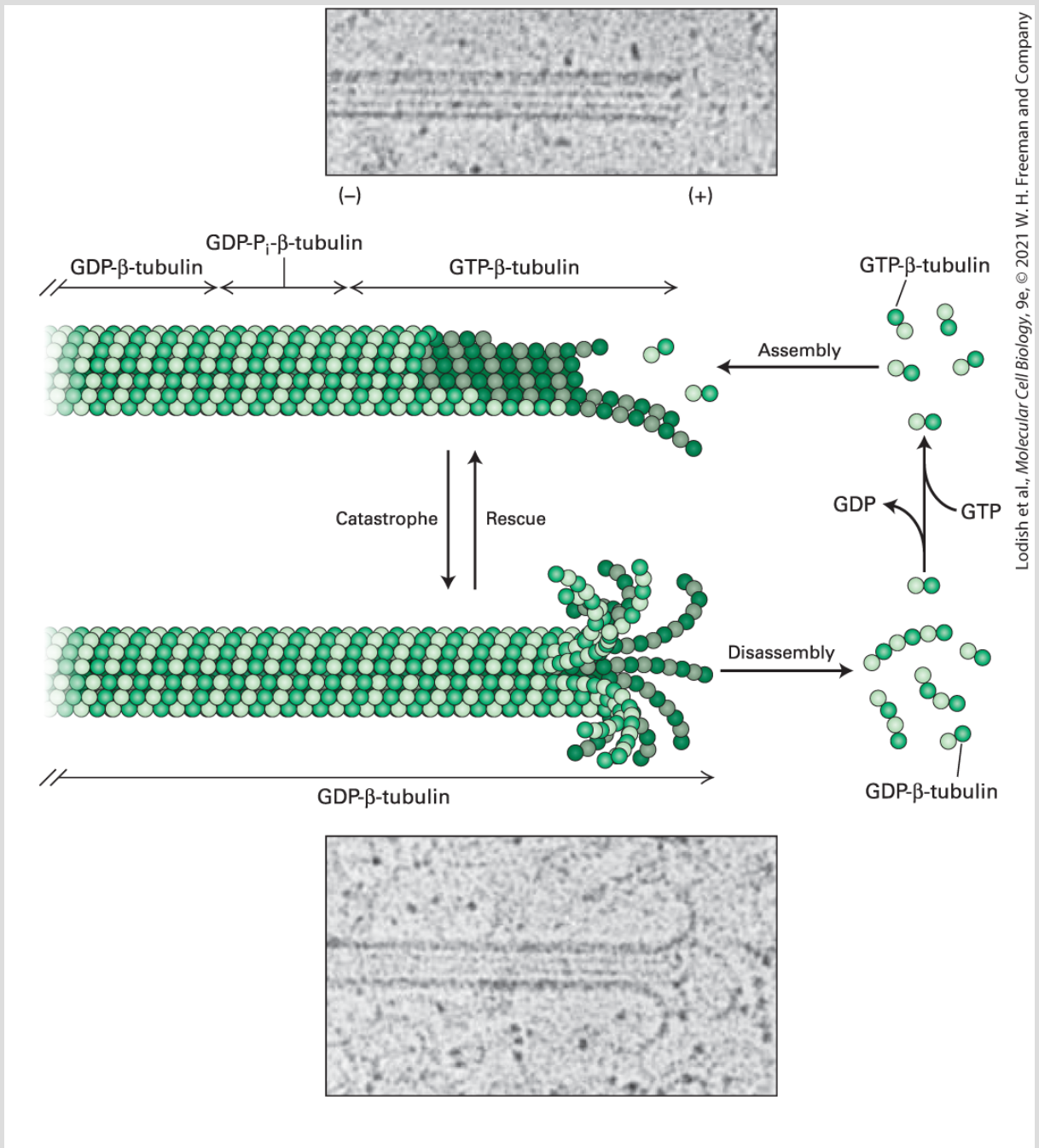


FIGURE 18-11 Dynamic instability depends on the presence or absence of a GTP-β-tubulin cap. Images taken in the electron microscope of frozen samples of a growing microtubule (*upper*) and a shrinking microtubule (*lower*). Notice that the end of the growing microtubule has a blunter end, whereas the end of the shrinking one has curls like a ram's horns. As the diagram shows, a microtubule with several GTP-β-tubulin dimers on the end of each protofilament is strongly favored to grow. However, a microtubule with GDP-β-tubulin at the ends forms highly curved protofilaments and will undergo rapid disassembly. Switches between growing and shrinking phases, called *rescues* and *catastrophes*, can occur, and the rate of switching is regulated by associated proteins.

[E. M. Mandelkow, E. Mandelkow, and R. A. Milligan, 1991, *J. Cell Biol.* **114**(5): 977–991; https://doi.org/10.1083/jcb.114.5.977.]

A simple structural difference accounts for the morphology of these two different classes of microtubule (+) ends. Using a GDP analog, researchers have found that artificially made *single* protofilaments — which are not exposed to lateral interactions — made up of repeating αβ-tubulin dimers containing GDP-β-tubulin are curved, like a ram's horn. However, artificially made single protofilaments made up of αβ-tubulin dimers in which β-tubulin has a bound GTP analog are only slightly curved. The assembling end of a microtubule contains these slightly curved nascent protofilaments, giving rise to the gently curved protofilaments characteristic of growing ends (see Figure 18-11). The gently curved protofilaments containing GTP-β-tubulin then zip up by lateral interactions to form straight protofilaments within the microtubule. Such a growing microtubule is said to have a GTP cap. By contrast, shrinking microtubules with highly curled ends are made up of GDP-β-tubulin. Therefore, if the GTP molecules in the terminal β-tubulins become hydrolyzed on a

microtubule that has stopped growing, a formerly blunt-ended microtubule will curl and a catastrophe will ensue. These relationships are summarized in Figure 18-11.

These results have an additional and fascinating implication, but to understand it we must consider the growing microtubule in more detail. The addition of dimers with GTP-β-tubulin to the (+) end of protofilaments on a growing microtubule will generate the GTP cap on each protofilament. These elongated protofilaments then associate by lateral interactions to become straight and form the microtubule wall. This transition from a curved to straight protofilament induces a strained conformation in the tubulin dimer. The interaction between the β-subunit of one dimer and the α-subunit of the next enhances hydrolysis of GTP in the β-subunit to GDP-P_i. The inorganic P_i is then slowly released. In this way, a growing microtubule has a GTP cap, followed by a region in which β-tubulin has GDP-P_i, followed by the older region that has just GDP-β-tubulin (see Figure 18-11).

As we mentioned above, an *isolated* protofilament containing GDP-β-tubulin is curved along its length, so when it is present in a microtubule, why doesn't it break out and peel away? The lateral protofilament-protofilament interactions in the GTP-β-tubulin cap are sufficiently strong that they do not allow the microtubule to unpeel at its end — and so the protofilaments behind the GTP-β-tubulin cap are constrained from unpeeling (see Figure 18-11). The

energy released by GTP hydrolysis in the subunits behind the cap is stored within the lattice as structural strain waiting to be released when the GTP-β-tubulin cap is lost. This implies that if the GTP-β-tubulin cap is lost, the stored energy can do work if some structure, such as a chromosome, is attached to the disassembling microtubule end. As we will see, this stored energy drives the movement of chromosomes during the anaphase stage of mitosis.

How can a disassembling microtubule suddenly be rescued to grow again? One possible answer to this perplexing problem has been suggested. Using an antibody that recognizes GTP-β-tubulin but not GDP-β-tubulin, researchers have found that islands of GTP-β-tubulin can persist along the length of an assembled microtubule. It seems likely that when a disassembling microtubule encounters one of these GTP-β-tubulin islands, disassembly pauses, and a rescue may be provoked.

Localized Assembly and *Search and Capture* Help Organize Microtubules

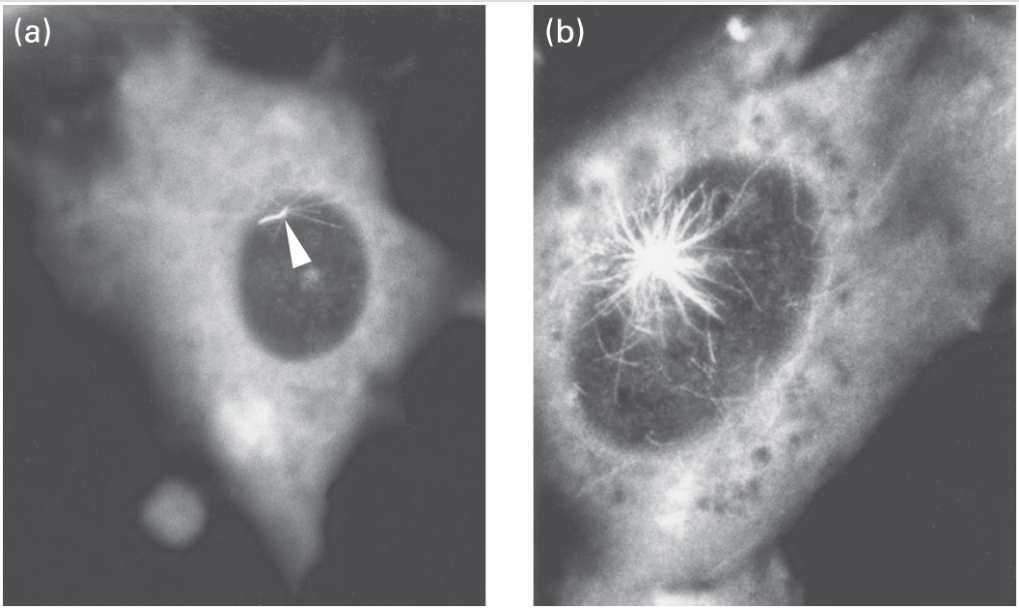
We have now presented two major concepts relating to microtubule organization and (+) end dynamics: microtubules are assembled starting from MTOCs, and individual microtubules can undergo dynamic instability. Together, these two processes contribute to the distribution of microtubules in cells.

In an interphase cell growing in culture, microtubules are constantly being nucleated from the centrosome and spreading out, randomly *searching* the cytoplasmic space. The frequency of catastrophes and rescues, together with growth and shrinkage rates, determines the length of each microtubule. If the microtubule is subject to a high catastrophe frequency and a low rescue frequency, it will shrink back to the centrosome and disappear, whereas if it has few catastrophes and is readily rescued, it will continue to grow. If the searching microtubule encounters an appropriate target on a cell structure or organelle, the microtubule end may become attached to that structure. Organelle or cell-structure *capture* by the microtubule can stabilize its (+) end and protect it from catastrophes, whereas unattached microtubules have a greater frequency of disassembly. So the dynamics of the microtubule (+) end is a very important determinant of microtubule life cycle and function. *Search and capture* is part of the mechanism determining the overall organization of microtubules in a cell. Moreover, by changing the rate of nucleation or local microtubule dynamics and capture sites, a cell can rapidly change its overall microtubule distribution. We will see later that this is what happens as cells enter mitosis.

Drugs Affecting Tubulin Polymerization Are Useful

Experimentally and in Treatment of Diseases

The conserved nature of tubulins and their essential role in critical processes such as mitosis make them prime targets for both naturally occurring and synthetic drugs that affect polymerization or depolymerization. The first of these drugs was colchicine (present in extracts of the meadow saffron), which binds tubulin dimers so that they cannot polymerize into microtubules. Since most microtubules are in a dynamic state between dimers and polymers, the addition of colchicine sequesters all free dimers in the cytoplasm, resulting in loss of microtubules due to their natural turnover. Treatment of cultured cells with colchicine for a short time results in the depolymerization of all the cytoplasmic microtubules, leaving the more stable tubulin-containing centrosome (Figure 18-12a). When the colchicine is washed out to allow regrowth of the microtubules, they can be seen to grow from the centrosome, revealing its ability to nucleate the assembly of new microtubules (Figure 18-12b).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-12 Microtubules grow from the MTOC. To investigate the assembly of microtubules in vivo, a cultured fibroblast was treated with colchicine until almost all the cytoplasmic microtubules were disassembled. The cell was then stained with antibodies to tubulin and viewed by immunofluorescence microscopy (a). In a separate experiment, the colchicine was washed out to allow reassembly of the microtubules. Panel (b) shows the first stages of reassembly, revealing microtubules growing from the MTOC in the central region above the nucleus (dark area). Note in panel (a) the remaining primary cilium (arrowhead; discussed in Section 18.5) associated with the centrosome, which is not depolymerized by colchicine treatment under these conditions. Note also the fluorescence from the cytoplasm, which is from unpolymerized αβ-tubulin dimers.

[E. M. Mandelkow, E. Mandelkow, and R. A. Milligan, 1991, *J. Cell Biol.* **114**(5):977–991; https://doi.org/10.1083/jcb.114.5.977.]

Colchicine has been used for hundreds of years to relieve the joint pain of acute gout. A famous patient was King Henry VIII of England, who was treated with colchicine to relieve this ailment. A low level of colchicine relieves the inflammation caused in gout by

reducing the microtubule dynamics of white blood cells, rendering them unable to migrate efficiently to the site of inflammation.

In addition to colchicine, a number of other drugs bind the tubulin dimer and restrain it from forming polymers. These drugs include podophyllotoxin (from juniper) and nocodazole (a synthetic drug).

Taxol, a plant alkaloid from the Pacific yew tree, binds microtubules and stabilizes them against depolymerization. Because taxol stops cells from dividing by inhibiting mitosis, it has been used to treat some cancers, such as those of the breast and ovary, where the cells are especially sensitive to the drug. ■

KEY CONCEPTS OF SECTION 18.2

Microtubule Dynamics

- Individual microtubule (+) ends can undergo dynamic instability, showing alternating periods of growth or rapid shrinkage (see Figure 18-10).
- Most of the β-tubulin in microtubules contains bound GDP. In growing microtubules, the (+) ends have a cap of many dimers containing GTP-β-tubulin, resulting in slightly splayed ragged protofilament ends. Shrinking microtubules have lost the cap, causing the protofilaments at the end, containing GDP-β-tubulin, to peel outward and disassemble (see Figure 18-11).
- Growing microtubules store the energy derived from GTP hydrolysis in the microtubule lattice, so they have the potential to do work when disassembling.
- Microtubules assembled from the centrosome and exhibiting dynamic instability can search the cytoplasm for structures or organelles with appropriate targets and capture them, which results in stabilization of the microtubule (+) end. In this way, assembly coupled with search and capture can contribute to the overall distribution of microtubules in a cell.

18.3 Regulation of Microtubule Structure and Dynamics

The regulation of microtubule structure and dynamics is critical for proper cell function. As we will see later in this chapter, microtubules are the major organizers of organelles in animal cells, and the stability and dynamics of microtubules are tailored for the specific function of the cell at any given time. For example, the dynamics of microtubules increase dramatically as a cell enters mitosis to allow the cell to build the mitotic spindle.

The wall of a microtubule is built from αβ-tubulin dimers, and highly purified αβ-tubulin will assemble in vitro into microtubules. Assembly of microtubules in vitro is greatly enhanced by the presence of certain microtubule-associated proteins (MAPs). MAPs regulate microtubule dynamics in a number of ways. Broadly, some stabilize microtubules, other MAPs destabilize them, and a third group of MAPs alters the growth properties of microtubules. We discuss the various examples of MAPs in this section, which are summarized in Table 18-1.

TABLE 18-1 • Summary of the Microtubule-Associated Proteins (MAPs) Discussed in This Section

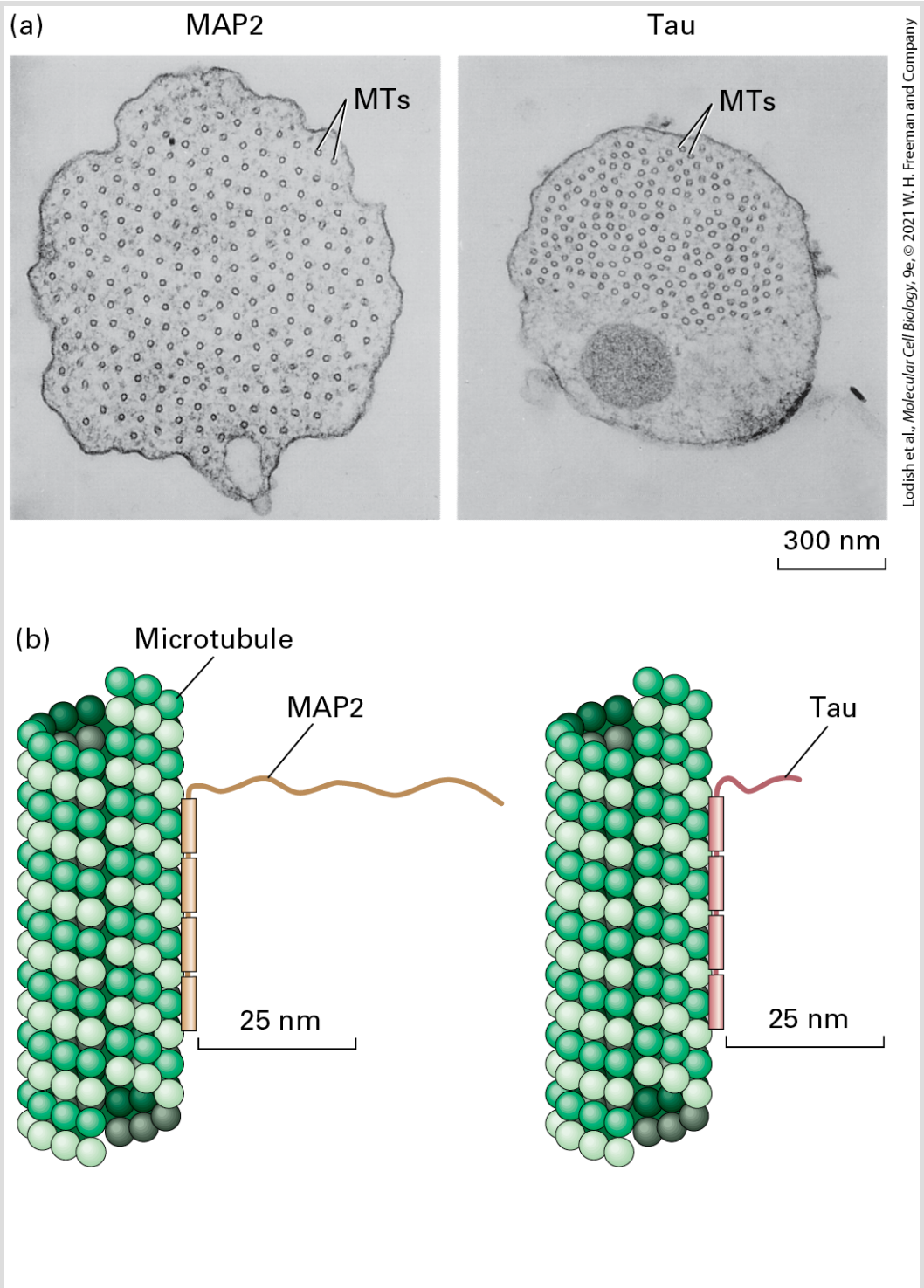
Name	Function
MAP2, MAP4, Tau	Stabilizing

EB1	+TIP enhances catastrophes
XMAP215	+TIP promotes (+) end growth
CLASP	+TIP suppresses (+) end catastrophes
Kinesin-13	Promotes end disassembly
Op18/Statmin	Promotes (+) end disassembly
Katanin	AAA ATPase severing

Microtubules Are Stabilized by Side-Binding Proteins

Several classes of proteins stabilize microtubules, many of which show cell-type-specific expression. Among the best-studied MAPs are members of the tau family of proteins, which includes tau itself as well as proteins called *MAP2* and *MAP4*. Tau and MAP2 are neuronal proteins, while MAP4 is expressed by other cell types and is generally not present in neurons. These proteins have a modular design with two key domains. The first domain consists of a positively charged 18-residue sequence, repeated three to four times, that binds to the negatively charged microtubule surface. The second domain projects outward at a right angle from the microtubule (Figure 18-13b). Tau proteins stabilize microtubules and also act as spacers between them. MAP2 is found only in dendrites of neurons, where it forms fibrous cross-bridges between

microtubules and links microtubules to intermediate filaments. Tau, which is much smaller than most other MAPs, is present in both axons and dendrites. The basis for this specificity is still a mystery.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

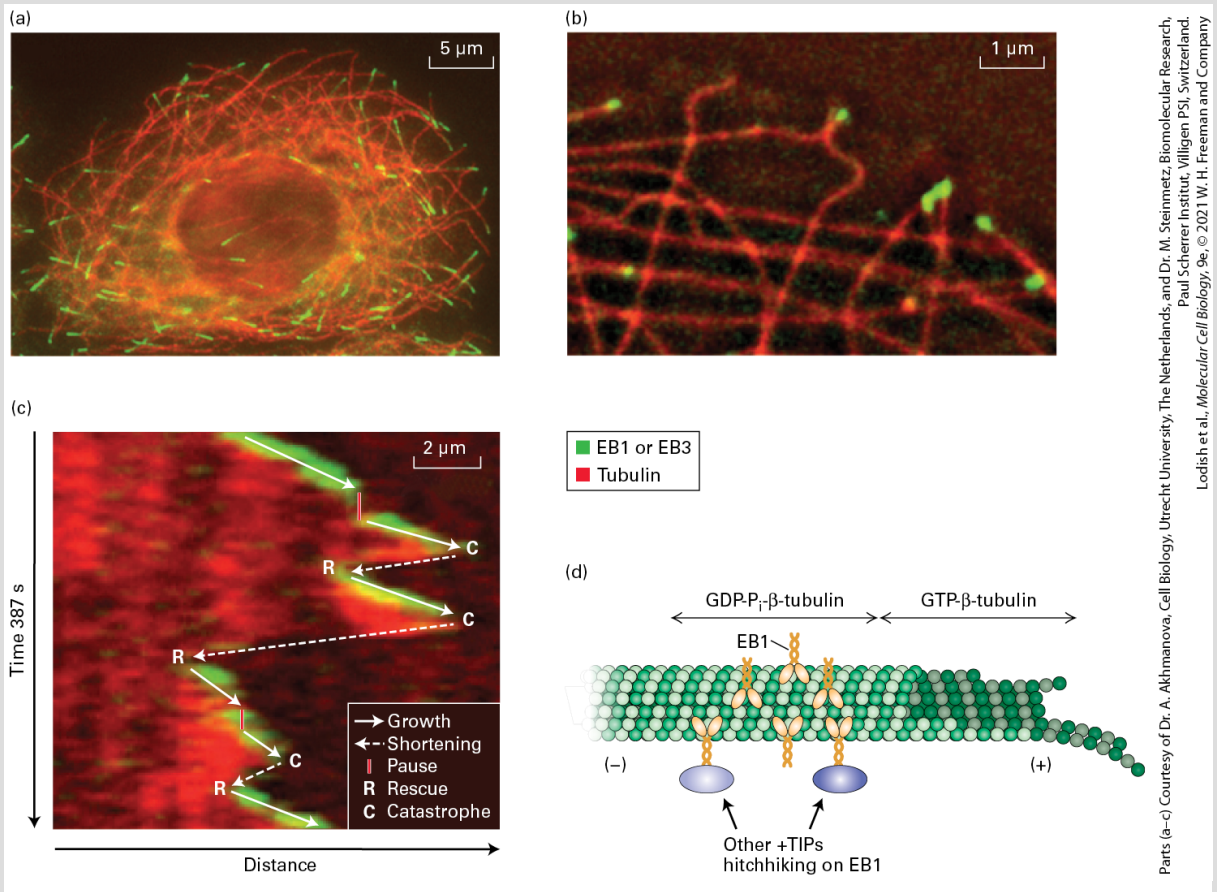
EXPERIMENTAL FIGURE 18-13 Spacing of microtubules depends on the length of the projection domains of microtubule-associated proteins. (a) Insect cells transfected to express MAP2, which has a long arm, or to express tau protein, which has a short arm, grow long axon-like processes. These electron micrographs show cross sections through the processes induced by the expression of MAP2 (*left*) and tau (*right*) in transfected cells. Note that the spacing between microtubules (MTs) in MAP2-containing cells is greater than in tau-containing cells. Both cell types contain approximately the same number of microtubules, but the effect of MAP2 is to enlarge the caliber of the axon-like process. (b) Diagrams of association between microtubules and MAPs. Note the difference between the lengths of the projection domains in MAP2 and in tau.

[Part (a) Reprinted by permission from Nature Publishing Group, from J. Chen et al., 1992, "Projection Domains of MAP2 and Tau Determine Spacings Between Microtubules in Dendrites and Axons," *Nature*, **360**(6405):674–676; permission conveyed throght Copyright Clearance Center, Inc.]

When stabilizing MAPs coat the outer wall of a microtubule, they can increase the growth rate of microtubules or reduce catastrophe frequency. In many cases, the activity of the MAPs is regulated by the reversible phosphorylation of their projection domains. Phosphorylated MAPs are unable to bind to microtubules; thus phosphorylation promotes microtubule disassembly. For example, microtubule-affinity-regulating kinase (MARK/Par-1) is a key modulator of tau proteins. Some MAPs, including MAP4, are also phosphorylated by a cyclin-dependent kinase (CDK) that plays a major role in controlling the activities of proteins in the course of the cell cycle (see Chapter 19).

+TIPs Regulate the Properties and Functions of the Microtubule (+) End

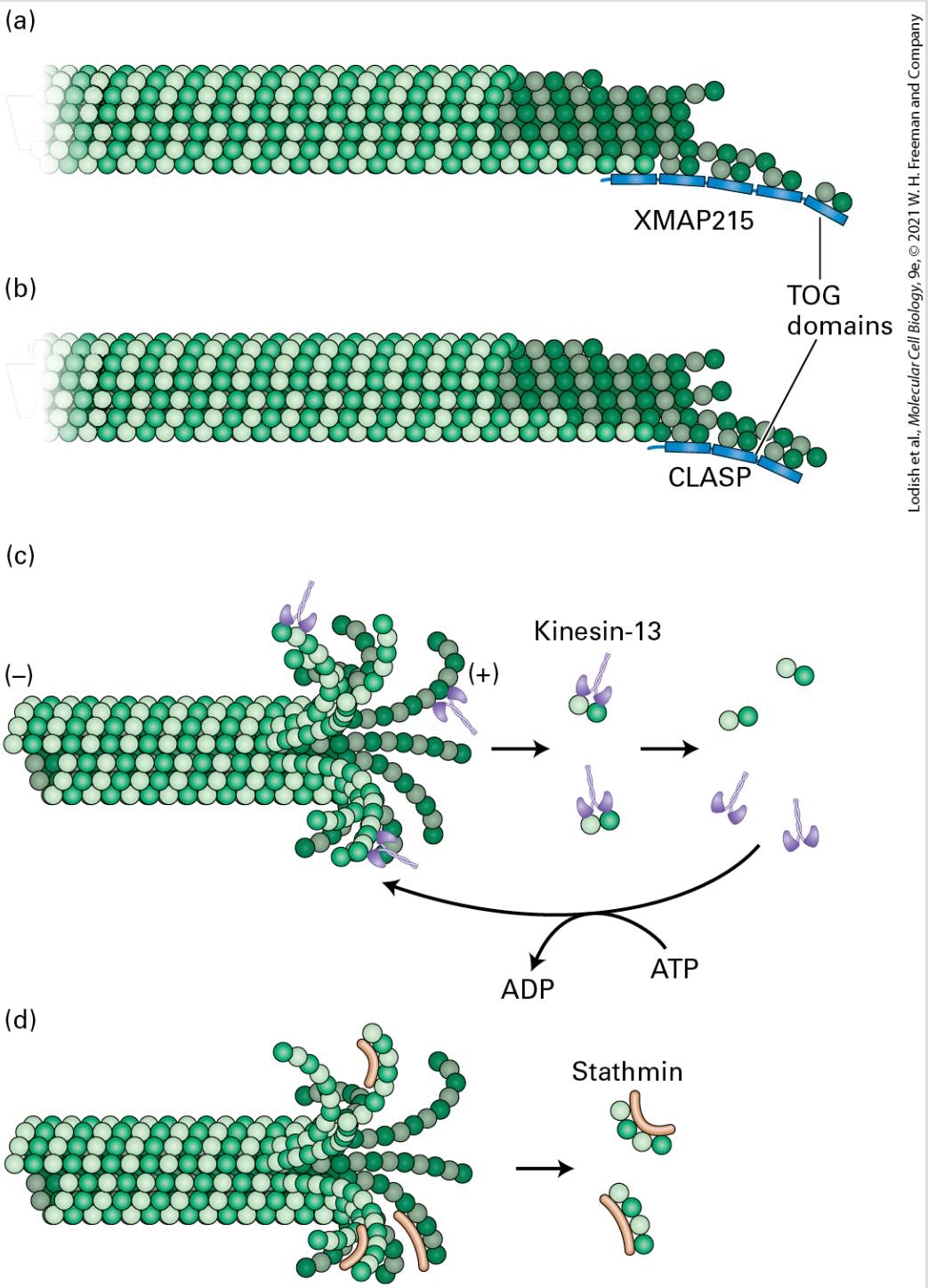
In addition to the side-binding MAPs, MAPs that associate with a region near the (+) end of microtubules have been identified. In many cases, they associate only with (+) ends that are growing (Figure 18-14a, b). The MAPs in this class are known as *+TIPs*, for plus-end tracking proteins. The best studied of these is between the +TIP called *EB1* (end binding-1) and the related *EB3* and a growing microtubule (+) end. High-resolution microscopy has revealed that EB1 associates with the microtubule tip just behind the GTP-β-tubulin-containing cap, in the region containing GDP-P_i-β-tubulin (see Figure 18-11). Binding of EB1 induces a twist in the tubulin subunits that is thought to transmit a long-range mechanical signal to enhance GTP hydrolysis in the cap. Thus at least in some in vitro studies EB1 is a destabilizer of microtubules leading to a higher rate of catastrophe (Figure 18-14c). Some +TIPs associate with the (+) end either by binding EB1 or by requiring EB1 for their association with the (+) end, and are generally said to be "hitchhiking" on EB1 (Figure 18-14d).



Parts (a–c) Courtesy of Dr. A. Akhmanova, Cell Biology, Utrecht University, The Netherlands, and Dr. M. Steinmetz, Biomolecular Research, Paul Scherrer Institut, Villigen PSI, Switzerland.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-14 The +TIP protein EB1 associates near the (+) ends of microtubules. (a) A cultured cell stained with antibodies to tubulin (red) and the +TIP protein EB1 (green). EB1 is enriched in the region of the microtubule (+) end. (b) Edge of a live cell expressing EB3-GFP (green) and mCherry-α-tubulin (red). EB3, which is closely related to EB1, is found very close to the end of some microtubules. (c) EB3-GFP selectively associates with growing microtubules, as seen in this so-called kymograph. In this figure, the dynamics of a single microtubule (red) and EB3 (green) in a live cell like that shown in part (b) is followed by taking the same region from sequential frames of a movie and lining them up top to bottom. At the top, one sees the start of the movie with the microtubule capped by EB3. Moving down the figure, one can track the dynamics of the microtubule over time as it grows and shrinks. When the microtubule grows, it remains capped by EB3. When microtubule growth pauses or the microtubule shrinks, EB3 is no longer associated with the end, but it becomes reassociated when growth resumes. A diagrammatic summary of the microtubule dynamics overlies the kymograph. (d) A possible mechanism for EB1 binding to a growing microtubule and “hitchhiking” by other proteins on EB1.

Other +TIPs, that bind independently of EB1, can promote microtubule growth either by enhancing assembly or by suppressing catastrophes. For example, a protein called *XMAP215* contains five so-called TOG domains. These domains have the ability to bind free αβ-tubulin dimers as well as the gently curved regions of protofilaments at the growing end of a microtubule. By binding to the growing end and bringing more αβ-tubulin dimers there, XMAP215 effectively increases the local αβ-tubulin concentration to enhance microtubule assembly (Figure 18-15a). Another class of proteins, called *CLASPs*, have related TOG domains but do not enhance assembly. Instead, they bind to the gently curved growing end and suppress catastrophes (Figure 18-15b). Another function of +TIPs is to link the microtubule (+) end to other cellular structures, such as the cell cortex, F-actin, and as we will see later during our discussion of mitosis, chromosomes; a key feature of this dynamic system is that when a +TIP at the end of a searching microtubule encounters an appropriate target, the microtubule can become captured and stabilized. Yet other +TIPs link microtubule (+) ends to membranes; for example, linkage to the endoplasmic reticulum transmembrane protein *STIM* promotes microtubule-dependent extension of the tubular endoplasmic reticulum.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-15 Proteins that regulate the ends of microtubules. (a) The protein XMAP15 has five TOG domains, some of which bind the gently curved protofilaments of a growing microtubule end, whereas others bind αβ-tubulin dimers to enhance gowth. (b) CLASP also has TOG domains that stabilize the

ends of microtubules to suppress catastrophes. (c) A member of the kinesin-13 family enriched at a microtubule end can enhance the disassembly of that end. [Although depolymerization of the (+) end is shown, kinesin-13 can also depolymerize the (−) end.] These proteins are ATPases, and ATP enhances their activity by dissociating them from the αβ-tubulin dimer. (d) Op18/stathmin binds selectively to curved protofilaments and enhances their dissociation from a microtubule end.

Other End-Binding Proteins Also Promote Microtubule Disassembly

Various mechanisms for microtubule destabilization are known. One of these involves the kinesin-13 family of proteins. As we will see in Section 18.4, most kinesins are molecular motors, but the kinesin-13 proteins are a distinct class of kinesin that bind to and curve the end of the tubulin protofilaments into the GDP-β-tubulin conformation. They then facilitate the removal of terminal tubulin dimers, thereby greatly enhancing the frequency of catastrophes (Figure 18-15c). They act catalytically in the sense that they need to hydrolyze ATP to remove sequential terminal tubulin dimers.

Another protein, known as *Op18/stathmin*, also enhances the rate of catastrophes. It was originally identified as a protein highly overexpressed in certain cancers; hence part of its name (Oncoprotein 18). Op18/stathmin is a small protein that binds two tubulin dimers in a curved, GDP-β-tubulin-like conformation (Figure 18-15d). It may function by enhancing the hydrolysis of the GTP in

the terminal tubulin dimer and aiding in its dissociation from the end of the microtubule. As might be expected for a regulator of microtubule ends, it is subject to negative regulation by phosphorylation by a large variety of kinases. In fact, it has been found that Op18/stathmin is inactivated by phosphorylation near the leading edge of a motile cell, which contributes to preferential growth of microtubules toward the front of the cell.

Severing Proteins Also Regulate Microtubule Dynamics

Animals and plants also have microtubule-severing proteins. All microtubule-severing proteins are members of the AAA ATPases, enzymes that have the ability to use the power of ATP hydrolysis to disassemble protein complexes. The first to be isolated was called *katanin* after the Japanese samurai sword. Katanin forms a six-membered ring that can sever a microtubule by literally pulling subunits out of the microtubule, resulting in its destabilization and severing. This results either in more rapid depolymerization of the microtubule, or in the majority of cases, the microtubule is repaired with GTP-tubulin to generate a new growing end. Although microtubule-severing proteins clearly have an important function in animal cells, it has been difficult to pinpoint exactly what that is. However, in plant cells, microtubule severing of cortical microtubules has been clearly documented.

KEY CONCEPTS OF SECTION 18.3

Regulation of Microtubule Structure and Dynamics

- Microtubules can be stabilized by side-binding microtubule-associated proteins (MAPs) (see Figure 18-13).
- Some MAPs, called *+TIPs,* bind selectively to growing (+) ends of microtubules and can alter the dynamic properties of the microtubule or attach components of the cell to the searching (+) end of the microtubule (see Figure 18-14).
- Proteins, such as XMAP15, enhance the growth of microtubule (+) ends, or stabilize them against catastrophes, such as CLASP (see Figure 18-15a, b).
- Microtubule ends can be destabilized by some proteins, such as the kinesin-13 family of proteins and Op18/stathmin, to enhance the frequency of catastrophes (see Figure 18-15c, d).
- MAPs, such as katanin, are AAA ATPases that sever microtubules.

18.4 Kinesins and Dyneins: Microtubule-Based Motor Proteins

Organelles in cells are frequently transported distances of many micrometers along well-defined routes in the cytoplasm and delivered to particular intracellular locations. Diffusion alone cannot account for the rate, directionality, and destinations of such transport processes. Findings from early experiments with fish-scale pigment cells and neurons demonstrated that microtubules function as tracks in the intracellular transport of various types of so-called cargo.

As already discussed, polymerization and depolymerization of microtubules can do work using the energy provided by GTP hydrolysis. In addition, **motor proteins** move along microtubules, powered by ATP hydrolysis. Two families of motor proteins — **kinesins** and **dyneins** — are known to mediate transport along microtubules. In this section, we first discuss the initial evidence for the existence of these motors and then how they work and the roles they perform in interphase cells. In subsequent sections, we discuss their functions in cilia and flagella, and in mitosis.

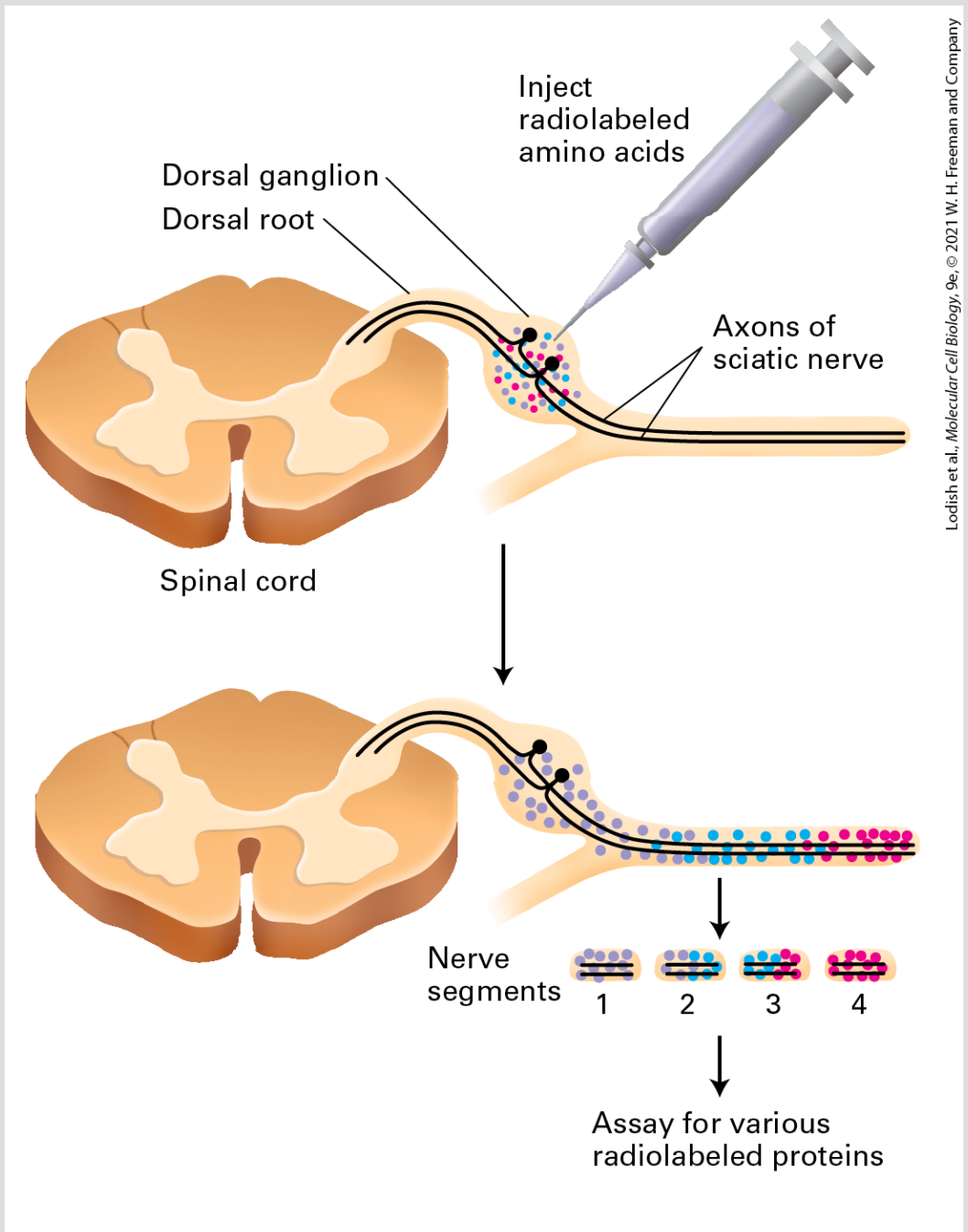
Organelles in Axons Are Transported Along Microtubules in Both Directions

A neuron must constantly supply new materials — proteins and membranes — to its axon terminal to replenish those lost in the exocytosis of neurotransmitters at the junction (synapse) with another cell (see Chapter 23). Because proteins and membranes are primarily synthesized in the cell body, these materials must be transported down the length of the axon, which can be as long as a meter in some neurons. This movement of materials is accomplished on microtubules, which are all oriented with their (+) ends toward the axon terminal (see Figure 18-5e).

The results of classic experiments, in which radioactive amino acids were microinjected into the dorsal-root ganglia near the spinal cord to allow for their incorporation into proteins in spinal neurons, and the radioactivity then tracked along the axons of those cells, showed that **axonal transport** occurs from the cell body down the axon. Other experiments showed that transport could also occur in the reverse direction, toward the cell body. *Anterograde* transport proceeds from the cell body to the axon terminal and is associated with axonal growth and the delivery of synaptic vesicles. *Retrograde* transport moves recycling membranes from the axon terminal along the axon toward the cell body.

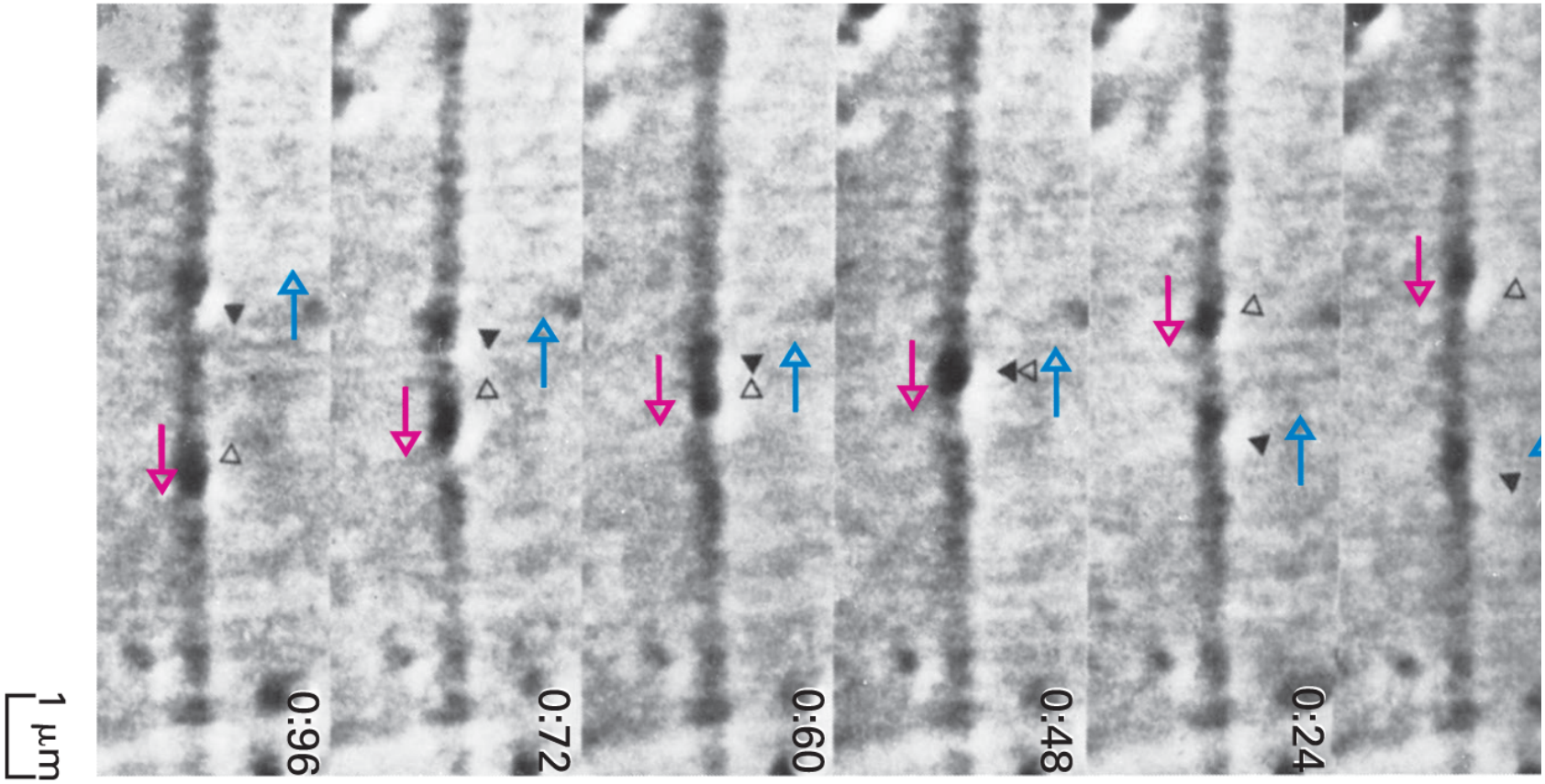
These experiments also revealed that different materials move at different speeds (Figure 18-16). The fastest moving material, consisting of membrane-bound vesicles, has a velocity of about 3 μm/s, or 250 mm/day — requiring about 4 days to travel from a cell

body in your back down an axon that terminates in your big toe. The slowest moving material, comprising tubulin subunits and neurofilaments (the intermediate filaments found in neurons), moves only a fraction of a millimeter per day. Organelles such as mitochondria move down the axon at an intermediate rate.



EXPERIMENTAL FIGURE 18-16 The rate of axonal transport in vivo can be determined by radiolabeling and gel electrophoresis. The cell bodies of neurons in the sciatic nerve are located in dorsal-root ganglia (near the spinal cord). Radioactive amino acids injected into these ganglia in experimental animals are incorporated into newly synthesized proteins, which are then transported down the axon to the synapse. Animals are sacrificed at various times after injection and the dissected sciatic nerve is cut into small segments to see how far radioactively labeled proteins have been transported; these proteins can be identified after gel electrophoresis and autoradiography. The red, blue, and purple dots represent groups of proteins that are transported down the axon at different rates, red most rapidly, purple least rapidly.

Neurobiologists have long made extensive use of the squid giant axon for studying organelle movement along microtubules. Involved in regulating the squid's water propulsion system, the aptly named giant axon can be up to 1 mm in diameter, which is about 100 times wider than the average mammalian axon. Moreover, squeezing the axon like a tube of toothpaste pushes out the cytoplasm (also known as *axoplasm*), which can then be observed by video microscopy. The movement of vesicles along microtubules in this cell-free system requires ATP, moving at a rate similar to that of axonal transport in intact cells, and can proceed in both anterograde and retrograde directions (Figure 18-17). Collectively, these experiments showed that there must be ATP-dependent motors capable of moving cargo along microtubules in anterograde and retrograde directions. Moreover, electron microscopy of squid giant axon cytoplasm showed organelles attached to individual microtubules.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Comp

EXPERIMENTAL FIGURE 18-17 DIC microscopy demonstrates microtubule-based vesicle transport in vitro. Cytoplasm was squeezed from a squid giant axon with a roller onto a glass coverslip. After buffer containing ATP was added to the preparation, it was viewed by differential interference contrast (DIC) microscopy, and the images were recorded on videotape. In the sequential images shown, the two organelles indicated by open and solid triangles move in opposite directions (indicated by colored arrows) along the same filament, pass each other, and continue in their original directions. Elapsed time in seconds appears at the upper-right corner of each video frame.

[Republished with permission from Elsevier, from B. J. Schnapp et al., 1985. "Single Microtubules from Squid Axoplasm Support Bidirectional Movement of Organelles," *Cell*, **40**(2):455–462; permission conveyed through the Copyright Clearance Center, Inc.]

In other experiments, neurofilaments were tagged with green fluorescent protein (GFP) and then injected into cultured cells. The results of these experiments suggest that neurofilaments pause frequently as they move down an axon. Although the peak velocity of neurofilaments is similar to that of fast-moving vesicles, their numerous pauses lower the average rate of transport. These findings suggest that there is no fundamental difference between fast and slow axonal transport.

Kinesin-1 Powers Anterograde Transport of Vesicles Down Axons Toward the (+) Ends of Microtubules

The protein responsible for anterograde organelle transport was first purified from axonal extracts. Researchers found that when they mixed three components — purified organelles from squid axons, an organelle-free cytosolic axonal extract, and taxol-stabilized microtubules — organelles could be seen moving on the microtubules in an ATP-dependent manner. However, if they omitted the axonal extract, the organelles neither bound nor moved along the microtubules. The researchers concluded that the extract contributes a protein that both attaches organelles to a microtubule and transports them along it — that is, a motor protein.

Their strategy for purifying the motor protein was based on additional observations of organelles moving on microtubules. It was known that if ATP was hydrolyzed to ADP, the organelles fell off the microtubules. However, if the nonhydrolyzable ATP analog AMPPNP was added, the organelles remained associated with the microtubules, but did not move. These observations suggested that the motor protein linked the organelles to the microtubules very tightly in the presence of AMPPNP, but when AMPPNP was replaced by ATP, which was subsequently hydrolyzed to ADP, then the organelles separated from the microtubules. By looking for a protein that binds microtubules in the presence of AMPPNP and is released upon the addition of ATP, researchers were able to purify the motor protein, which they called *kinesin*.

Many kinesins have now been purified. Kinesin-1 isolated from squid giant axons is a dimer of two heavy chains, each associated with a light chain, with a total molecular weight of about 380,000 Da. The molecule contains a pair of globular N-terminal *head domains* connected by a short flexible *linker domain* to a long *central stalk*, and ends in a pair of small globular *tail domains*, which associate with the light chains (Figure 18-18). Each domain carries out a particular function: the head domain binds microtubules and ATP and is responsible for the motor activity of kinesin; the linker domain is critical for forward movement; the stalk domain is involved in dimerization through a coiled-coil interaction (see Figure 3-7a) of the two heavy chains; and the tail domain is responsible for binding to receptors on the membrane vesicles containing cargo.

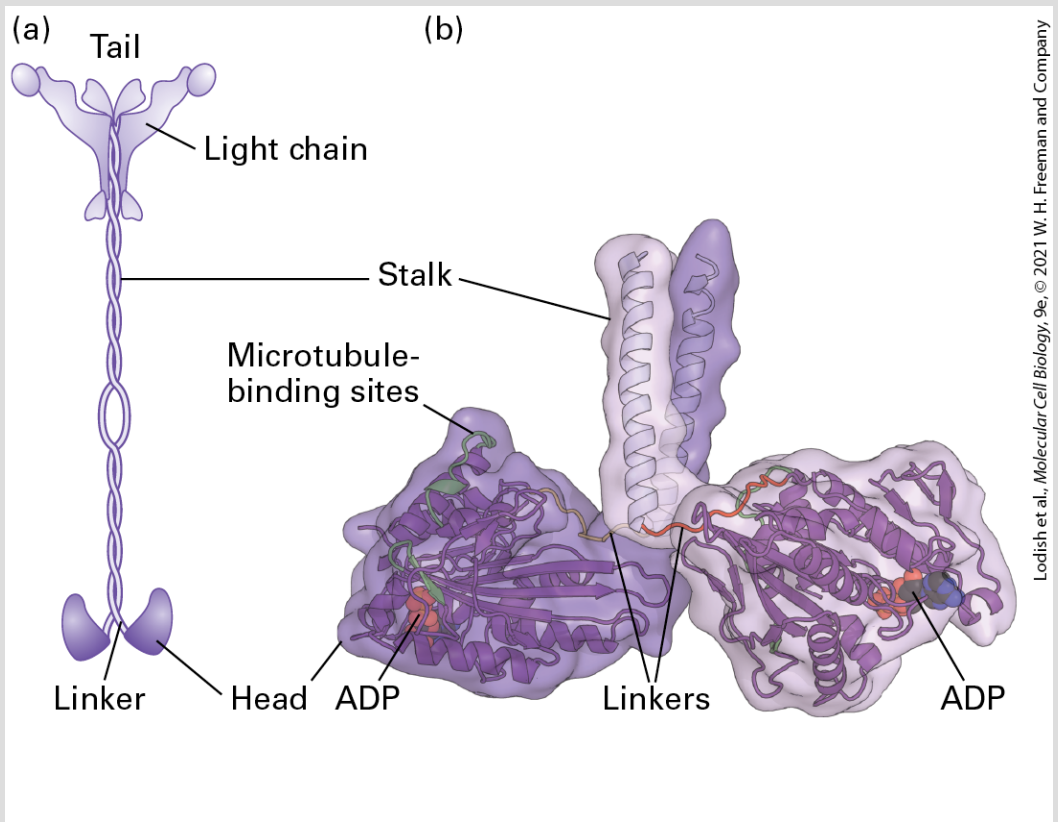


FIGURE 18-18 Structure of kinesin-1. (a) Representation of kinesin-1 showing the two intertwined heavy chains, each with a motor domain in the head region. Each head is attached to the coiled-coil stalk by a flexible linker domain. Two light chains associate with the tail of the heavy chain. See R. D. Vale, 2003, *Cell* **112**:467. (b) X-ray structure of the kinesin heads with the microtubule-binding and nucleotide-binding sites (containing ADP) indicated, including the linkers and the beginning of the stalk region. See M. Thormahlen et al., 1998, *J. Struct. Biol.* **122**:30.

[Data from F. Kozielski et al., 1997, *Cell* **91**:985–994, PDB ID 3kin.]

Kinesin-1-dependent movement of vesicles can be tracked by in vitro motility assays similar to those used to study myosin-dependent movements (see Figure 17-23). In one type of assay, a vesicle or a bead coated with kinesin-1 was added to a glass slide along with a preparation of stabilized microtubules and observed in a microscope. In the presence of ATP, the beads were seen to move in one direction along a microtubule. Researchers found that beads coated with kinesin-1 always moved from the (−) to the (+) end of a microtubule (Figure 18-19). Thus kinesin-1 is a (+) end–directed microtubule motor protein, and additional evidence shows that it mediates anterograde transport of organelles along the axon.

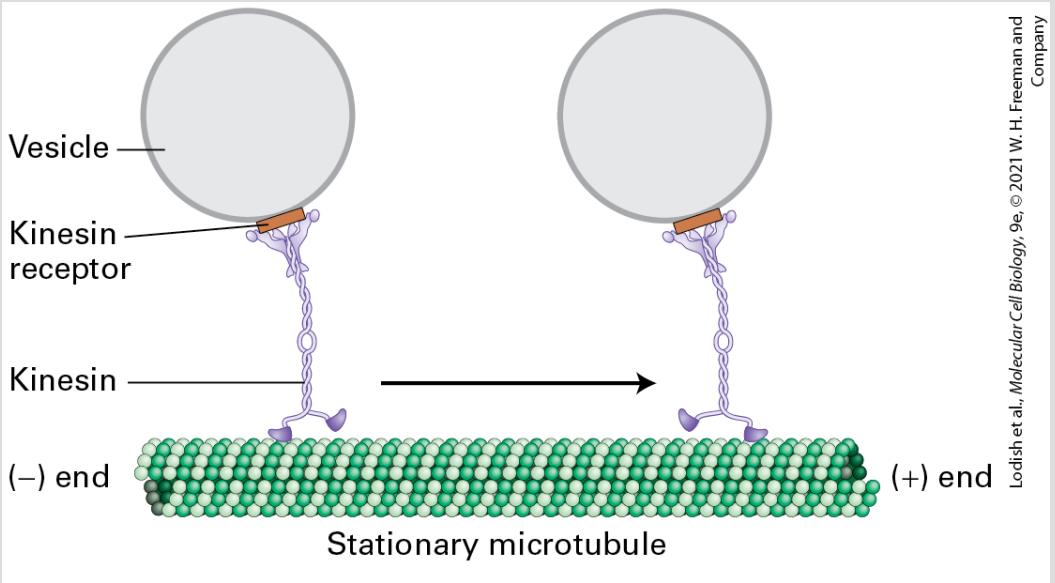


FIGURE 18-19 Model of kinesin-1-catalyzed vesicle transport. Kinesin-1 molecules, attached to receptors on the vesicle surface, transport the vesicles from the (−) end to the (+) end of a stationary microtubule. ATP is required for movement. See R. D. Vale et al., 1985, *Cell* **40**:559; and T. Schroer et al., 1988, *J. Cell Biol.* **107**:1785.

The Kinesins Form a Large Protein Superfamily with Diverse Functions

Following the discovery of kinesin-1, a number of proteins with similar motor domains were identified by both genetic screens and molecular biology approaches. There are now 14 known classes of kinesins in animals, defined as sharing amino acid sequence homology with the motor domain of kinesin-1. Proteins of the kinesin superfamily are encoded by about 45 genes in the human genome. Although the functions of all these proteins have not yet been elucidated, some are involved in organelle, mRNA, and

chromosome transport; microtubule sliding; and microtubule depolymerization.

As with the different classes of myosin motors, in the kinesin families the conserved motor domain is fused to a variety of class-specific nonmotor domains (Figure 18-20). Whereas kinesin-1 has two identical heavy chains and two identical light chains, members of the kinesin-2 family (also involved in organelle transport) have two different heavy-chain motor domains and a third polypeptide that associates with the tail and binds cargo. Members of the bipolar kinesin-5 family have four heavy chains, forming bipolar motors that can cross-link antiparallel microtubules and, by walking toward the (+) end of each microtubule, slide them past each other. The kinesin-14 motor proteins are the only known class to move toward the (−) end of a microtubule; this class functions in mitosis. Members of the kinesin-13 family have two subunits, but with the conserved kinesin domain in the middle of the polypeptide. Kinesin-13 proteins do not have motor activity; instead, they are special ATP-hydrolyzing proteins that can enhance the depolymerization of microtubule ends (see Figure 18-15).

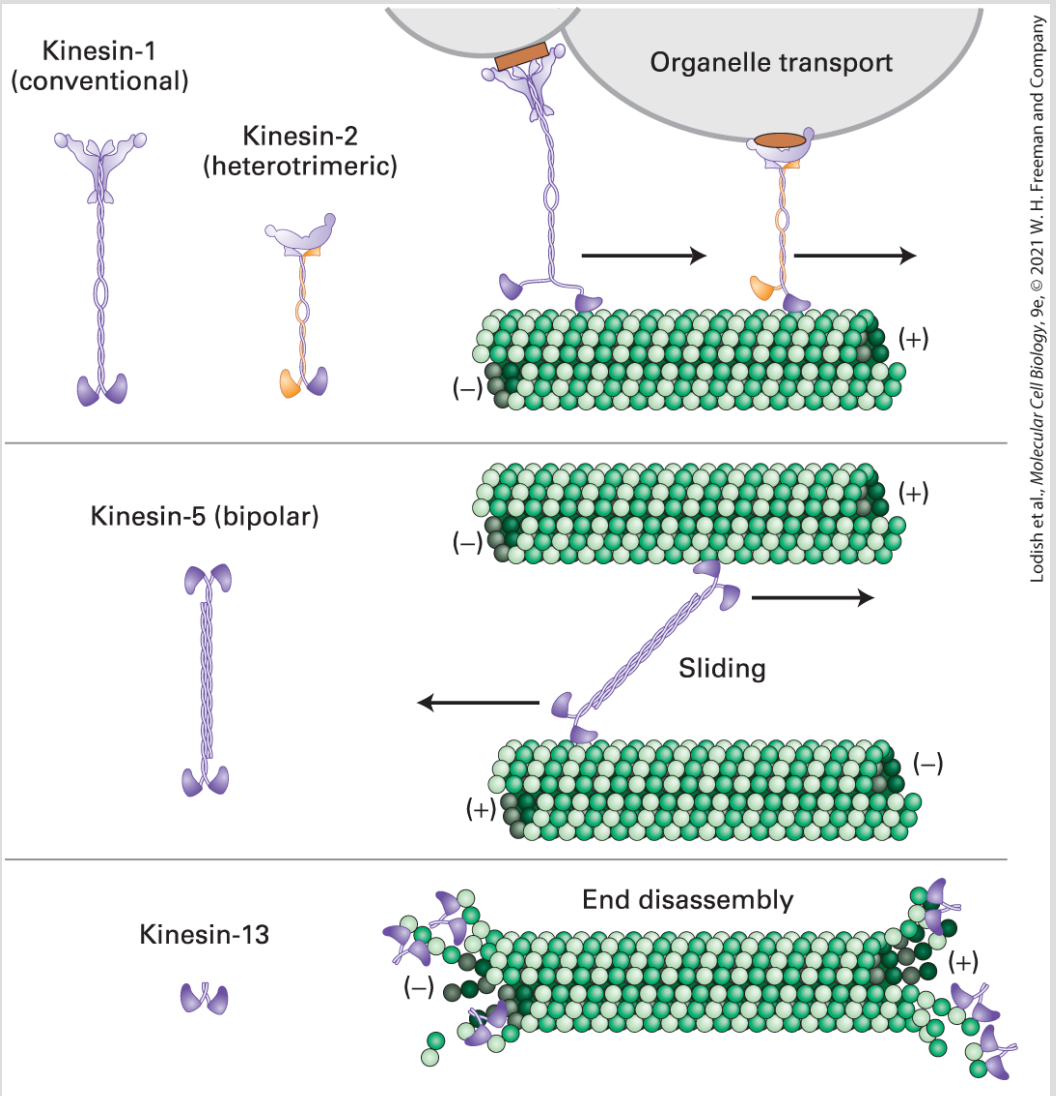


FIGURE 18-20 Structure and function of selected members of the kinesin superfamily. Examples of the structure and mode of action of some kinesin family members. Different kinesins have been given many different names; we use the unified nomenclature described in C. J. Lawrence et al., 2004, *J. Cell Biol.* **167**:19–22. See R. D. Vale, 2003, *Cell* **112**:467.

Kinesin-1 Is a Processive Motor

Optical trap and fluorescence-labeling techniques similar to those used to characterize myosin (see Figures 17-28 and 17-29) have been used to study how kinesin-1 moves down a microtubule and how ATP hydrolysis is converted into mechanical work. Such experiments demonstrate that kinesin-1 is a highly processive motor, meaning that it takes hundreds of steps down a microtubule without dissociating from it. During this process, the double-headed molecule moves along the same protofilament within the microtubule. The two heads work in a coordinated manner so that one is always attached to the microtubule.

The kinesin-1 heads bind a microtubule very tightly when they are in either an ATP-bound or nucleotide-free state, but binding is weak when ADP occupies the nucleotide-binding sites. It is these properties that account for the ability of kinesin-1 to move cargo in a processive manner. The ATP cycle of kinesin-1 movement is most easily understood if we begin just after the motor has taken a step (Figure 18-21a). At this point, the motor has a nucleotide-free leading head, which is strongly bound to a tubulin dimer in a protofilament, and an ADP-bound trailing head, which is weakly associated with the protofilament (Figure 18-21a, state **1**). ATP then binds to the leading head (Figure 18-21a, state **2**), and this binding induces a conformational change so that the neck linker of this leading head is no longer free and pointing backward, but now swings forward and *docks* into the head. This swinging motion causes the linker domain to rotate forward, which throws the

trailing head — like throwing a ballet dancer — into position to become the leading head (Figure 18-21a, state **3**). The new leading head finds the next binding site 16 nm ahead on the microtubule (Figure 18-21a, state **4** and Figure 18-21b). This step, which moves the cargo 8 nm, is the power stroke. The interaction of the leading head with the microtubule induces coordinated events between the two heads. First, it induces the leading head to release ADP to generate the nucleotide-free state and therefore bind the microtubule tightly. Second, it induces the trailing head to hydrolyze ATP to ADP and P_i, and upon releasing P_i transition into the weakly bound state (Figure 18-21a, step **5**). ATP can now bind to the leading head to repeat the cycle and allow the protein to take another step down the microtubule. During this cycle, one head is always in the ATP or nucleotide-free state, both of which result in tight binding to the microtubule, so kinesin very rarely dissociates as it walks along a microtubule, thereby accounting for its ability to take thousands of steps along a microtubule without disassociating from it.

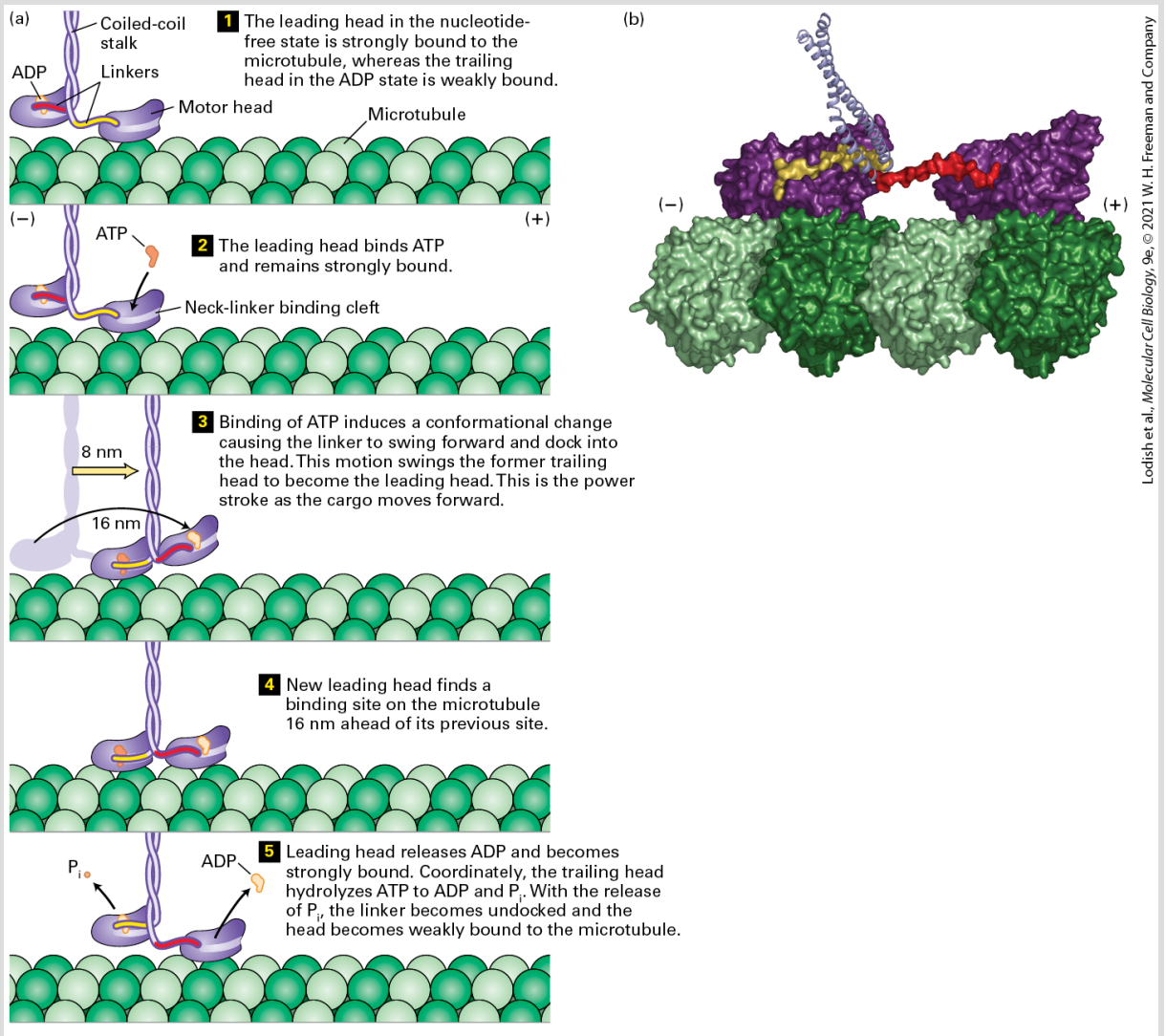


FIGURE 18-21 Kinesin-1 uses ATP to walk down a microtubule. (a) In this diagram, the two kinesin heads are shown with differently colored linker domains (yellow and red) to distinguish them. The cycle is shown starting after kinesin has taken a step, with the leading head in the nucleotide-free state and tightly bound to the microtubule, while the trailing head is weakly bound to the microtubule because it has ADP bound (state 1). The leading head then binds ATP (state 2), which induces a conformational change that causes the yellow linker region to swing forward and dock into its associated head domain, thereby thrusting the trailing head forward (state 3). The new leading head now finds a binding site 16 nm down the microtubule, to which it binds weakly (state 4). The leading head releases ADP and binds tightly to the microtubule, and coordinately the trailing head hydrolyzes ATP to ADP and P_i, and releases P_i. This converts the trailing head into a weak binding state, and also releases the docked linker region (state 5). The cycle now repeats

itself for another step. See R. D. Vale and R. A. Milligan, 2000, *Science* **288**:88. (b) Structural model of two kinesin heads (purple) bound to a protofilament in a microtubule. The trailing head, at left, has bound ATP and has thrust the other head into the leading position. Notice how the linker region (yellow) is docked into the trailing head, whereas the linker region (red) of the leading head is still free.

[Part (b) Data from E. P. Sablin and R. J. Fletterick, 2004, *J. Biol. Chem.* **279**:15707 and custom PDB files based on 3kin, 1mkj, and 1jff.]

As a transporter of organelles, kinesin-1 must bind and transport the correct cargo. It does this through receptor proteins on the appropriate organelle that bind to the tail domain of the motor protein. Since kinesin-1 is an ATPase, it is important that this activity be inactivated to conserve energy when it is not needed and also after releasing the organelle following transport. To accomplish this, kinesin-1 can fold into an inactive state in which the tail domain interacts with the motor head domain to inhibit both its ATPase activity and microtubule binding (Figure 18-22). When the tail binds the cargo-associated receptor protein, the motor unfolds and is activated (step **1**). When the cargo-activated motor complex encounters a microtubule, it transports the organelle (step **2**). Upon organelle delivery, kinesin-1 is released from the cargo and is inactivated by refolding into the head-to-tail inhibited state.

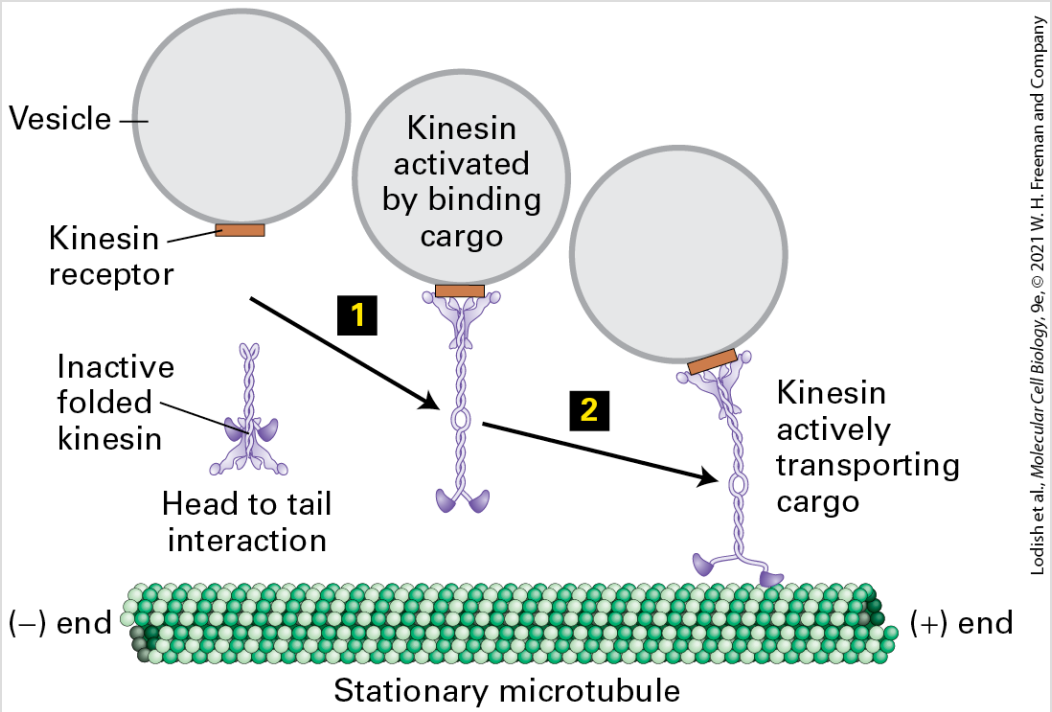


FIGURE 18-22 Kinesin-1 is regulated by a head-to-tail interaction. In its inhibited form, the head of kinesin-1 folds back and interacts with the tail. This interaction inhibits the ATPase activity of kinesin-1. When the motor encounters an appropriate receptor, here shown on a vesicle, it unfolds (step 1) and is now able to transport the cargo toward the (+) end of a microtubule (step 2). It is not yet clear how the motor dissociates from the cargo when they arrive at the destination, but it involves cargo release and folding back into the inhibited state.

When the x-ray crystallographic structure of the kinesin head was determined, it revealed a major surprise — the catalytic core has the same overall structure as myosin's (Figure 18-23)! This is the case despite the fact that there is no amino acid sequence conservation between the two proteins, arguing strongly that convergent evolution has twice generated a fold that can use the hydrolysis of ATP to generate work. Moreover, the same type of three-

dimensional structure is seen in small GTP-binding proteins, such as Ras, that undergo a conformational change upon GTP hydrolysis.

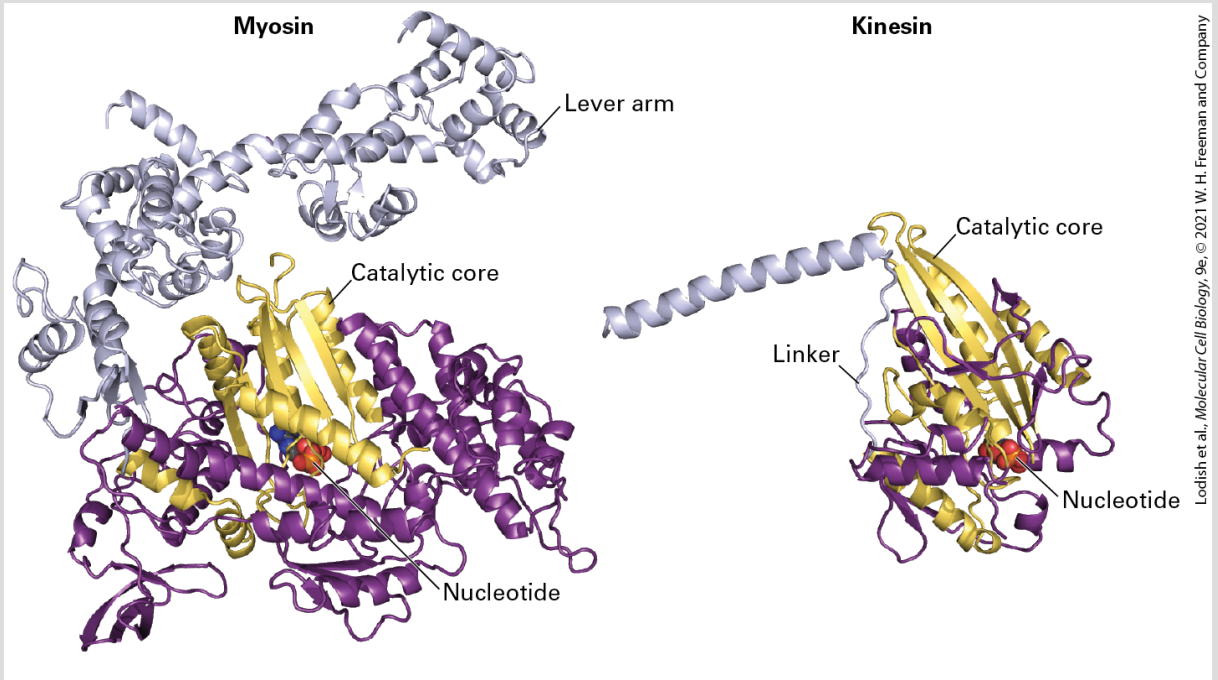


FIGURE 18-23 Convergent structural evolution of the ATP-binding cores of myosin and kinesin heads. The common catalytic cores of myosin and kinesin are shown in yellow, the nucleotide in red, and the lever arm (for myosin-II) and linker domain (for kinesin-1) in light purple. See R. D. Vale and R. A. Milligan, 2000, *Science* **288**:88.

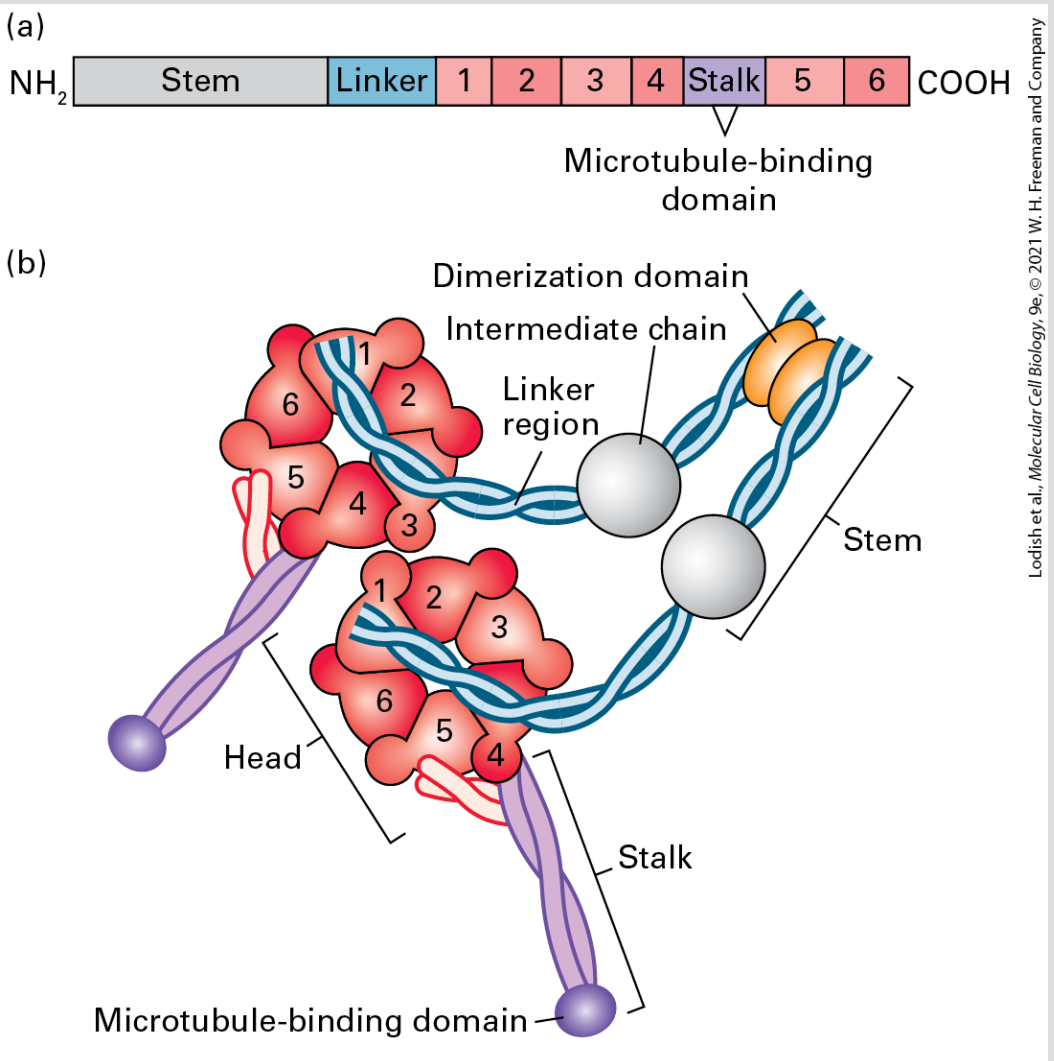
[Data from F. Kozielski et al., 1997, *Cell* **91**:985–994, custom PDB file (*left*) and PDB ID 3kin (*right*).]

Dynein Motors Transport Organelles Toward the (−) Ends of Microtubules

Cells use a different motor protein to transport organelles in a retrograde direction toward the (−) ends of microtubules.

Cytoplasmic dynein is very large; the human version is a 1.4-MDa complex consisting of two heavy, two intermediate, two light intermediate, and six light chains. In addition to its role in retrograde transport of organelles toward the (−) ends of microtubules in axons, dynein has many other functions, which we consider later in the chapter. Compared with the myosins and the kinesins, the family of dynein-related proteins is not very diverse. For example, only one gene encodes the human dynein heavy chain.

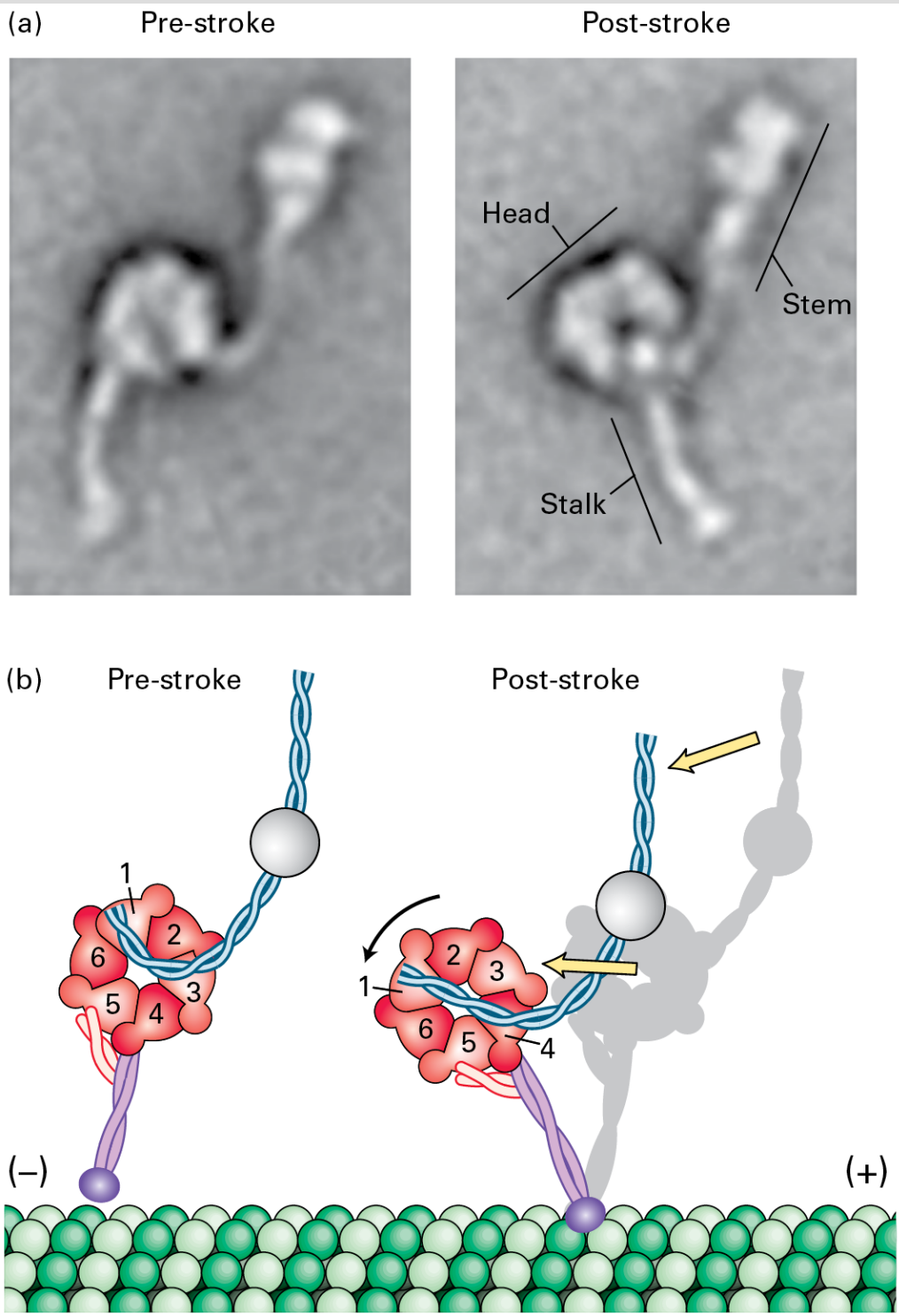
Like kinesin-1, cytoplasmic dynein is a two-headed molecule, built around two heavy chains. A single dynein heavy chain consists of a number of distinct domains (Figure 18-24a). The first is the *stem*, to which the other dynein subunits bind, and which associates with cargo through another protein complex, dynactin, that we discuss shortly. The next part of the heavy chain is a *linker* that plays a critical role during ATP-dependent motor activity. A large part of the heavy chain makes up the *head* containing the *AAA ATPase* domain, consisting of six repeats that assemble into a flowerlike structure (Figure 18-24b), within which lies the ATPase activity. Embedded between the fourth and fifth AAA repeats is the *stalk*, which protrudes from the structure and contains the microtubule-binding region.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-24 The domain structure of cytoplasmic dynein. (a) The dynein heavy chain, consisting of over 4000 amino acid residues, has several distinct domains. Following the stem and linker domains are six AAA repeats (peach, numbered 1–6), with the stalk and its microtubule-binding domain between repeats 4 and 5. (b) The six AAA repeats of each heavy chain assume a structure like petals on a flower. Emerging from this structure is a coiled-coil stalk domain with a microtubule-binding site at the end. A number of additional subunits associate with the stem region and can link dynein to cargo through dynactin (see Figure 18-26). See R. D. Vale, 2003, *Cell* **112**:467.

Electron microscopy, combined with x-ray images of the structure of a dynein heavy chain, provides a glimpse of how dynein might work. The first AAA repeat is probably the only one involved in converting the hydrolysis of ATP into mechanical work. In the absence of a bound nucleotide, dynein binds to microtubules. Note that the linker is straight, lies across the AAA domain, and associates with the first and fifth AAA repeats. Upon ATP binding, dynein dissociates from the microtubule and the linker becomes bent. The linker now crosses between the second and third AAA repeats ("pre-stroke"; Figure 18-25a and b, *left panels*). Interaction with a microtubule, ATP hydrolysis, and release of P_i cause the linker to straighten. This straightening is the power stroke that moves the cargo toward the (−) end of the microtubule ("post-stroke"; Figure 18-25a and b, *right panels*). The ADP is released, and the motor head remains bound to the microtubule. When ATP binds, the head is released, and the cycle repeats for another step.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-25 The power stroke of dynein. (a) Multiple images of purified single-headed dynein molecules in their pre-stroke and post-stroke states were recorded in an electron microscope and then averaged. The image at the left shows dynein in the pre-stroke state, and the image at the right shows it in the

post-stroke state. (b) A comparison of the microscopic images combined with structural data shows that the force-generation mechanism involves a change in the position of the linker, which causes a movement of the microtubule-binding stalk. For clarity, only one of the two dynein heads is shown. See G. Bhabha et al., 2014, *Cell* **159**:857.

[Part (a) Reprinted by permission from Nature Publishing Group, from S. A. Burgess et al., 2003, "Dynein Structure and Power Stroke," *Nature* **421**(6924):715–718; permission conveyed through Copyright Clearance Center, Inc.]

Unlike kinesin-1, cytoplasmic dynein cannot mediate cargo transport by itself. Rather, dynein-related transport requires a large protein complex, called *dynactin*, that links dynein to its cargo (Figure 18-26). Dynactin consists of 11 different types of subunits, functionally organized into two domains. One domain is built around one actin subunit and eight copies of the actin-related protein Arp1 assembled into a short filament. The end corresponding to the (+) end of this filament is capped by CapZ, the same capping protein that binds the (+) end of an actin filament (see Figure 17-12). A number of other subunits are associated with the (−) end. This Arp1-containing domain is responsible for binding cargo. The second domain of dynactin consists of a long protein called $p150^{Glued}$, which contains the dynein-binding site and has a microtubule-binding site at one end. Holding the two dynactin domains together is a protein called *dynamitin* — so named because when it is overexpressed, it dissociates (or "blows apart") the two domains, making a nonfunctional complex. This feature has been very useful experimentally because it has allowed researchers to

identify processes that are dependent on the interaction of dynein and dynactin, which are disrupted in cells overexpressing dynamitin.

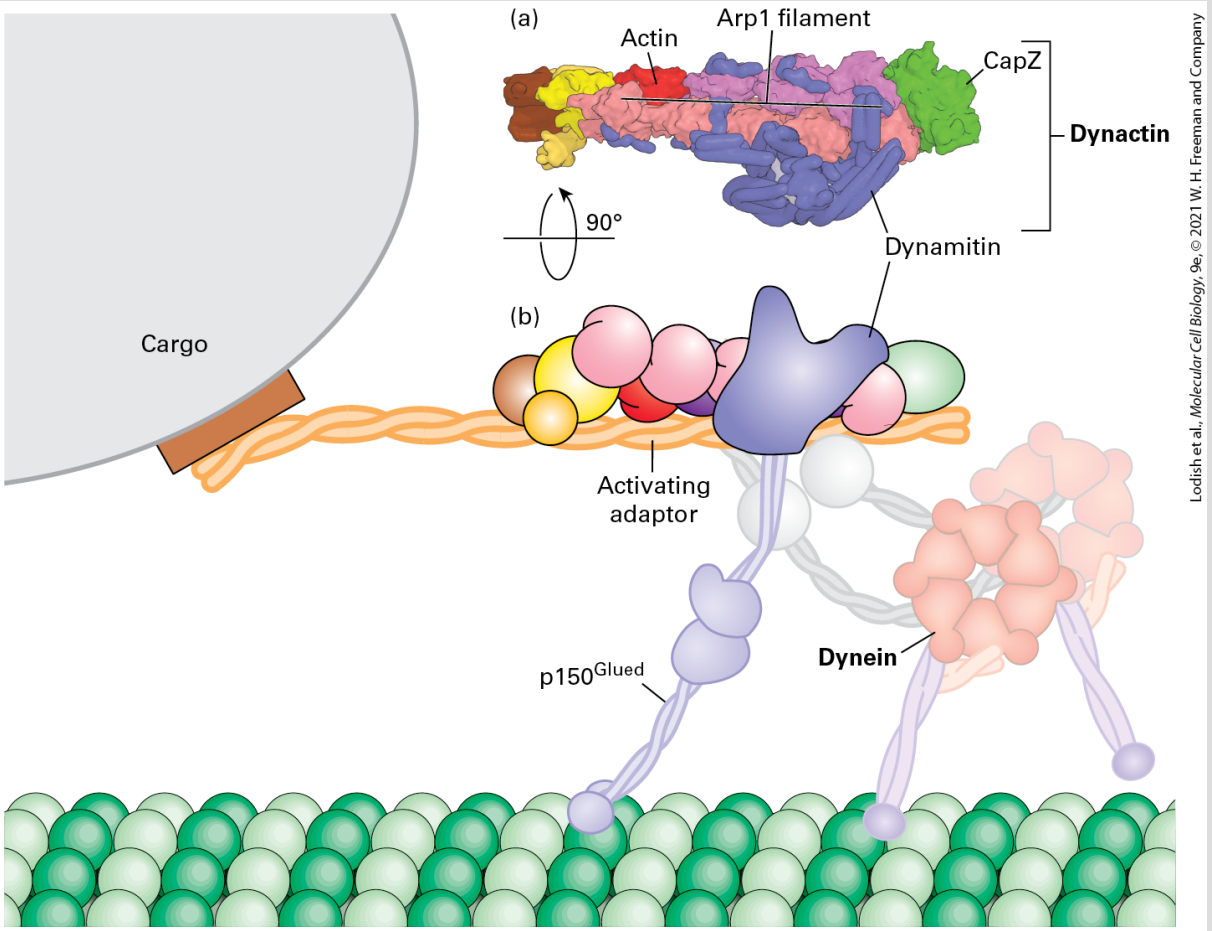
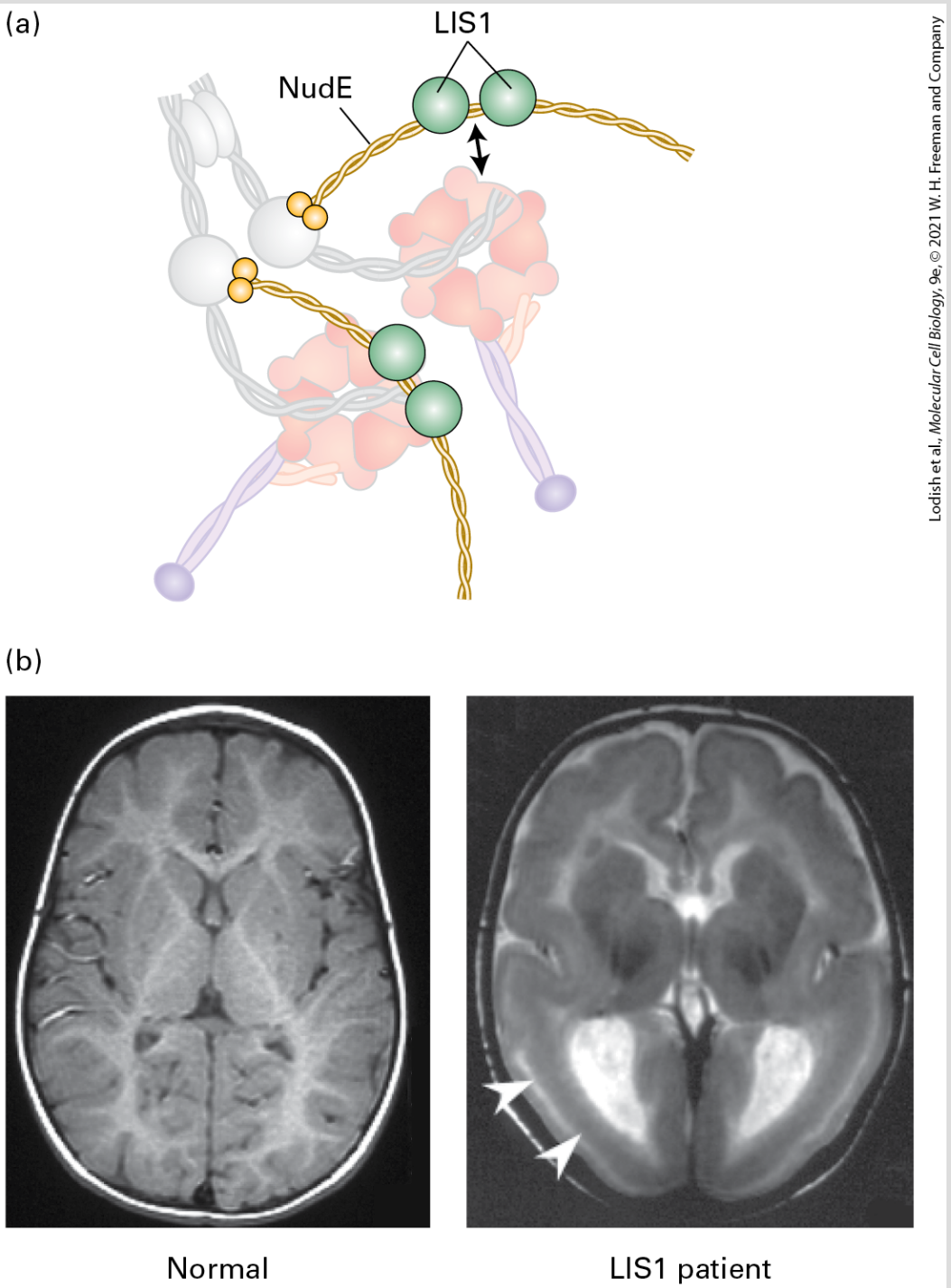


FIGURE 18-26 The dynactin complex functions with dynein to transport cargo. (a) One domain of the dynactin complex is built around a short filament made up of eight subunits of the actin-related protein Arp1 and one actin subunit, capped by CapZ. Another domain consists of the protein $p150^{Glued}$, which has a microtubule-binding site on its distal end and is also involved in attaching cytoplasmic dynein to the complex. Dynamitin holds the two parts of the dynactin complex together. (b) Diagram of how dynactin and dynein interact with each other and with an activating adaptor and a microtubule. See L. Urnavicius et al., 2015, *Science* **347**:1441.

Recent work has shown that the dynein-dynactin complex exists in an inactive form until it encounters activating adaptors, which both activate the motor and link it to cargo. There are many activating adaptors; each functions with a specific cargo. All of the adaptors are dimeric proteins with a distinct cargo-binding site and a long coiled-coil region that interacts with dynactin (Figure 18-26b). In some circumstances, dynein must be transported in an inactive form to a specific location, where it is then activated. For example, it has been found that the dynactin $p150^{Glued}$ subunit binds +TIP protein EB1, allowing dynein to become associated with the growing (+) end of a microtubule. Why would dynein associate with the microtubule (+) end if it is a (−) end–directed motor? Recent work suggests that when dynein is associated with the (+) end of a microtubule via the dynactin-EB1 interaction, it is held in an inactive conformation. When the growing microtubule reaches the cell cortex, the inactive dynein and dynactin encounter an activator that is localized there. The dynein now becomes active, associating with the cortex and pulling on the microtubule that delivered it there! This mechanism has been shown to help orient the mitotic spindle in yeast, and it probably applies to other situations as well.

In addition to dynactin, other proteins function with dynein. One group of proteins, which includes LIS1 and NudE, is involved in regulating the activity of dynein during brain development. NudE links the dynein intermediate and light chains to LIS1 (Figure 18-27a). LIS1 then interacts with the ATPase domain of dynein to

lengthen the duration of the power stroke, making the motor more processive under conditions of high load. Defects in LIS1 cause the fatal disease Miller-Dieker lissencephaly (from which the protein got its name); *lissencephaly* means "smooth brain," as cortical folds and grooves are lacking in the brains of patients with this condition (Figure 18-27b). Mutations in LIS1 result in defects in both neuronal mitoses and neuronal migration in early development, resulting in the smooth-brain phenotype and defects in mental development as well as many other abnormalities. ■



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-27 The LIS1 protein, which regulates dynein, is required for brain development. (a) Model of how NudE associates with dynein to allow LIS1 to interact with the ATPase domain of the dynein heavy chain. See R. J. McKenny et al., 2010, *Cell* **141**:304–314. (b) Magnetic resonance images (MRIs) of a normal brain (*left*) and a brain from a patient with Miller-Dieker

lissencephaly (*right*), which results from a lack of LIS1 function. Notice the absence of folding in the patient's brain at the locations marked by the arrows.

[Part (b) Reprinted with permission from Oxford University Press, from K. Mitsuhiro and W. B. Dobyns, 2003, "Lissencephaly and the Molecular Basis of Neuronal Migration," *Hum. Mol. Genet*. **12**:R89–R96; permission conveyed through Copyright Clearance Center, Inc.]

Kinesins and Dyneins Cooperate in the Transport of Organelles Throughout the Cell

Dyneins and kinesins play important roles in the microtubule-dependent organization of organelles in cells (Figure 18-28). Recall that the orientation of microtubules is fixed by the microtubule organizing center, or MTOC. So the direction of transport — toward or away from the cell center — depends on the motor proteins. For example, the Golgi complex collects in the vicinity of the centrosome, where the (−) ends of microtubules lie, and is driven there by dynein-dynactin. When secretory cargo emerges from the endoplasmic reticulum, it is transported to the Golgi by dynein-dynactin. The endoplasmic reticulum is spread throughout the cytoplasm and is transported there by kinesin-1, which moves toward the (+) ends of microtubules, located toward the cell's periphery. Some organelles of the endocytic pathway, including late endosomes and lysosomes, are associated with dynein-dynactin complexes. Kinesins transport mitochondria as well as

nonmembranous cargo such as mRNAs encoding proteins that need to be localized to specific regions of the cell during development.

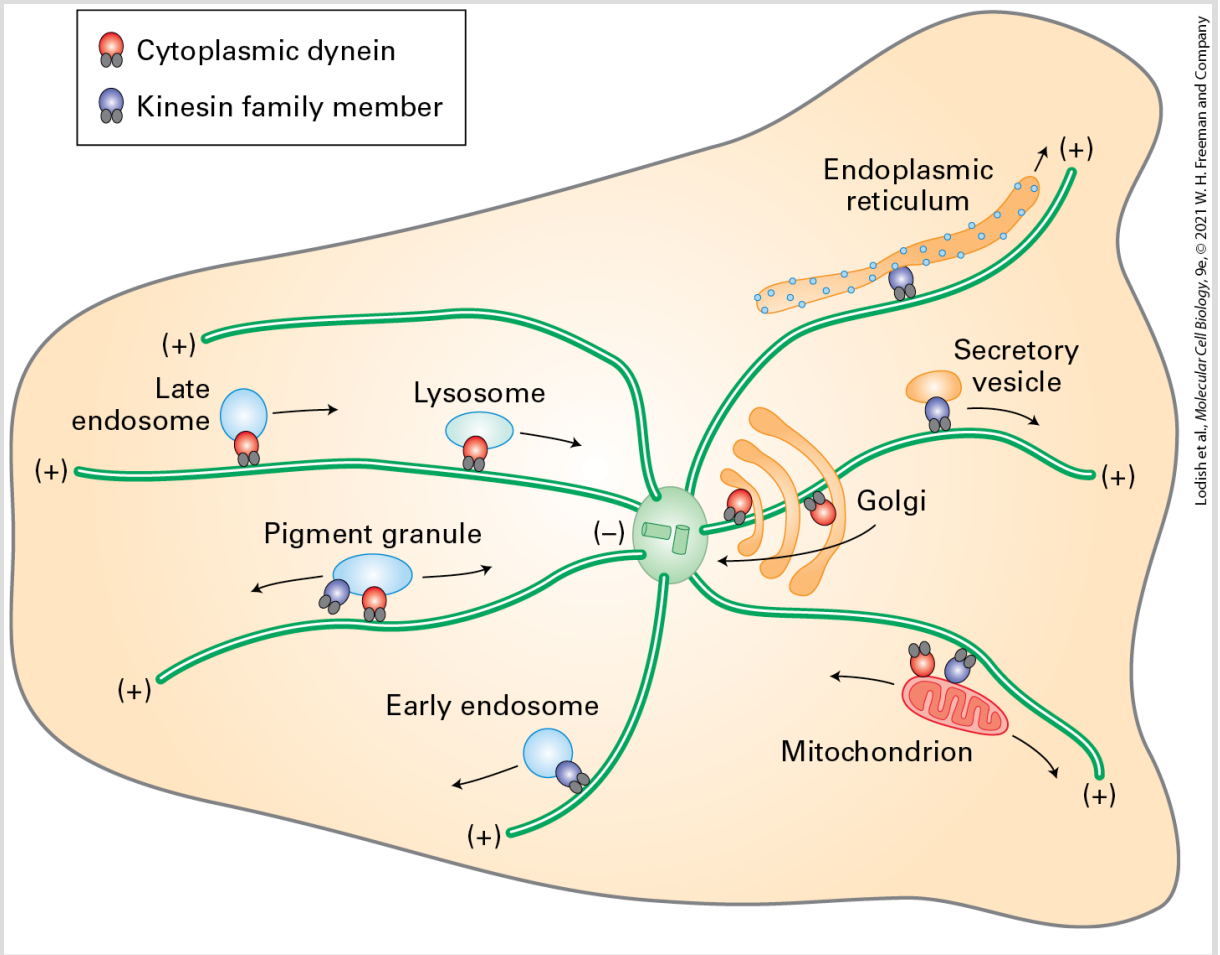
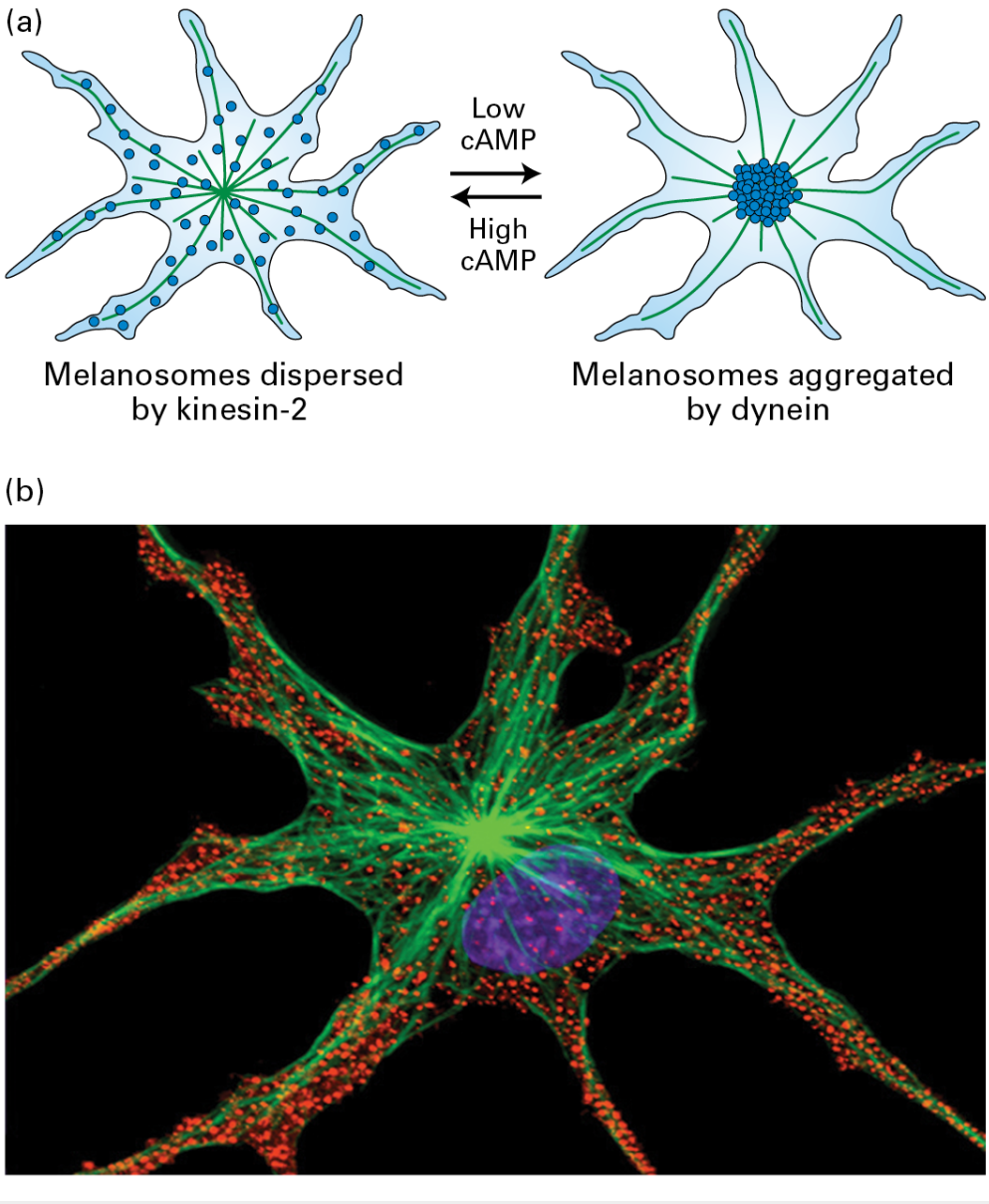


FIGURE 18-28 Organelle transport by microtubule motors. Cytoplasmic dyneins (red) mediate retrograde transport of organelles toward the (−) ends of microtubules (cell center); kinesins (purple) mediate anterograde transport toward the (+) ends (cell periphery). Most organelles have one or more microtubule-based motors associated with them. It should be noted that the association of motors with organelles varies by cell type, so some of these associations may not exist in all cells, whereas others not shown here also exist.

We have seen how kinesin-1 transports organelles in an anterograde direction down axons. What happens to the motor when it gets to

the end of the axon? It is carried back in a retrograde direction on organelles transported by cytoplasmic dynein. Thus kinesin-1 and dynein can associate with the same organelle, and a mechanism must exist that turns one motor off while activating the other, although these mechanisms are not yet understood fully.

Much of what we know about the regulation of microtubule-based organelle transport comes from studies using fish (e.g., angelfish) or frog melanophores. Melanophores are cells of the vertebrate skin that contain hundreds of dark, melanin-filled pigment granules called *melanosomes*. Melanophores either have their melanosomes dispersed, in which case they make the skin darker, or aggregated at the cell center, which makes the skin paler (Figure 18-29). These changes in skin color, mediated by neurotransmitters in the fish and regulated by hormones in the frog, serve to camouflage the fish and enhance social interactions in the frog. The movement of the melanosomes is mediated by changes in intracellular cAMP and is dependent on microtubules. Studies investigating which motors are involved have shown that melanosome dispersion requires kinesin-2, whereas melanosome aggregation requires cytoplasmic dynein-dynactin. The first hints of how these activities might be coordinated came from the finding that overexpression of dynamitin inhibited melanosome transport in both directions. This surprising result was explained when it was found that dynactin can bind not only to cytoplasmic dynein but also to kinesin-2 — and may coordinate the activity of the two motors.



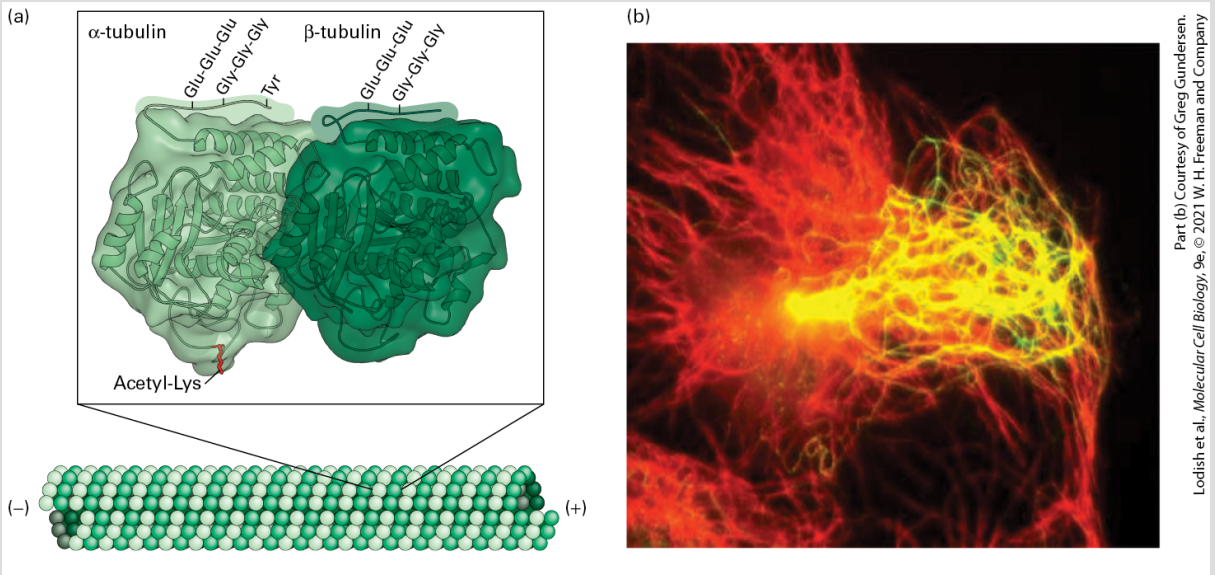
Part (b) courtesy of Steve Rogers.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-29 Movement of pigment granules in frog melanophores. (a) Diagram of the microtubule-based reorganization of melanosomes according to the level of cAMP. Melanosomes are aggregated by cytoplasmic dynein and dispersed by kinesin-2. (b) Melanosomes (red) in the dispersed state as seen by immunofluorescence microscopy. Microtubules are labeled green, and the DNA in the nucleus is labeled blue.

The association of dynein and kinesin-2 with the same organelle is not limited to melanosomes. It has been suggested that these motors may cooperate to appropriately localize late endosomes, lysosomes, and mitochondria in some cells. Thus the association of organelles with a number of distinct motors is not the exception but an emerging theme.

Tubulin Modifications Distinguish Different Classes of Microtubules and Their Accessibility to Motors

The stability and functions of different classes of microtubules are influenced by post-translational modifications to tubulin. Although multiple types of modifications have been detected, we restrict our discussion here to those that are the best understood — lysine acetylation, detyrosylation, polyglutamylation, and polyglycylation (Figure 18-30a).



EXPERIMENTAL FIGURE 18-30 Post-translational modifications of tubulin affect the stability and function of microtubules. (a) Structure of α- and β-tubulin, showing the sites of lysine acetylation on the inner surface of the microtubule and polyglutamylation, polyglycylation, and detyrosylation on the outer surface. Note that polyglutamylation and polyglycylation are likely mutually exclusive, so they would not normally occur at the same time. See J. W. Hamond, D. Cai, and K. J. Verhey, 2008, *Curr. Opin. Cell Biol.* **20**:71–76. (b) Detyrosylated microtubules are preferentially oriented toward the leading edge of a moving cell. A cell migrating toward the right is stained for total microtubules (red) and detyrosylated microtubules (green). The resulting merged image shows the detyrosylated microtubules enriched toward the front of the cell in yellow, which is a combination of red and green.

[Part (a) data from E. Nogales, S. G. Wolf, and K. H. Downing, 1998, *Nature* **391**:199–203, PDB ID 1tub.]

Two of these modifications are found only on α-tubulin. The first is *acetylation* of the ε-amino group of a specific lysine residue of α-tubulin that lies on the inside of the microtubule. Microtubules with this acetylated lysine are stabilized against breakage and are found on microtubule structures such as centrioles, basal bodies, and primary cilia (primary cilia are discussed in Section 18.5). Indeed,

cells unable to acetylate tubulin have defective primary cilia, whereas cells in which the acetylation cannot be removed have unusually stable primary cilia. The second modification of α-tubulin involves its C-terminal tyrosine. This tyrosine can be removed by a carboxypeptidase that functions only when bound to the surface of a microtubule, where it sequentially removes the C-terminal tyrosines from α-tubulin subunits. Such *detyrosylated* microtubules are more stable, because they are more resistant to depolymerization by kinesin-13 and related proteins. Moreover, in migrating cells, these more stable microtubules are generally oriented toward the front of the cell (Figure 18-30b). When a detyrosylated microtubule depolymerizes, the α-tubulin subunit of the αβ-dimer has the C-terminal tyrosine added back by a tyrosine ligase that acts only on soluble tubulin, and the αβ-tubulin dimer can now be used during elongation of another growing microtubule.

The C-terminal regions of both α- and β-tubulin are rich in glutamic acid residues, and specific enzymes can modify these residues. These modifications occur only after assembly of the microtubule. The tubulin tails can be modified by *polyglutamylation*, in which a chain of glutamic acid residues is linked to a specific glutamate residue, or by *polyglycylation*, in which a chain of glycine residues is added to a different glutamate residue. These two modifications are likely mutually exclusive, so that a tubulin subunit modified by polyglycylation is protected from polyglutamylation, and vice versa.

Like detyrosylation, both polyglycylation and polyglutamylation can enhance microtubule stability.

These post-translational modifications of tubulin not only affect microtubule stability but can also influence the ability of molecular motors to interact with microtubules. Kinesin-1 associates preferentially with detyrosylated and acetylated microtubules, so these modifications may be important in recruiting this motor for axonal transport in neurons. As is illustrated in Figure 18-5e, neurons have different microtubule organizations in their dendrites and their axons. The microtubules in the axon are stabilized by acetylation and detyrosylation, which allows kinesin-1 to associate preferentially with them for axonal transport. Polyglutamylation has a key role in the beating of cilia and flagella, which we discuss in the next section.

Research elucidating the effects of post-translational modifications of tubulin on microtubule function and microtubule-based motors is all quite recent. We can expect future studies to reveal so-called codes used by the cell that distinguish among different classes of microtubules and impart specific functions to the different types of microtubules.

KEY CONCEPTS OF SECTION 18.4

Kinesins and Dyneins: Microtubule-Based Motor Proteins

- Kinesin-1 is a microtubule $(+)$ end–directed ATP-dependent motor protein that transports membrane-bounded organelles (see Figure 18-19).
- Kinesin-1 consists of two heavy chains, each with an N-terminal motor domain, and two light chains that associate with cargo (see Figure 18-18).
- The kinesin superfamily includes motors that function in interphase and mitotic cells, transporting organelles and sliding antiparallel microtubules past each other. The superfamily includes one class, kinesin-13, that is not motile but instead destabilizes microtubule ends (see Figure 18-20).
- Kinesin-1 is a highly processive motor because it coordinates ATP hydrolysis between its two heads so that one head is always firmly bound to a microtubule (see Figure 18-21).
- Kinesin-1 can exist in an inactive folded conformation or in a cargo-associated extended conformation (see Figure 18-22).
- Cytoplasmic dynein is a microtubule $(-)$ end–directed ATP-dependent motor that associates with the dynactin complex and activating adapters to transport cargo (see Figures 18-24 and 18-26).
- Kinesins and dyneins associate with many different organelles to organize their locations in cells (see Figure 18-28).
- Post-translational modifications of tubulin can affect the stability of microtubules and regulate their ability to interact with microtubule-based motors.

18.5 Cilia and Flagella: Microtubule-Based Surface Structures

Cilia and **flagella** are microtubule-based and membrane-enveloped extensions of the plasma membrane that project from many protozoan and most animal cells. Abundant motile cilia are found on the surfaces of specific epithelia, such as those that line the trachea (see Figure 4-35), where they beat in an orchestrated, wavelike fashion to move fluids. Animal-cell flagella, which are longer than cilia but have a similar structure, can propel a cell, such as a sperm, through liquid. Cilia and flagella contain many different microtubule-based motors: axonemal dyneins are responsible for the beating of flagella and cilia, whereas kinesin-2 and cytoplasmic dynein are responsible for flagellum and cilium assembly and turnover.

Eukaryotic Cilia and Flagella Contain Long Doublet Microtubules Bridged by Dynein Motors

Cilia and flagella range in length from a few micrometers to more than 2 mm for some insect sperm flagella. They possess a central bundle of microtubules, called the **axoneme**, which consists of a

9 + 2 arrangement of nine doublet microtubules surrounding a central pair of singlet microtubules (Figure 18-31a, c). Each of the nine outer doublets consists of an A microtubule with 13 protofilaments and a B microtubule with 10 protofilaments (see Figure 18-4). All the microtubules in cilia and flagella have the same polarity: the (+) ends are located at the distal tip. At its point of attachment in the cell, the axoneme connects with the *basal body*, a complicated structure containing nine triplet microtubules (Figure 18-31a, b).

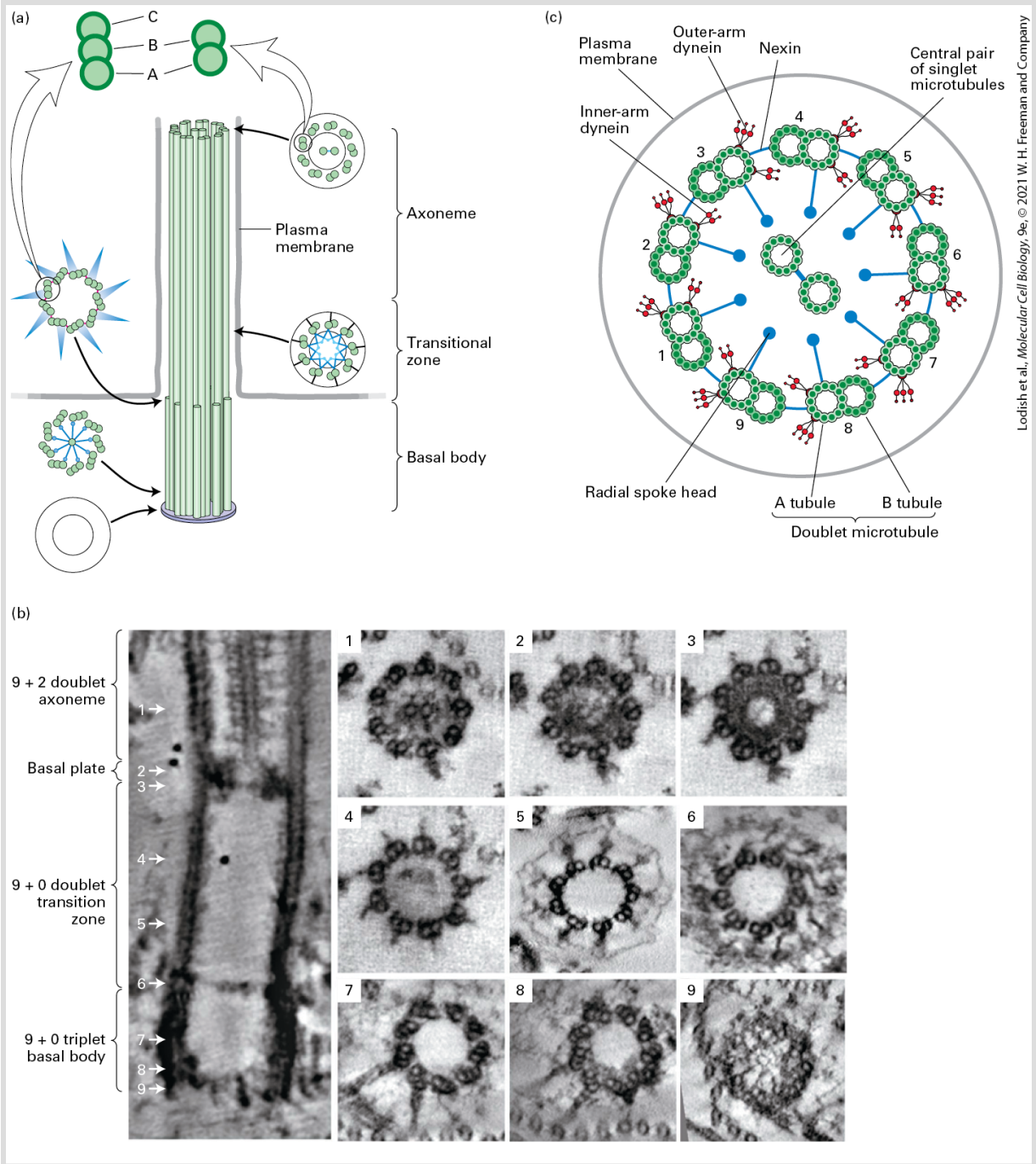


FIGURE 18-31 Structural organization of cilia and flagella. (a) Cilia and flagella are assembled from a basal body, a structure built around nine linked triplet microtubules. Continuous with the A and B microtubules of the basal body are the A and B tubules of the axoneme — the membrane-enveloped core of the cilium or flagellum. Between the basal body and axoneme is the transitional zone. (b) Electron micrograph of a longitudinal section (*left*) and indicated transverse sections (*right*) through a basal body and part of the axoneme of a cilium showing the intricate structure of a basal body, the transitional zone, and

axoneme. (c) Diagram of a transverse section of a cilium, to show the identity of the structures. See S. K. Dutcher, 2001, *Curr. Opin. Cell Biol.* **13**:49–54.

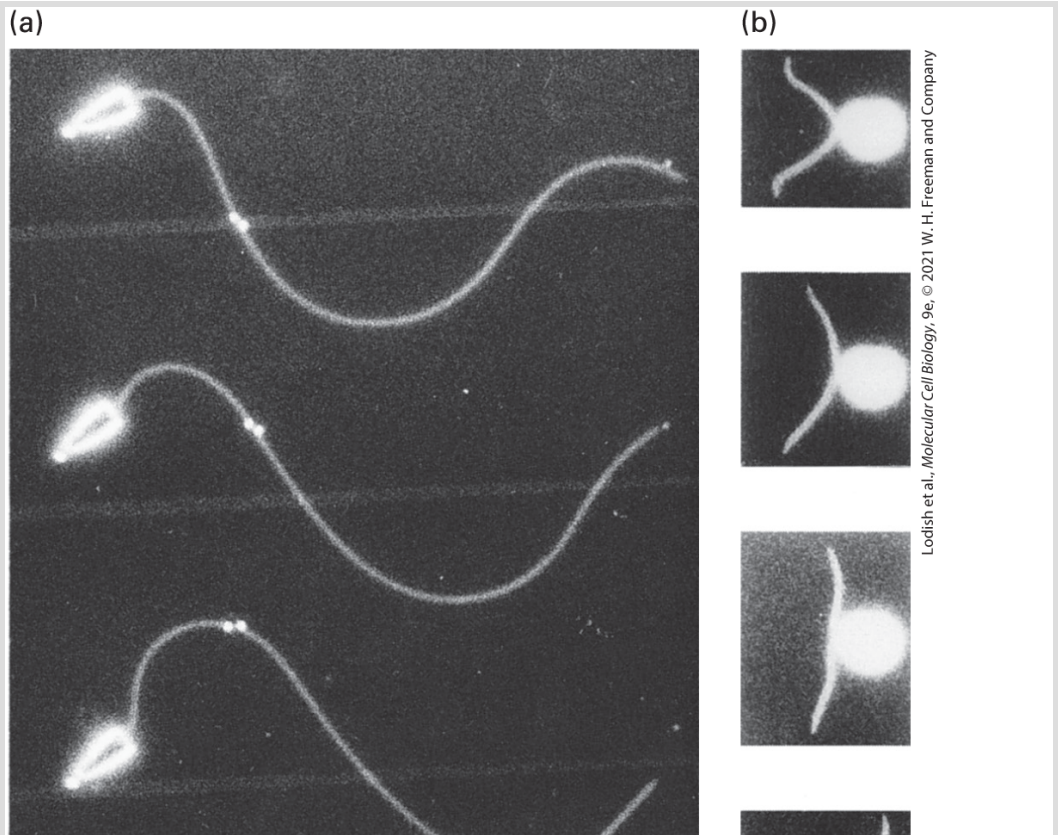
[Part (b) S. Vaughan, K. Gull, 2015, "Basal Body Structure and Cell Cycle-Dependent Biogenesis in *Trypanosoma brucei*," *Cilia* **5**:5, https://doi.org/10.1186/s13630-016-0023-7.]

The axoneme is held together by three sets of protein cross-links (Figure 18-31c). The two central singlet microtubules are connected to each other by periodic bridges, like rungs on a ladder. A second set of linkers, composed of the protein *nexin*, joins adjacent outer doublet microtubules to each other. *Radial spokes* project from each A tubule of the outer doublets toward the central pair.

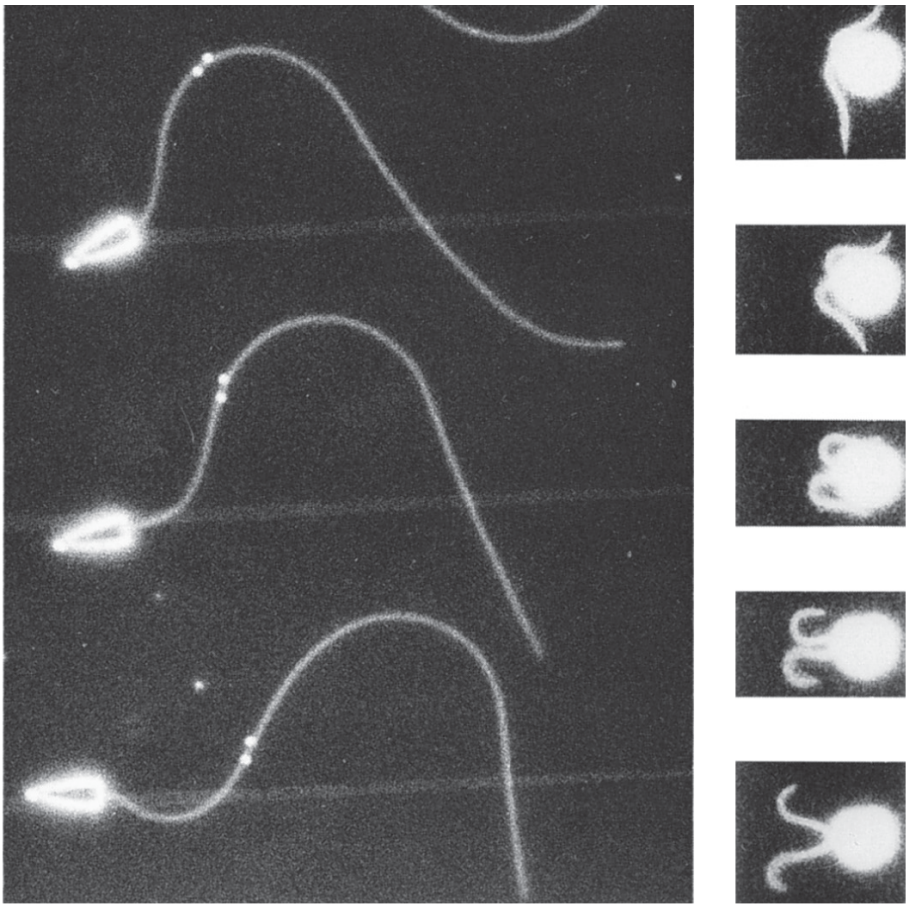
The major motor protein present in cilia and flagella is *axonemal dynein*, a large, multisubunit protein related to cytoplasmic dynein. Two rows of axonemal dynein motors are attached periodically down the length of each A tubule of the outer doublet microtubules; these motors are called the *inner-arm* and *outer-arm* dyneins (see Figure 18-31c). It is the interaction of these dynein motors with the B tubule in the adjacent doublet that brings about ciliary and flagellar bending.

Ciliary and Flagellar Beating Are Produced by Controlled Sliding of Outer Doublet Microtubules

Activation of the axonemal dynein motors causes cilia and flagella to bend. A close examination of this movement using live cell imaging reveals that a bend starts at the base of a cilium or flagellum and then propagates along the structure (Figure 18-32). A clue to how this occurs came from studies of isolated axonemes. In these classic experiments, axonemes were gently treated with a protease that cleaves only the nexin linkers. When ATP was added to the treated axonemes, the doublet microtubules slid past one another as dynein, attached to the A tubule of one doublet, walked down the B tubule of the adjacent doublet (Figure 18-33b, c). In an axoneme with intact nexin linkers, the action of dynein induces flagellar bending because the microtubule doublets are connected to one another (Figure 18-33a).

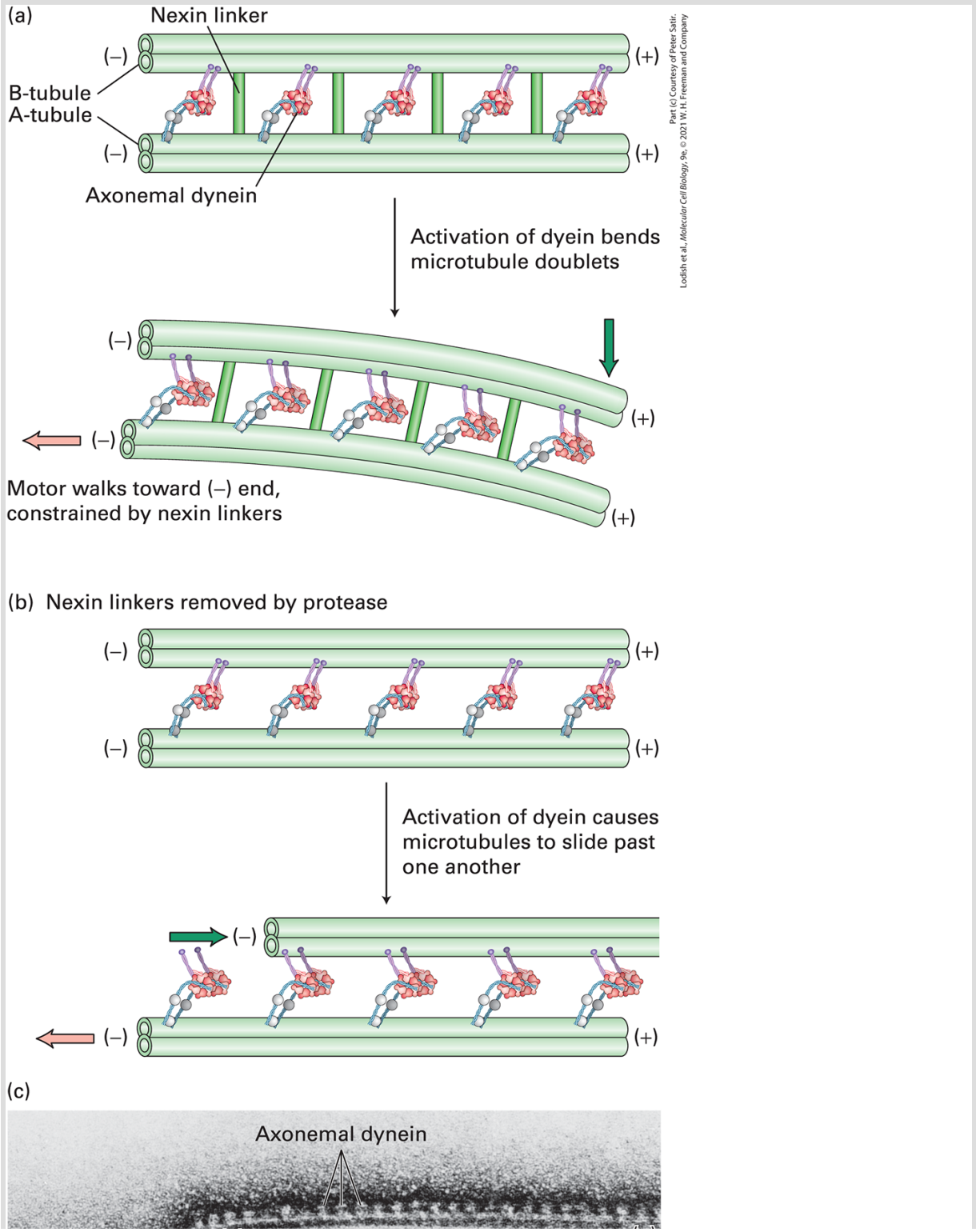


Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company



EXPERIMENTAL FIGURE 18-32 Live cell imaging shows flagellar movements that propel sperm and *Chlamydomonas* forward. In both cases, the cells are moving to the left. (a) In the typical sperm flagellum, successive waves of bending originate at the base and are propagated out toward the tip; these waves push against the surrounding fluid and propel the cell forward. In this sequence, a bend at the base of the flagellum in the first (*top*) frame has moved distally halfway along the flagellum by the last frame. A pair of gold beads attached to the flagellum are seen to slide apart as the bend moves through their region. (b) Beating of the two flagella on *Chlamydomonas* occurs in two stages, called the *effective stroke* (top three frames) and the *recovery stroke* (remaining frames). The effective stroke pulls the organism through the water. During the recovery stroke, a different wave of bending moves outward from the bases of the flagella, pushing the flagella along the surface of the cell until they reach the position to initiate another effective stroke. Beating commonly occurs 5–10 times per second.

[Part (a) C. J. Brokaw, 1991, *J. Cell Biol.* **114**(6):1201–1215; https://doi.org/10.1083/jcb.114.6.1201.]



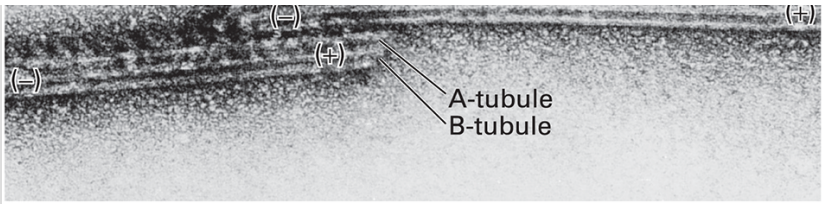


FIGURE 18-33 Ciliary and flagellar bending is mediated by axonemal dynein. (a) Axonemal dynein attached to an A tubule of an outer doublet pulls on the B tubule of the adjacent doublet trying to move toward the (−) end. Because the adjacent tubules are tethered by nexin, the force generated by dynein bends the cilium or flagellum. (b) Experimental evidence for the model in (a). When the nexin linkers are cleaved with a protease and ATP added to induce dynein activity, the microtubule doublets slide past one another. (c) Electron micrograph of two doublet microtubules in a protease-treated axoneme incubated with ATP. In the absence of cross-linking proteins, doublet microtubules slide excessively. The dynein arms can be seen projecting from A tubules and interacting with B tubules of the left microtubule doublet.

We don't yet fully understand how specific subsets of dynein are activated and how a wave of activation is propagated down the axoneme, but post-translational modifications of tubulin may play a role. Recall from Section 18.4 that post-translational modifications of tubulin subunits can affect the interactions between microtubules and motor proteins. The B tubules of the outer axoneme doublets are often polyglutamylated, and this modification strongly affects the interaction of the inner-arm dynein with the B tubule. Because inner-arm dynein motors mainly affect the waveform of the ciliary beat, it is this aspect of ciliary function that is compromised in mutants unable to undergo polyglutamylation.

Intraflagellar Transport Moves Material Up and Down Cilia and Flagella

Although axonemal dynein is involved in bending cilia and flagella, another type of motility has been observed in these structures. Careful examination of flagella on the biflagellate green alga *Chlamydomonas reinhardtii* revealed cytoplasmic particles moving at about 2.5 μm/s toward the tip of a flagellum (anterograde movement) and other particles moving at about 4 μm/s from the tip to the base (retrograde movement). This movement, known as *intraflagellar transport (IFT)*, occurs in both cilia and flagella. Light and electron microscopy revealed that the particles move along a path that runs parallel to the outer doublet microtubules and the plasma membrane (Figure 18-34). Analysis of algal mutants demonstrated that the anterograde movement is powered by kinesin-2 and the retrograde movement by cytoplasmic dynein. Remarkably, anterograde-moving particles can move past retrograde-moving particles without interference. This is because the particles move along distinct microtubules of the outer doublets: the anterogade particles powered by kinesin-2 move along the B-tubule and the retrograde particles powered by dynein move along the A-tubule.

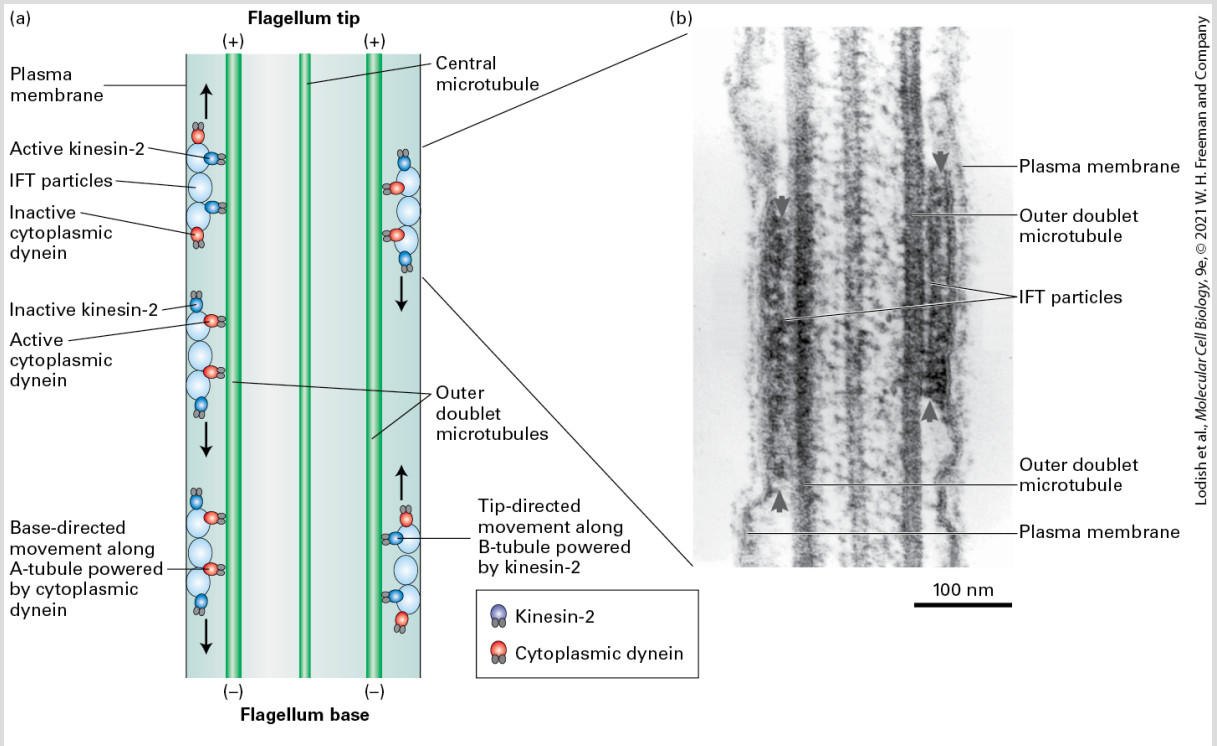


FIGURE 18-34 Intraflagellar transport. (a) Particles are transported between the plasma membrane and the outer doublet microtubules. Transport of the particles to the tip is dependent on kinesin-2, whereas transport toward the base is mediated by cytoplasmic dynein. (b) This thin-section electron micrograph shows IFT particles in a section of a *Chlamydomonas* flagellum.

[Part (b) Reprinted by permission from Nature Publishing Group, from J. L. Rosenbaum and G. B. Witman, 2002, "Intraflagellar Transport," *Nat. Rev. Mol. Cell Biol.* **3**:813–825; permission conveyed through Copyright Clearance Center, Inc.]

The anterograde and retrograde IFT particles transported in *Chlamydomonas* flagella have been isolated and their composition determined. They consist of two distinct protein complexes, called *IFT complex A* and *IFT complex B*. By analyzing the phenotypes of cells having mutations affecting these complexes, it has been found that complex B is necessary for anterograde IFT, whereas complex A is important for retrograde IFT. Despite this segregation of function,

both complexes are transported in both directions. All the components of IFT particles have homologs in organisms containing cilia, such as nematodes, fruit flies, mice, and humans, but these particles are absent from the genomes of yeasts and plants that lack cilia, suggesting that they are specific to IFT.

What is the function of IFT? Because all the microtubules in a flagellum have their growing (+) ends at the tip, that is the site where new tubulin subunits and flagellar structural proteins are added. Moreover, even in cells with flagella of uniform length, the microtubules are turning over, with assembly and disassembly occurring at the flagellar tip. In cells defective for kinesin-2, flagella shrink, suggesting that IFT transports new material to the tip for growth. Since IFT is a continually occurring process, what happens to the kinesin-2 molecules when they get to the tip, and where do the dynein motors come from to transport the particles retrogradely? Remarkably and similar to the situation in axonal transport, dynein is carried to the tip as cargo on the anterograde-moving particles, powered by kinesin-2, and then kinesin-2 becomes cargo as the particles are transported back to the base by dynein.

Primary Cilia Are Sensory Organelles on Interphase Cells

Many vertebrate cells bear a solitary nonmotile cilium known as the **primary cilium**. The primary cilium is a stable structure that is resistant to drugs such as colchicine that disassemble most microtubules. After colchicine treatment, the only remaining microtubules are found in the centrioles and the primary cilium (see Figure 18-12). Moreover, the tubulin in the primary cilium is highly acetylated, so using antibodies that specifically recognize acetylated α-tubulin readily identifies the single primary cilium on each interphase cell (Figure 18-35a).

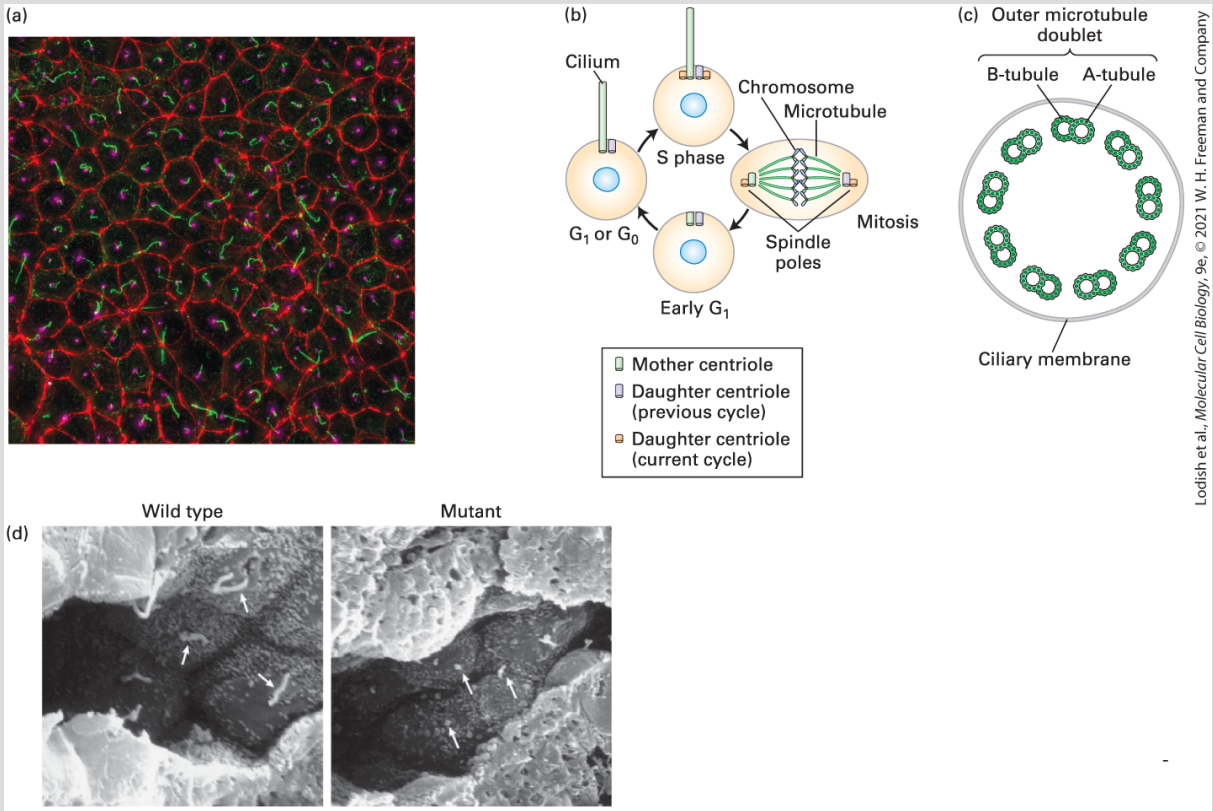


FIGURE 18-35 Many interphase cells contain a nonmotile primary cilium. (a) Fluorescence micrograph of mouse epithelial cells stained with antibodies to acetylated α-tubulin (green), which decorate the primary cilia; to pericentrin (magenta), which decorate the centrosome; and to ZO-1 (red), which label the tight junctions that encircle each cell. (b) The centrosome

contains two centrioles in the G1 or G0 stage of the cell cycle, with one of the centrioles functioning as a basal body for the primary cilium. During S phase, the centrosome duplicates in preparation for mitosis. As cells enter mitosis, the primary cilium is disassembled and the duplicated centrosomes form the spindle poles. (c) A diagram depicting a section through a nonmotile primary cilium, showing the lack of the central pair of microtubules and dynein arms typical of motile cilia and flagella. (d) Scanning electron micrographs of epithelial cells of a kidney collecting tubule from a wild-type mouse (*left*) and a mutant mouse defective in a component of the IFT particles. Arrows point to the primary cilia, which are short stubs in the mutant mouse.

[Part (a) Republished with permission from John Wiley & Sons, from B. D. Engel et al., 2011, "A Cell-Based Screen for Inhibitors of Flagella-Driven Motility in *Chlamydomonas* Reveals a Novel Modulator of Ciliary Length and Retrograde Actin Flow," *Cytoskeleton* **68**(3):188–203; permission conveyed through Copyright Clearance Center, Inc. Part (d) G. J. Pazour et al., 2000, *J. Cell Biol.* **151**(3):709–718; https://doi.org/10.1083/jcb.151.3.709.]

Terminally differentiated cells and dividing cells not in mitosis contain a primary cilium. In the latter case, the presence of the primary cilium is tied to the duplication cycle of the centrioles (discussed in Section 8.6), with the older, mother centriole functioning as the basal body for the assembly of the primary cilium (Figure 18-35b). Thus the centrosome, consisting of the two centrioles and pericentriolar material (see Figure 18-6), serves both as the basal body for the primary cilium and as the MTOC for the organization of the cytoplasmic array of microtubules.

The primary cilium is nonmotile because it lacks the central pair of microtubules and the dynein arms that are found in other cilia and in flagella (Figure 18-35c), but it does retain the IFT transport system. Work in the past several years has shown that the primary

cilium is a sensory organelle, acting as the cell's antenna by detecting extracellular signals. For example, our sense of smell is due to odorants binding to receptors located in the primary cilia of olfactory sensory neurons in the nose (see Chapter 23). In another example, the rod and cone cells of the eye have a primary cilium with a greatly expanded tip to accommodate the proteins involved in photoreception. The retinal protein opsin moves through the primary cilium at about 2000 molecules per minute, transported by kinesin-2 as part of the IFT system. Defects in this transport cause retinal degeneration. In vertebrates, the primary cilium is the site of action of the Hedgehog signaling pathway, as described in Chapter 16.

The primary cilium has a diffusion barrier at its base so that only the appropriate proteins can enter it: globular proteins of 10 kDa have ready access, but proteins above 40 kDa are excluded. Remarkably, transport through this barrier has similarities to transport into the nucleus through nuclear pores, and indeed, these two forms of transport have some components in common. Recall that import though nuclear pores requires a gradient of the Ran GTPase and involves binding of cargo proteins to importins (see Section 13.6). Transport of at least some proteins, such as kinesin-2, across the base and into a primary cilium requires both a gradient of the Ran GTPase and importins.

Defects in Primary Cilia Underlie Many Diseases

For many years, the existence and function of the primary cilium was ignored. However, this situation has changed dramatically over the last decade as we have come to appreciate that defects in intraflagellar transport result in the loss of primary cilia in mice (Figure 18-35d), and many diseases have now been traced to defects in primary cilia and IFT. One of the first clues came from the discovery that loss of a mammalian homolog of a *Chlamydomonas* IFT protein results in defects in the primary cilia and causes autosomal dominant polycystic kidney disease (ADPKD). It is hypothesized that the primary cilia on the epithelial cells of the kidney collecting tubule act as mechanochemical sensors to measure the rate of fluid flow by the degree to which they are bent.

In another example, patients with Bardet-Biedl syndrome have retinal degeneration, polydactyly (from the Greek for "many fingers"), and obesity. The syndrome can be caused by mutations in any one of 14 genes and has been traced to defects in the function of primary cilia. Many of these genes encode subunits of the BBsome, an octameric complex that forms a coat with structural elements in common with COPI, COPII, and clathrin (described in Chapter 14) and that traffics membrane proteins to cilia. While defects in many of the BBsome's components do not affect the structure of the primary cilium itself, they result in a lack of specific membrane

receptors that would normally be delivered to primary cilia through the interaction of the BBsome with the IFT apparatus. For example, the polydactyly seen in patients with Bardet-Biedl syndrome is due to a loss of localized Hedgehog signaling (see Chapter 16) in the primary cilium that is necessary for patterning during embryogenesis. ■

KEY CONCEPTS OF SECTION 18.5

Cilia and Flagella: Microtubule-Based Surface Structures

- Cilia and flagella are microtubule-based cell-surface structures with a characteristic central pair of singlet microtubules and nine sets of outer doublet microtubules (see Figure 18-31).
- All cilia and flagella grow from basal bodies, structures with nine sets of outer triplet microtubules that are closely related to centrioles.
- Axonemal dyneins attached to the A tubule on one doublet microtubule interact with the B tubule of another to bend cilia and flagella.
- Cilia and flagella have a mechanism, intraflagellar transport (IFT), in which material is transported to the tip by kinesin-2 and from the tip back to the base by cytoplasmic dynein. IFT regulates the function and length of cilia and flagella.
- Many cells have on their surface a single nonmotile primary cilium, which lacks the normal central pair of microtubules and the dynein arms of motile cilia. The primary cilium functions as a sensory organelle, with receptors for extracellular signals localized to its plasma membrane.
- Due to its sensory function, many diseases result from defects in receptor localization or in the structure of the primary cilium itself.

18.6 Mitosis

Of all the processes that permit the existence and perpetuation of life, perhaps the most critical is the ability of cells to accurately duplicate and then faithfully segregate their chromosomes at each cell division. During the **cell cycle**, a highly regulated process discussed in Chapter 19, cells precisely duplicate their chromosomes once during a period known as *S phase* (for DNA synthesis phase). After duplication, the two copies of each duplicated chromosome are held together by proteins called *cohesins*. The cells then pass through a period called G_2 (for gap 2) before entering **mitosis**, the process by which the duplicated chromosomes are segregated to the daughter cells.

Mitosis has to be precise — loss or gain of a chromosome can be lethal to the cell or cause severe complications for the cell. It is estimated that a yeast cell only mis-segregates 1 of its 16 chromosomes every 100,000 cell divisions, which makes mitosis one of the most accurate processes in biology. To achieve this level of accuracy, the process is highly regulated so that it proceeds in an orderly series of steps that are free of errors. In addition, error detection mechanisms proofread the steps. When errors are detected, the process pauses so they can be corrected. The timing and mechanisms that ensure the fidelity of mitosis are regulated by the cell cycle circuitry that we discuss in detail in Chapter 19. Here

we limit our discussion of this circuitry as it applies to microtubules and the mechanics of mitosis.

In outline, the cell has to first organize the assembly of a machine (the spindle) to segregate the chromosomes, then attach the duplicated chromosomes to the spindle so that during segregation one copy of each chromosome moves toward each daughter cell, and finally the cell divides in two so that each daughter cell inherits a complete set of chromosomes.

Centrosomes Duplicate Early in the Cell Cycle in Preparation for Mitosis

In preparing to separate the chromosomes during mitosis, cells duplicate their MTOCs — their centrosomes — in a manner that is coordinated with the duplication of their chromosomes in S phase (Figure 18-36). As we discuss in Chapter 19, the cell cycle is largely driven by the association of cell-cycle-specific cyclins with cyclin-dependent kinases (CDKs). Two kinases — G_1/S phase CDK and Plk4 — promote centrosome duplication. The duplicated centrosomes separate as the cell enters mitosis in a process known as *centrosome disjunction* and then move to opposite sides of the nucleus to become the two poles that will establish the mitotic spindle. The number of centrosomes in animal cells is carefully controlled to ensure the assembly of an appropriate spindle. In fact, many tumor

cells have more than two centrosomes, which contributes to genetic instability resulting from mis-segregation of chromosomes and hence **aneuploidy** (unequal numbers of chromosomes). The reasons why aneuploidy results in cancer are discussed in Chapter 25.

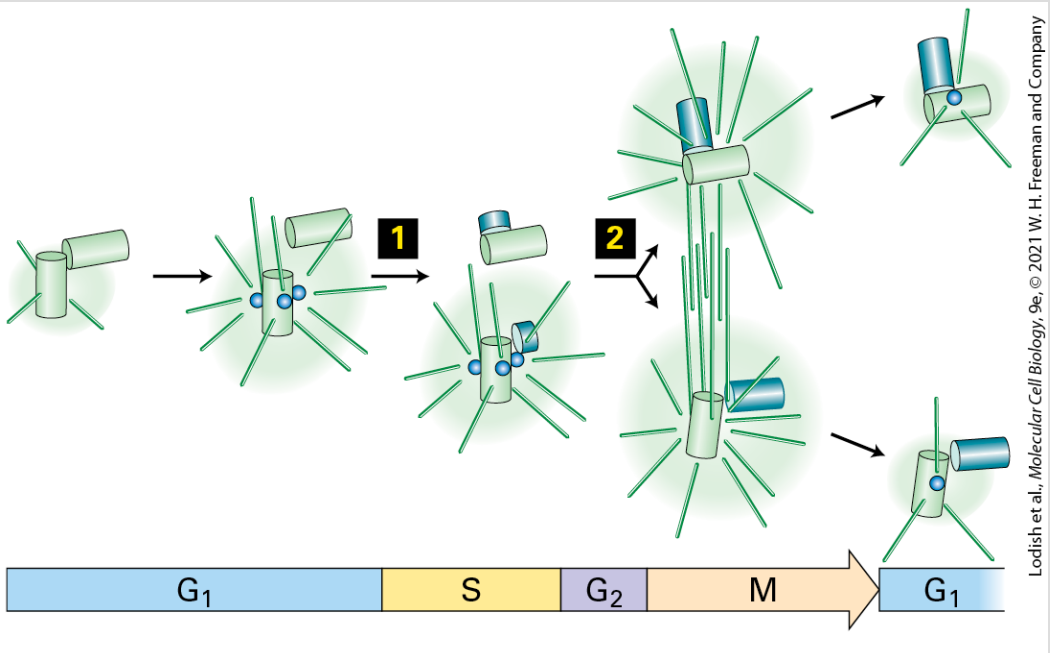
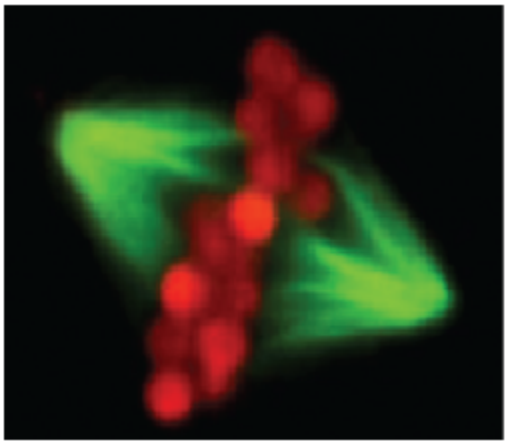


FIGURE 18-36 Relation of centrosome duplication to the cell cycle. Centrosome duplication, which is initiated by the G_1/S phase CDKs and Plk4 (step **1**), results in the pair of centrioles (green) separating and a daughter centriole (blue) budding from each. By the G_2 phase, growth of the daughter centrioles is complete, but the two pairs of centrioles remain within a single centrosomal complex. Early in mitosis, driven by the activation of M phase CDKs (step **2**), the centrosome splits, each half nucleates assembly of microtubules, and the two centriole pairs are moved to opposite sides of the nucleus. The amount of pericentriolar material (green area around the centrioles) and the microtubule nucleation activity of the centrosomes increase greatly in mitosis. In mitosis, each MTOC is called a spindle pole.

There are a number of cases in which spindles form in the absence of centrosomes, including plant-cell mitosis and animal-cell meiosis in females. This observation implies that nucleation of microtubules from centrosomes is not the only way in which a spindle can form. Studies exploiting mitotic extracts from frog eggs — extracts that do not contain centrosomes — show that the addition of beads covered with DNA is sufficient to assemble a relatively normal mitotic spindle (Figure 18-37). In this system, the beads induce the assembly of microtubules, and factors in the extract cooperate to make a spindle. One of the factors necessary for this reaction is cytoplasmic dynein, which is proposed to bind to two microtubules and migrate to their (−) ends, thereby drawing them together.



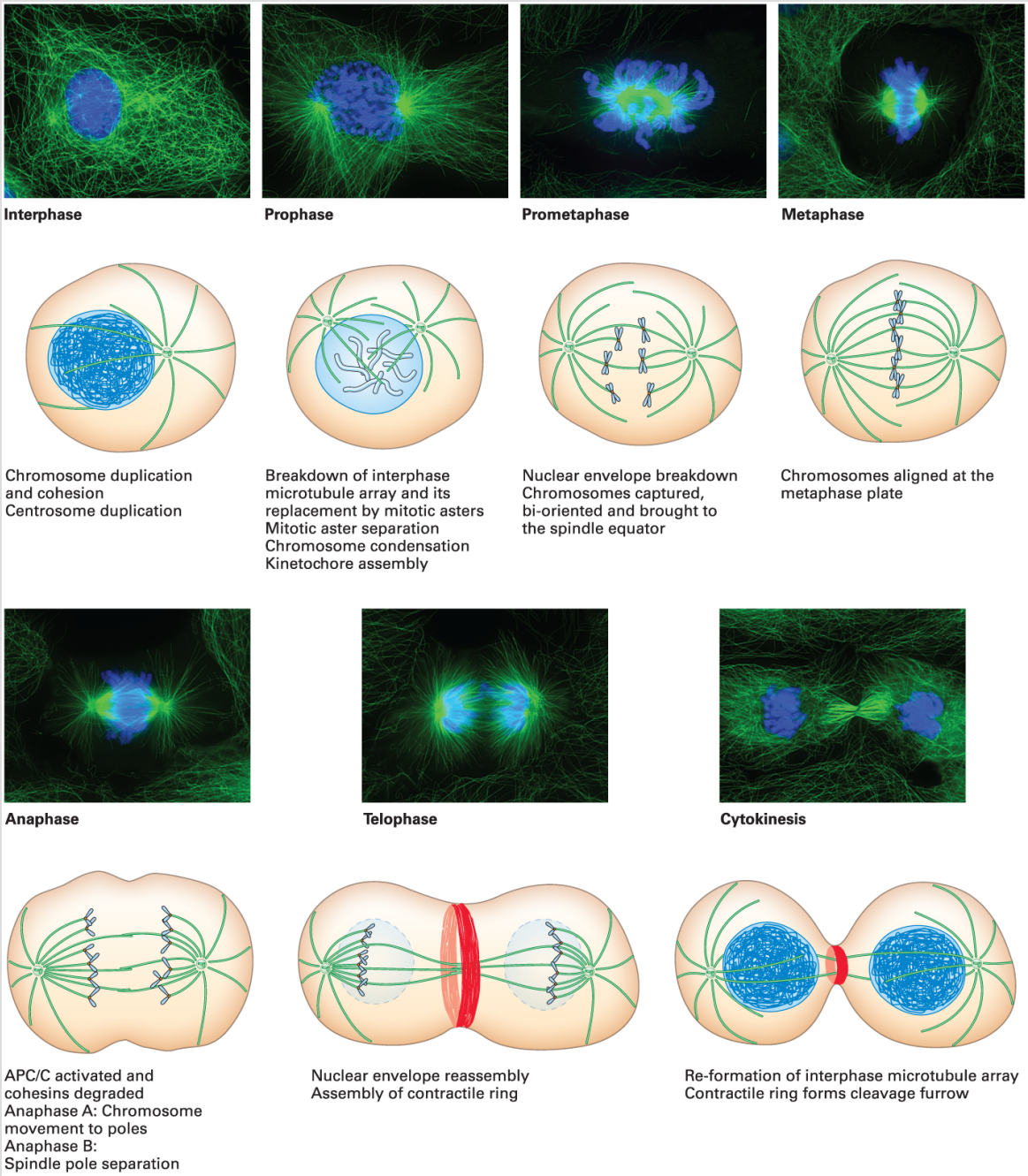
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-37 Mitotic spindles can form in the absence of centrosomes. Centrosome-free extracts can be isolated from frog oocytes arrested in mitosis. When fluorescently labeled tubulin (green) is added to the extracts together with beads covered with DNA (red), mitotic spindles spontaneously form around the beads from randomly nucleated microtubules. See K. Kinoshita et al., 2002, *Trends Cell Biol.* **12**:267–273; and C. Antonio et al., 2000, *Cell* **102**:425.

[Republished with permission from Nature Publishing Group, from R. Heald et al., 1996, "Self-Organization of Microtubules into Bipolar Spindles Around Artificial Chromosomes in *Xenopus* Egg Extracts," *Nature* **382**(6590):420–425; permission conveyed through Copyright Clearance Center, Inc.]

Mitosis Can Be Divided into Five Stages

Mitosis is a carefully orchestrated process that proceeds through several stages, driven by the cell cycle circuitry, with each stage dependent on completion of the previous one (Figure 18-38). We first review events that occur in each of the stages, and then discuss some of the mechanisms that bring them about.



Torsten Wittmann, University of California, San Francisco.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-38 The stages of mitosis, followed by cytokinesis, in animal cells. Upper panels show stages in cultured cells stained blue for DNA and green for tubulin. Lower diagrams show the different stages and the events that occur in each.

Prophase is the first stage of mitosis. It is preceded by a number of coordinated and dramatic events that are necessary to assemble the spindle and prepare the cell for chromosome segregation and cell division. Some of these processes directly relate to chromosome segregation, while others facilitate mitosis or prepare the cell for the formation of two daughter cells. We discuss these two categories of processes in turn.

To prepare the microtubules for mitosis, the interphase array of microtubules is replaced as the duplicated centrosomes gain more pericentriolar material and become more active in microtubule nucleation. Because the assembly of microtubules radiating from these two MTOCs now resembles stars, they are often called mitotic **asters**. The dynamics of the growing microtubules increases due to changes in the activities of + TIPs at their (+) ends. Bipolar kinesin-5 (see Figure 18-20) walks toward the (+) ends of interdigitating microtubules, which slides them apart until the two asters come to lie on opposite sides of the nucleus (Figure 18-39a). The separated centrosomes will become the poles of the **mitotic spindle**, the microtubule-based structure that separates chromosomes. The cohesins holding together each pair of duplicated chromosomes, or **sister chromatids** as they are called at this stage, degrade except at the centromeric region, where the two sister chromatids remain linked by intact cohesins (Figure 18-39b). Also during prophase, specialized structures called **kinetochores**, which will become sites

of microtubule attachment, assemble at the centromeric region of each sister chromatid.

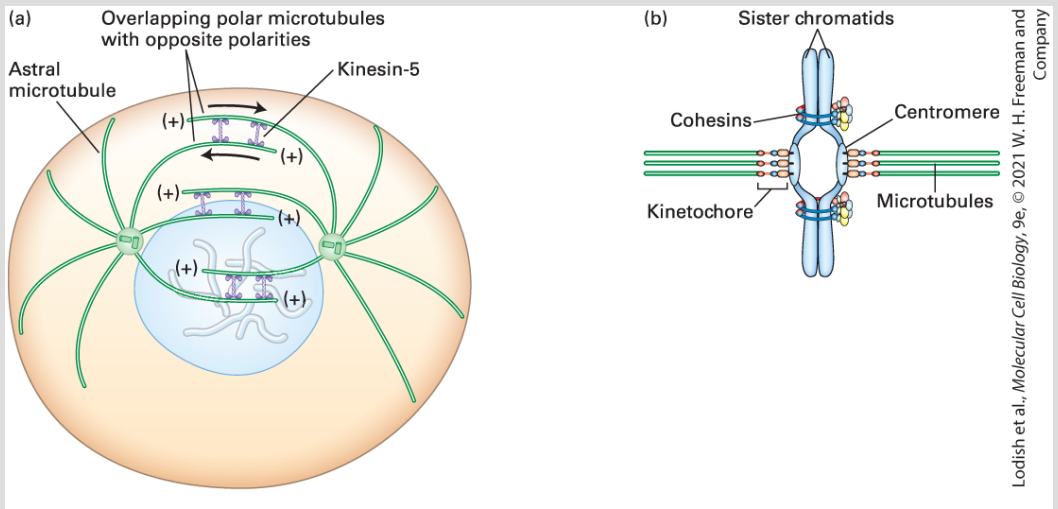


FIGURE 18-39 Separation of mitotic asters in prophase and parts of a condensed chromosome. (a) In prophase, the asters are separated by the bipolar kinesin-5 interacting with overlapping microtubules of opposite polarities. (b) Parts of a condensed chromosome in mitosis. The duplicated chromosome consists of two sister chromatids (each is a single DNA duplex), held together by cohesins at a constricted region called the *centromere*. The centromere is the site where the kinetochore, which will make attachments to the kinetochore microtubules, forms.

As noted, many cellular processes that are not directly related to chromosome segregation are dramatically altered in preparation for cell division. The internal membrane organization (which is normally dependent on the interphase array of microtubules) disassembles in preparation for the equal segregation of organelles into the two daughter cells. Endocytosis and exocytosis cease, and microfilaments rearrange to give rise to a rounded cell. In the

nucleus, the nucleolus breaks down and chromosomes begin to condense.

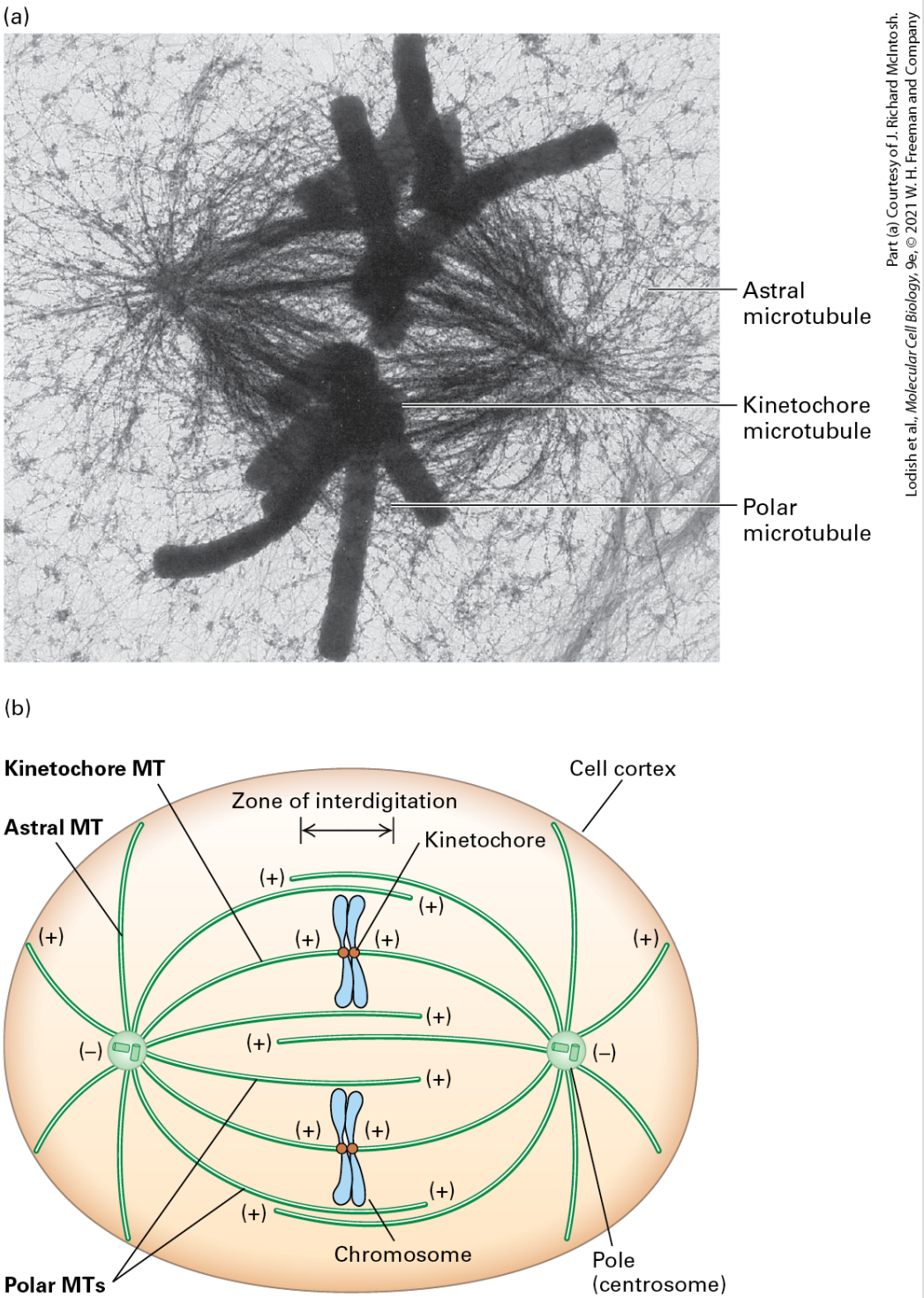
The next stage of mitosis, **prometaphase**, is signaled by disassembly of the nuclear pores and the nuclear lamina that underlies the nuclear envelope. This leads to breakdown of the nuclear envelope and its retraction into the endoplasmic reticulum. Loss of the nuclear envelope enables microtubules assembled from the spindle poles to search for and capture chromosome pairs by associating with their kinetochores. Each chromatid has a kinetochore, so a sister chromatid pair has two kinetochores (see Figure 18-39b). Microtubules from each spindle pole attach to one of the kinetochores. This attachment process is a critical step, which we discuss in detail shortly. The next stage, **metaphase**, is achieved when all the chromosome pairs align at a point equidistant between the two spindle poles.

When the cell detects that all chromosomes are attached to the spindle correctly, the next stage, **anaphase**, is induced by activation of the anaphase-promoting complex or cyclosome (APC/C) (discussed in Chapter 19). The activated APC/C ultimately leads to the destruction of the cohesins that were holding the sister chromatids together, so that each chromatid can be pulled to its respective pole by the microtubules attached to its kinetochore. This movement is known as *anaphase A*. A separate and distinct movement also occurs. The poles move farther apart, which

elongates the spindle in a process known as *anaphase B*. Now that the chromosomes have separated, the cell enters **telophase**, during which the nuclear envelope and nuclear lamina re-form, the chromosomes decondense, and the spindle disperses. Finally, the cell is pinched into two daughter cells by the contractile ring during **cytokinesis**.

The Mitotic Spindle Contains Three Classes of Microtubules

Before we discuss the mechanisms involved in the remarkable process of mitosis, it is important to examine the three classes of microtubules that emanate from the spindle poles. The *astral microtubules* extend from the spindle poles to the cell cortex (Figure 18-40). By interacting with the cortex, the astral microtubules perform the critical function of orienting the spindle to the axis of cell division. The *kinetochore microtubules* link the spindle poles to the kinetochores on sister chromatid pairs. During anaphase A, the kinetochore microtubules transport the newly separated chromosomes to their respective poles. The *polar microtubules* extend from each spindle pole toward the opposite pole and opposite-oriented polar microtubules interact in an antiparallel manner. As we mentioned, it is kinesin-5 walking along the polar microtubules that pushes the duplicated centrosomes apart during prophase.



Part (a) Courtesy of J. Richard McIntosh.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-40 Mitotic spindles have three distinct classes of microtubules. (a) In this high-voltage electron micrograph, a cell in metaphase was made permeable and microtubules were stained with biotin-tagged anti-tubulin antibodies to increase their visibility. The large cylindrical objects are

chromosomes. (b) Schematic diagram corresponding to the metaphase cell in (a). Three sets of microtubules (MTs) make up the mitotic apparatus. All the microtubules have their (−) ends at the poles. Astral microtubules project toward the cortex and are linked to it. Kinetochore microtubules are connected to chromosomes. Polar microtubules project toward the cell center with their distal (+) ends overlapping. The spindle pole with its associated microtubules is also known as a *mitotic aster*.

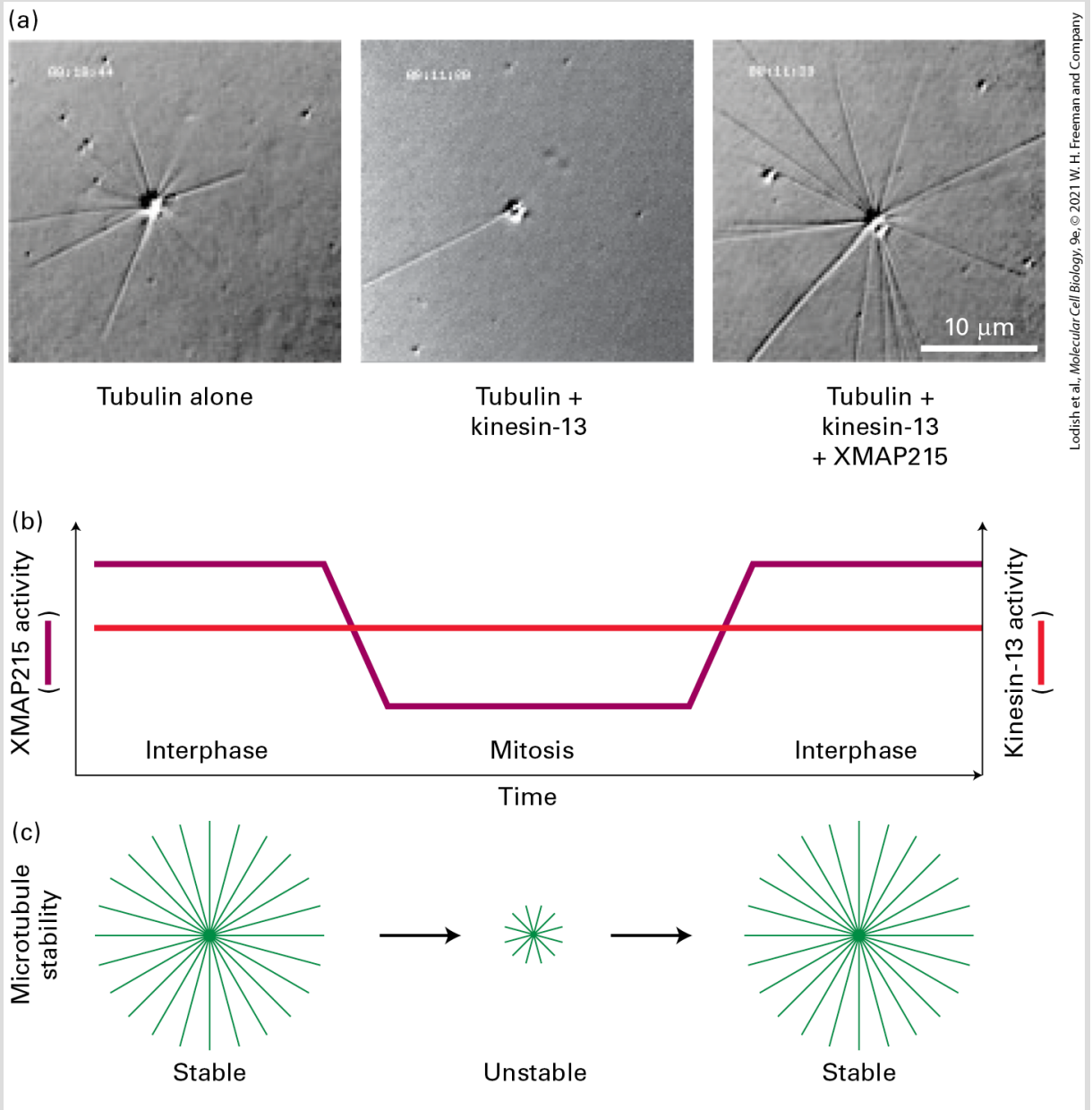
Microtubule Dynamics Increase Dramatically in Mitosis

Although we have drawn static images of the stages of mitosis, throughout mitosis microtubules are highly dynamic. As we have seen, as cells enter mitosis the ability of their centrosomes to nucleate the assembly of microtubules increases significantly (see Figure 18-36). In addition, microtubules become much more dynamic, allowing them to capture chromosomes and contribute to assembly of the spindle. How was this determined? In principle, one could label microtubules with a fluorescent tag and watch their individual behaviors, but practically speaking, there are too many microtubules in a mitotic spindle to follow individual ones. To get an average value for the dynamic instability of these microtubules, researchers introduced fluorescently labeled tubulin into cells undergoing mitosis, which became incorporated randomly into the cells' microtubules. They then bleached the fluorescent label in a small region of the mitotic spindle and measured the rate at which fluorescence came back using the technique known as *fluorescence*

recovery after photobleaching (FRAP) (see Figure 4-22). Since fluorescence recovery is due to the assembly of new microtubules from soluble fluorescent tubulin, the rate of recovery represents the average rate at which microtubules turn over. In a mitotic spindle, microtubule half-life is about 15 seconds, whereas in an interphase cell, it is about 5 minutes. It should be noted that these are bulk measurements and that individual microtubules can be more stable or dynamic, as we will see.

What makes microtubules more dynamic as the cell enters mitosis? As we discussed in Section 18.2, dynamic instability is a measure of the relative contributions of growth rates, shrinkage rates, catastrophes, and rescues (see Figure 18-9). Analysis of microtubule dynamics in vivo shows that the enhanced instability of individual microtubules in mitosis is generated mainly by an increase in catastrophes and a decrease in rescues, with little change in rates of growth (i.e., lengthening) or shrinkage (i.e., shortening). Studies with extracts from frog oocytes have suggested that the main factor enhancing catastrophes in both interphase and mitotic extracts is depolymerization by kinesin-13 proteins. This can be seen in an in vitro assay in which microtubule assembly from pure tubulin is nucleated from purified centrosomes (Figure 18-41a). If kinesin-13 is added to the assay, many fewer microtubules are formed. However, if the protein XMAP215, which enhances assembly at the (+) end, is added with the kinesin-13, many microtubules are formed due to a dramatic reduction in catastrophe frequency. It turns out that the

activity of kinesin-13 does not change significantly during the cell cycle, whereas the activity of XMAP215 is inhibited by phosphorylation during mitosis (Figure 18-41b). This results in much more unstable (more dynamic) microtubules as the cell enters mitosis (Figure 18-41c). This increase in dynamics is necessary to enhance the probability of chromosome capture by microtubules as described in the next section.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-41 Microtubule dynamics increase in mitosis due to loss of a stabilizing MAP. (a) These three panels reveal the ability of centrosomes to assemble microtubules under various conditions: with pure tubulin (*left*); with tubulin and the destabilizing protein kinesin-13 (*middle*); and with tubulin, kinesin-13, and the stabilizing protein XMAP215 (*Xenopus* MAP of 215 kDa) (*right*). Further analysis shows that the major effect of XMAP215 is to suppress catastrophes induced by kinesin-13. (b) The increased dynamics of microtubules in mitosis is due to the inactivation of XMAP215 by phosphorylation. (c) Diagram comparing the stabilities of microtubules in interphase and in

mitosis. Note that in addition to the decrease in *stability* in mitosis, the ability of MTOCs to *nucleate* microtubules increases dramatically in mitosis.

[Part (a) Republished with permission from AAAS, from K. Kinoshita et al., 2001, "Reconstitution of Physiological Microtubule Dynamics Using Purified Components," *Science* **294**(5545):1340–1343; permission conveyed through Copyright Clearance Center, Inc. Part (b) Data from K. Kinoshita, B. Habermann, and A. A. Hyman, 2002, *Trends Cell Biol.* **12**(6):267–273.]

Chromosomes Are Captured and Oriented During Prometaphase

Kinetochores, the structures that mediate attachment between chromosomes and microtubules, assemble on each sister chromatid during prophase at a region called the centromere. The **centromere** is a constricted region of the condensed chromosome defined by a centromeric DNA sequence. Centromeric DNA can vary enormously in size; in budding yeast it is about 125 bp, whereas in humans it is on the order of 1 Mb (see Chapter 7). Kinetochores contain many protein complexes that facilitate the linkage between centromeric DNA and microtubules. In animal cells, the kinetochore consists of a centromeric DNA layer and inner and outer kinetochore layers, with the (+) ends of kinetochore microtubules terminating in the outer layer (Figure 18-42). Yeast kinetochores are attached by a single microtubule to their spindle pole, human kinetochores are attached by about 30 microtubules, and plant chromosomes are attached by hundreds.

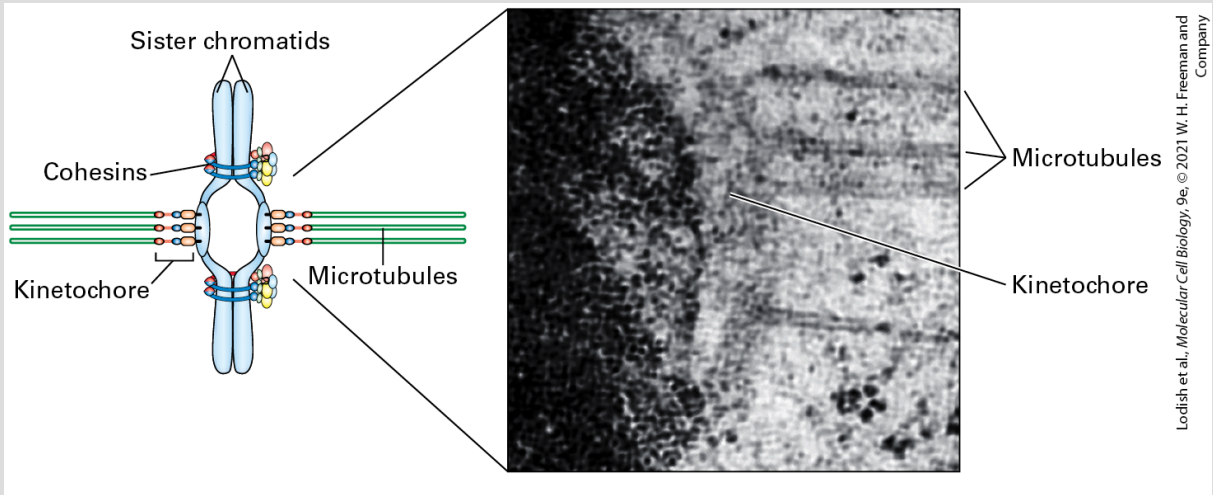


FIGURE 18-42 The structure of a mammalian kinetochore. Diagram and electron micrograph of a mammalian kinetochore. The arrows indicate the ends of microtubules that insert into the kinetochore.

[Republished with permission from Springer, from B. F. McEwen et al., 1998, "A New Look at Kinetochore Structure in Vertebrate Somatic Cells Using High-Pressure Freezing and Freeze Substitution," *Chromosoma* **107**(6–7): 366–375; permission conveyed through Copyright Clearance Center, Inc.]

How does a kinetochore become attached to microtubules in prometaphase? Microtubules nucleated from the spindle poles are very dynamic, and when they contact a kinetochore, either laterally or at their ends, this contact can lead to chromosomal attachment (Figure 18-43a, steps **1a** and **1b**). Microtubules that capture kinetochores are selectively stabilized by a reduction in the rate of catastrophes, which increases the chance that the attachment will persist.

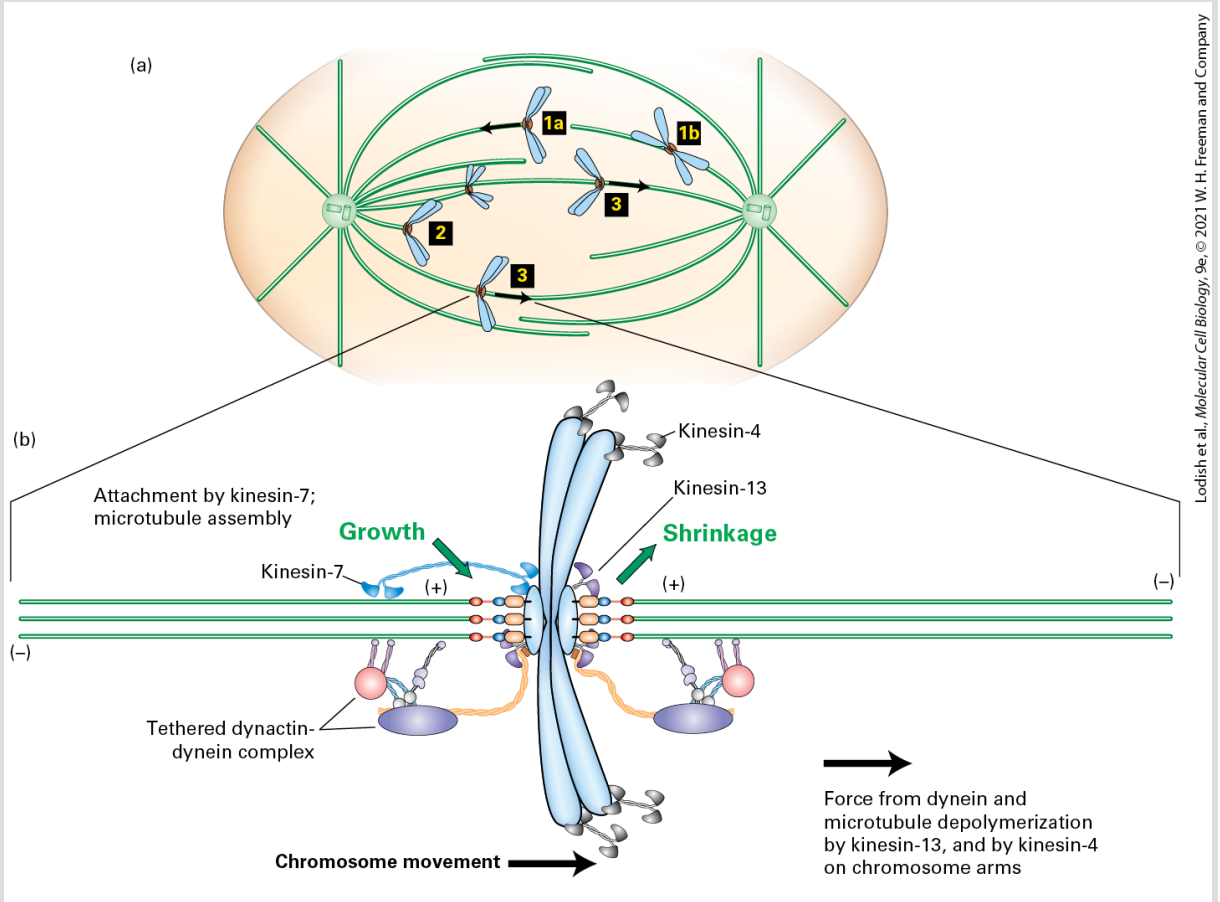


FIGURE 18-43 Chromosome capture and congression in prometaphase. (a) In the first stage of prometaphase, chromosomes become attached either to the end of a microtubule (**1a**) or to the side of a microtubule (**1b**). The chromosome is then drawn toward the spindle pole by dynein-dynactin that becomes associated with one of the kinetochores of the chromosome and moves toward the (–) end of the microtubule (**2**). Eventually, a microtubule from the opposite pole finds and becomes attached to the free kinetochore, and the chromosome is now said to be bi-oriented (**3**). The bi-oriented chromosomes then move to a central point between the spindle poles in a process known as *chromosome congression*. For ease of presentation, only single kinetochore microtubules are shown here. (b) Congression involves bidirectional oscillations of chromosomes, with one set of kinetochore microtubules shortening on one side of the chromosomes and the other set lengthening on the other. The kinetochore also contains many additional protein complexes not shown here.

Other studies have uncovered a mechanism involving Ran, a small GTPase, that enhances the chance that microtubules will encounter kinetochores. Recall that proteins destined for import into the nucleus associate with a nuclear import receptor, called an *importin* (see Figure 13-36). The complexes move from the cytosol through nuclear pores and into the nucleoplasm. The level of Ran·GTP is high in the nucleus, because its activator, the guanine nucleotide exchange factor for Ran (Ran-GEF), is localized there by binding to chromatin. In the nucleus, Ran·GTP binds to the importin, which changes the conformation of importin, causing it to release the transported protein. Then the importin-Ran·GTP complex leaves the nucleus through a nuclear pore to return to the cytosol.

In mitosis, Ran-GEF continues to associate with chromatin, so even after nuclear envelope breakdown there is a gradient of Ran·GTP around the chromosomes. This Ran·GTP releases a protein called *TPX* from importin. TPX binds the augmin complex and γ-TuRC and associates with the side of an existing microtubule to nucleate new microtubules at a shallow angle to the mother microtubule (see Figure 18-7c). This enhances the density of microtubules in the vicinity of chromosomes.

Once a kinetochore is attached laterally or terminally to a microtubule, dynein-dynactin associates with the kinetochore to move the duplicated chromosome down the microtubule toward the spindle pole. This movement eventually results in an end-on

attachment of the microtubule to one kinetochore (Figure 18-43a, step **2**). This movement helps orient the sister chromatid so that the unoccupied kinetochore on the opposite side is pointing toward the distal spindle pole. Eventually a microtubule from the distal pole will capture the free kinetochore; at this point the sister chromatid pair is said to be *bi-oriented* (Figure 18-43a, step **3**). With the two kinetochores attached to opposite poles, the duplicated chromosome is now under tension, being pulled in opposite directions by the two sets of kinetochore microtubules.

Duplicated Chromosomes Are Aligned by Motors and Microtubule Dynamics

During prometaphase, the chromosomes come to lie at the midpoint between the two spindle poles, called the *metaphase plate*, in a process known as chromosome *congression*. During this process, bi-oriented chromosome pairs often oscillate backward and forward before arriving at the metaphase plate. Chromosome congression involves the coordinated activity of several microtubule-based motors together with regulators of microtubule assembly and disassembly (Figure 18-43b). These regulators are localized at the kinetochores, but how they are maintained there is poorly understood — they are not part of the stable kinetochore complexes. The oscillating behavior of chromosomes involves lengthening of microtubules attached to one kinetochore and shortening of

microtubules attached to the other kinetochore, all without losing their attachments. In metazoans, several microtubule-based motors associated with the kinetochore contribute to this process. First, dynein-dynactin provides the strongest force, pulling the chromosome pair toward the more *distant* pole. This movement requires simultaneous shortening of the microtubule, which is enhanced by kinetochore-localized kinesin-13. The microtubules associated with the opposite kinetochore have to grow as the chromosome moves. Anchored at this kinetochore is the kinesin-related motor kinesin-7 (also known as *CENP-E*), which holds onto the growing (+) end of the lengthening microtubule. Also contributing to congression is another kinesin, chromokinesin/kinesin-4, which associates with the chromosome arms. Kinesin-4, a (+) end–directed motor, interacts with the polar and kinetochore microtubules to pull the chromosomes toward the center of the spindle and as a result, the chromosome arms point away from the closest spindle pole. When the chromosomes have congressed to the metaphase plate, dynein-dynactin is released from the kinetochores and streams down the kinetochore microtubules to the poles. These different activities and opposing forces work together to bring all the chromosomes to the metaphase plate, at which point the cell is ready for anaphase.

The Chromosomal Passenger Complex Regulates Microtubule Attachment at

Kinetochores

To have healthy daughter cells, the segregation of chromosomes at mitosis must be very accurate, so it is crucial that all chromosomes are bi-oriented before anaphase begins. During the random kinetochore-to-microtubule attachment process, it is possible for mistakes to be made. For example, both kinetochores of a sister chromatid pair might attach to microtubules from the same spindle pole. If such attachments persisted into anaphase, it would result in one cell missing a chromosome and the other having an extra one, which would either be lethal or very detrimental. Cells have two mechanisms to ensure that all chromosomes are correctly bi-oriented before anaphase begins. The first ensures that *each* individual chromosome is appropriately bi-oriented, and the second ensures that *all* of the chromosomes are properly bi-oriented.

The first mechanism works to ensure that the kinetochore-microtubule interactions are weak until bi-orientation occurs. When a chromosome is correctly bi-oriented, tension is produced across the chromosome, which leads to the kinetochore-microtubule attachments becoming stabilized. To understand how this works, we need to look a bit more closely at the molecular components that link a kinetochore to a microtubule. As we discussed in Chapter 7, kinetochores assemble on regions of chromosomal DNA marked by a centromere-specific H3 histone variant called *CENP-A*. This variant marks the site for kinetochore assembly. About half a dozen

distinct stable protein complexes, consisting of more than 40 different proteins, have been shown to associate with this centromeric region in yeast. Essentially all these protein complexes are conserved in humans, which is not surprising given the fundamental importance of kinetochores. One of these, the Ndc80 complex, is long and flexible, and many copies of it link the inner kinetochore with the (+) end of the microtubule in a sleevelike arrangement (Figure 18-44a). The function of Ndc80 and many of the associated factors at the kinetochore are regulated by the *chromosomal passenger complex (CPC)*. This complex associates with the inner kinetochore early in mitosis, and among its components is a protein kinase called *Aurora B*. Once the CPC associates with the kinetochore, Aurora B can phosphorylate several components in its vicinity, including the Ndc80 complex. This loosens the attachment of Ndc80 to the microtubule. The phosphorylation of these components is not stable, as the phosphatase PP1, which is associated with the outer kinetochore, can dephosphorylate them. Thus when the kinetochores on a pair of sister chromatids are not under tension, Ndc80 is continually phosphorylated by Aurora B and dephosphorylated by PP1. The result is a weakening of the interaction between the kinetochore and microtubule. However, when correct bi-orientation occurs, the microtubules pull in opposite directions on both kinetochores, creating tension across the chromosome. This extends the flexible Ndc80 complex, which increases the spacing between the inner and outer kinetochores (Figure 18-44b, c). As a result of these movements, Ndc80 cannot be

phosphorylated by Aurora B, and the dephosphorylated state of Ndc80 renders it more firmly attached to the microtubule. In this way, microtubule attachments to bi-oriented chromosomes are selectively stabilized.

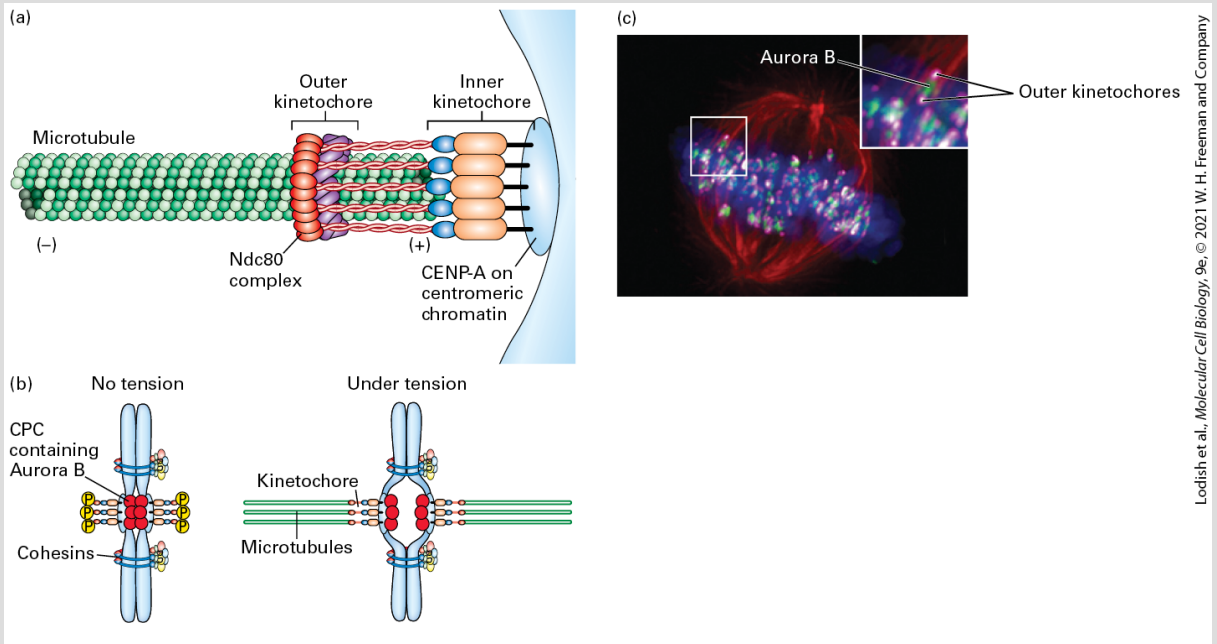


FIGURE 18-44 CPC regulation of microtubule-kinetochore attachment. The Ndc80 complex forms a critical and regulated attachment between the kinetochore and microtubule (+) end. (a) Diagram showing the sleevelike arrangement of the Ndc80 complex linking the inner kinetochore to the (+) end of a microtubule embedded in the outer kinetochore. See S. Santaguida and A. Musacchio, 2009, *EMBO J.* **28**:2511–2531. (b) Diagram of the relationship between the chromosomal passenger complex (CPC), which is associated with the inner kinetochore and contains the kinase Aurora B, and the outer kinetochore, where the phosphatase PP1 binds. Notice that when both kinetochores are under tension, the outer kinetochores move away from the CPC; as a result, Aurora B cannot phosphorylate components in the outer kinetochore, which include the microtubule-binding site of the Ndc80 complex. (c) Cell in metaphase stained for tubulin (red), DNA (blue), Aurora B kinase (green), and the outer kinetochore (magenta). Notice how the outer kinetochore is pulled away from Aurora B (inset).

[Part (c) Reprinted by permission from Nature Publishing Group, from S. Ruchaud, M. Carmena, and W. C. Earnshaw, 2007, “Chromosomal Passengers: Conducting Cell Division,”

Nat. Rev. Mol. Cell Biol. 2007, **8**:798–812; permission conveyed through Copyright Clearance Center, Inc.]

The CPC is important for bi-orientation of each individual chromosome, but it does not ensure that all chromosomes are bi-oriented before anaphase begins. The second mechanism to ensure correct chromosome segregation is the *spindle assembly checkpoint pathway*, a signaling pathway that stops the progression of the cell cycle into anaphase until tension is present at all the kinetochores. Even a single unattached, or inappropriately attached, kinetochore can activate the spindle assembly checkpoint pathway and pause the cell cycle until the error is corrected. An unattached kinetochore drives the assembly of the unstable mitotic checkpoint complex (MCC) as described in detail in Chapter 19. The MCC is an inhibitor of the APC/C and thereby delays entry into anaphase. When all kinetochores are appropriately attached, the MCC disassembles, the APC/C is activated, and the cell proceeds into anaphase.

Anaphase A Moves Chromosomes to Poles by Microtubule Shortening

The onset of anaphase A is one of the most dramatic movements that can be observed in the light microscope. When the spindle assembly checkpoint has been passed, APC/C activation induces proteolysis of the remaining cohesins holding the sister chromatids

together. Suddenly, the two paired sister chromatids separate from each other and are drawn to their respective poles. The movement is sudden because the kinetochore microtubules are under tension, and as soon as the cohesin attachments between the chromatids are removed, the separated chromatids are free to move.

Experiments with isolated metaphase chromosomes have shown that anaphase A movement can be powered by microtubule shortening, using the stored structural strain released after removal of the GTP-bound tubulin subunits at the microtubule tip. This mechanism can be nicely demonstrated in vitro. When metaphase chromosomes are added to purified microtubules, they bind preferentially to the (+) ends of the microtubules. Dilution of the mixture to reduce the concentration of free tubulin dimers results in the movement of the chromosomes toward the (−) ends by microtubule depolymerization at the chromosome-bound (+) ends. In addition, experiments have shown that in *Drosophila*, two members of the microtubule-depolymerizing kinesin-13 protein family (see Figure 18-15) also contribute to chromosome movement in anaphase A. One of these kinesin-13 proteins is localized at the kinetochore and enhances disassembly there (Figure 18-45, **A1**), and the other is localized at the spindle pole, enhancing depolymerization there (Figure 18-45, **A2**). Thus, at least in the fly where it has been studied, anaphase A is powered in part by kinesin-13 proteins specifically localized at the kinetochore and spindle pole to shorten the kinetochore microtubules at both their (+) and (−)

ends, drawing the chromosomes to the poles. In addition, it is hypothesized that the microtubule-severing protein, katanin, may sever microtubules from the γ-TuRC complex at the spindle pole (see Figure 18-7). This makes new (−) ends available for depolymerization, which contributes to microtubule shortening.

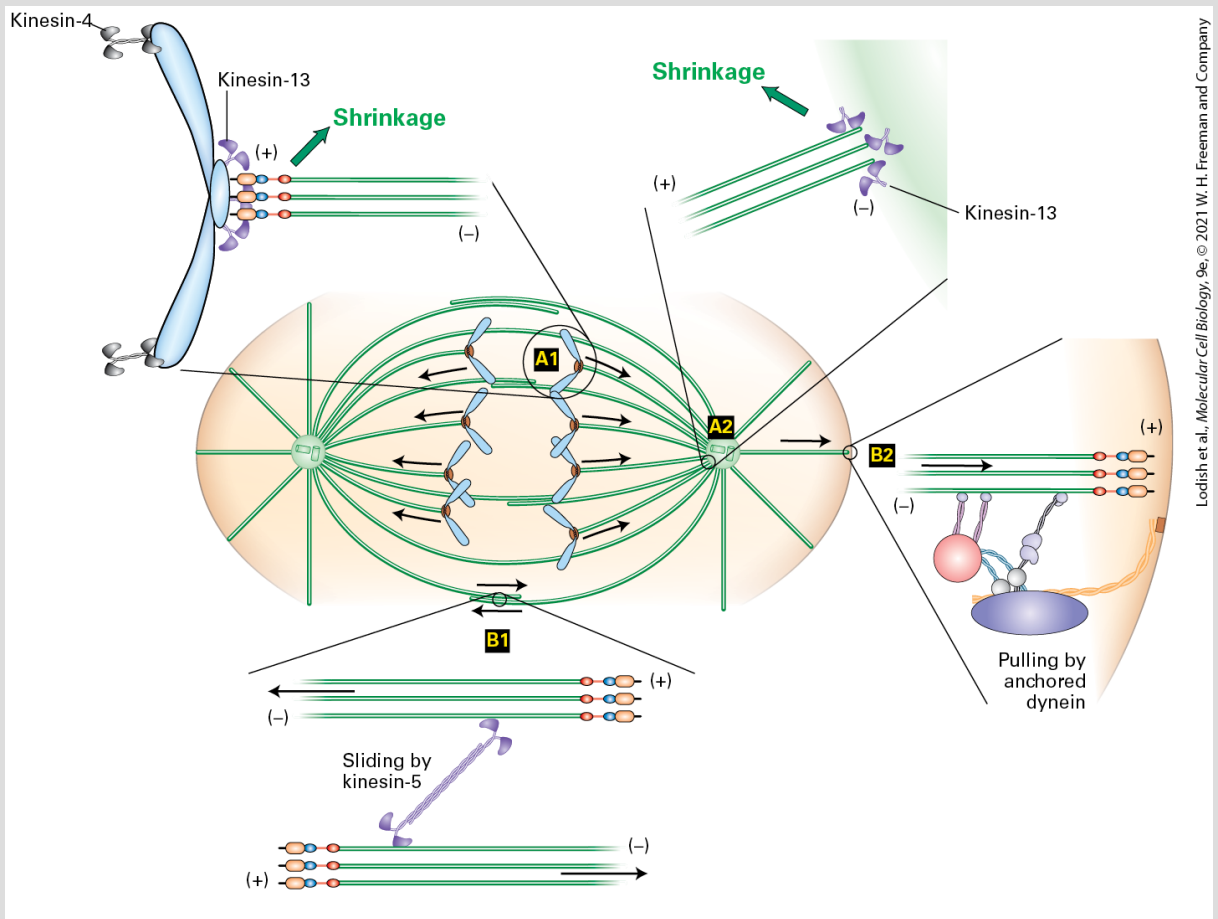


FIGURE 18-45 Chromosome movement and spindle pole separation in anaphase. Anaphase A movement is powered by microtubule-shortening kinesin-13 proteins at the kinetochore (**A1**) and at the spindle pole (**A2**). Note that the chromosome arms still point away from the spindle poles due to associated chromokinesin/kinesin-4 members, so the depolymerization force has to be able to overcome the force pulling the arms toward the center of the spindle. Anaphase B also has two components: sliding of antiparallel polar microtubules powered by a kinesin-5 (+) end–directed motor (**B1**), and pulling on astral

microtubules by dynein-dynactin located at the cell cortex (**B2**). Arrows indicate the direction of movement generated by the respective forces.

Anaphase B Separates Poles by the Combined Action of Kinesins and Dynein

Anaphase B causes the spindle to elongate by moving the spindle poles farther apart. A major contributor to this movement are the bipolar kinesin-5 proteins (Figure 18-45, **B1**); kinesin-5 is the same kinesin that is responsible for aster separation in prophase. These motors associate with the overlapping polar microtubules, and since they are (+) end–directed motors, they push the poles apart. While this is happening, the polar microtubules have to grow to accommodate the increased distance between the spindle poles. Another motor — the microtubule (−) end–directed cytoplasmic dynein/dynactin, localized and anchored on the cell cortex — pulls on the astral microtubules, which helps separate the spindle poles (Figure 18-45, **B2**).

The Spindle Is Centered and Oriented by a Dynein/Dynactin-Dependent Pathway

As we discuss shortly, the center of the spindle defines the plane where cell division will occur. Thus it is critical during prometaphase and metaphase that the spindle is placed centrally in the cell to ensure that the two daughter cells will be of equal size. This orientation and placement of the spindle is mediated by dynein/dynactin using two distinct mechanisms (Figure 18-46). Dynein is recruited to the cell cortex by a cortical anchor, allowing it to pull on astral microtubules to move the spindle. The regulation of the placement of the cortical anchor, and hence active dynein/dynactin, is what orients the spindle. First, the chromosomes send out a signal that inhibits the association of the cortical anchor in the vicinity of the chromosomes. This inhibitory signal is the Ran·GTP gradient that we described earlier in this section. The consequence of the Ran·GTP inhibitory signal is that cortical association of dynein/dynactin is restricted to the region of the cell around the spindle poles (Figure 18-46, signal **1**). Second, associated with each spindle pole is a kinase, called *polo-like* kinase, that can phosphorylate substrates in its vicinity, including the dynein/dynactin cortical anchor. Phosphorylation of the cortical anchor renders it unable to recruit dynein/dynactin to the cortex (Figure 18-46, signal **2**). Thus, as a spindle pole gets closer to the cortex, dynein/dynactin dissociates and the spindle moves back toward the center being pulled by dynein/dynactin on the more distal side. This tug of war ultimately moves the spindle into a central position in the cell. This function of dynein/dynactin pulling

on astral microtubules is the same mechanism that contributes to elongation of the spindle during anaphase B (see Figure 18-45).

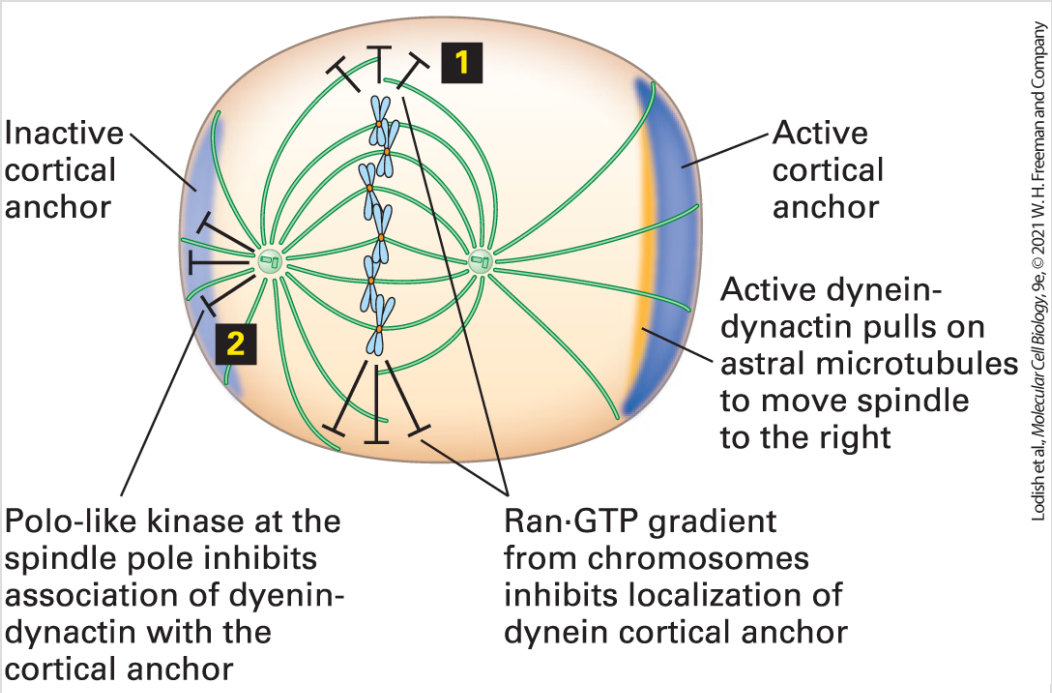


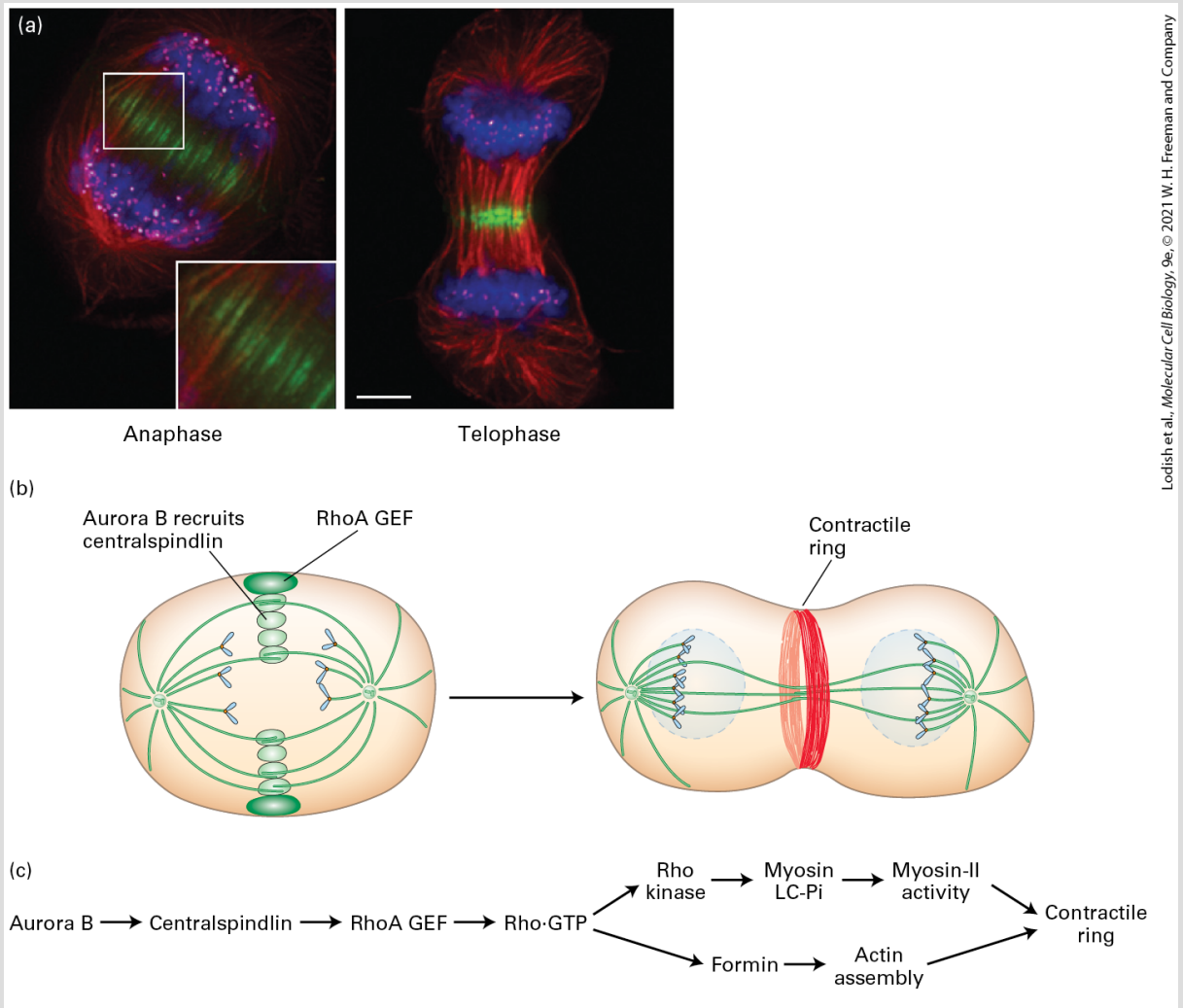
FIGURE 18-46 Dynein-dynactin centers and orients the spindle (a) During prometaphase and metaphase, the spindle is oriented and centered in the cell by two signals that regulate the location of the active cortical anchor for dynein. First, the chromosomes generate a Ran·GTP gradient (signal **1**) that inhibits association of the cortical anchor with the local cell cortex. Second, polo-like kinase at the spindle poles phosphorylates substrates in the local vicinity that inhibits the association of dynein-dynactin with the cortical anchor (signal **2**). See T. Kiyomitsu and I. M. Cheeseman, 2012, *Nat. Cell Biol.* **14**:311–317.

Cytokinesis Splits the Duplicated Cell in Two

During late anaphase and telophase in animal cells, the cell assembles a microfilament-based *contractile ring* attached to the plasma membrane that will eventually contract and pinch the cell into two, a process known as *cytokinesis* (see Figure 18-38). The contractile ring is a thin band of actin filaments of mixed polarity interspersed with myosin-II bipolar filaments (see Figure 17-35). On receiving a signal, the ring contracts, first to generate a *cleavage furrow* and then to pinch the cell into two.

Two aspects of the contractile ring are essential to its function. First, it has to be appropriately placed in the cell. It is known that this placement is determined by a signal provided by the spindle, so that the ring forms equidistant between the two spindle poles. The signal is provided, at least in part, by the chromosomal passenger complex (CPC) that regulates the attachment of microtubules to kinetochores during prometaphase (see Figure 18-44b). Up to anaphase, the CPC is associated with the inner kinetochores of unseparated chromatids. When anaphase begins, the CPC leaves the kinetochores and associates with the overlapping polar microtubules at the center of the spindle (Figure 18-47a). There the CPC recruits another protein complex, *centralspindlin*, that includes a (+) end–directed kinesin motor protein, which concentrates at the middle of the spindle due to its motor activity (Figure 18-47b). As anaphase B continues, centralspindlin recruits a guanine nucleotide exchange factor for RhoA to the adjacent region of the plasma membrane. Recall from Chapter 17 that Rho proteins are small GTP-

binding proteins that are activated by exchange factors to catalyze the exchange of GDP for GTP (see Figure 17-40). Once activated, the RhoA·GTP activates both a formin protein to drive the nucleation and assembly of actin filaments and myosin-II that together make up the contractile ring (Figure 18-47b, c). In this way, the position of the spindle directly defines the site of contractile ring formation, and hence cytokinesis.



Lodish et al., *Molecular Cell Biology, 9e*, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-47 The chromosomal passenger complex (CPC) remains at the spindle midzone during anaphase and telophase to direct the assembly of the contractile ring. (a) Micrographs of a cell in late anaphase (*left*) and telophase (*right*), showing

microtubules (red), DNA (blue), Aurora B kinase (green), and kinetochores (magenta). Notice how Aurora B, which is part of the CPC, concentrates in the region where the polar microtubules overlap and where the contractile ring will form. Scale bar $5\ \mu m$. (b) At anaphase, the CPC re-localizes to the region of overlapping microtubules and recruits the centralspindlin complex. This complex contains a kinesin motor that helps to concentrate it near microtubule $(+)$ ends. It recruits and activates a Rho guanine nucleotide exchange factor (RhoA GEF) at the adjacent plasma membrane, which activates RhoA to the GTP-bound form. (c) Pathway showing the interactions that results in the formation of Rho·GTP. This activates a formin to drive assembly of actin filaments, and also activates myosin II (discussed in Chapter 17). Together, the actin and myosin form the contractile ring that pinches the cell in two.

[Part (a) Reprinted with permission from Nature Publishing Group, from S. Ruchaud, M. Carmena, and W. C. Earnshaw, 2007, "Chromosomal Passengers: Conducting Cell Division," *Nat. Rev. Mol. Cell Biol.* **8**:798–812; permission conveyed through Copyright Clearance Center, Inc.]

The second important aspect of the contractile ring is the timing of its contraction. If it were to contract before all of the chromosomes had moved to their respective poles, disastrous genetic consequences would ensue. As we discuss in Chapter 19, a signaling pathway has been discovered in budding yeast called the *spindle position checkpoint*, which pauses the cell cycle to ensure that cytokinesis does not occur until the spindle is appropriately oriented. The mechanism of this coordination in animal cells is still being unraveled.

Plant Cells Reorganize Their Microtubules and Build a New Cell Wall

in Mitosis

Interphase plant cells lack a central MTOC that organizes microtubules into the radiating array typical of animal cells. Instead, numerous MTOCs containing γ-tubulin line the cortex of plant cells and nucleate the assembly of transverse bands of microtubules below the cell wall (Figure 18-48, *left*). These microtubules, which are of mixed polarity, are released from the cortical MTOCs by the action of katanin, a microtubule-severing protein; loss of katanin gives rise to very long microtubules and misshapen cells. The reason for this is that these cortical microtubules, which are cross-linked by plant-specific MAPs, aid in laying down extracellular cellulose microfibrils, the main component of the rigid cell wall (see Figure 20-43).

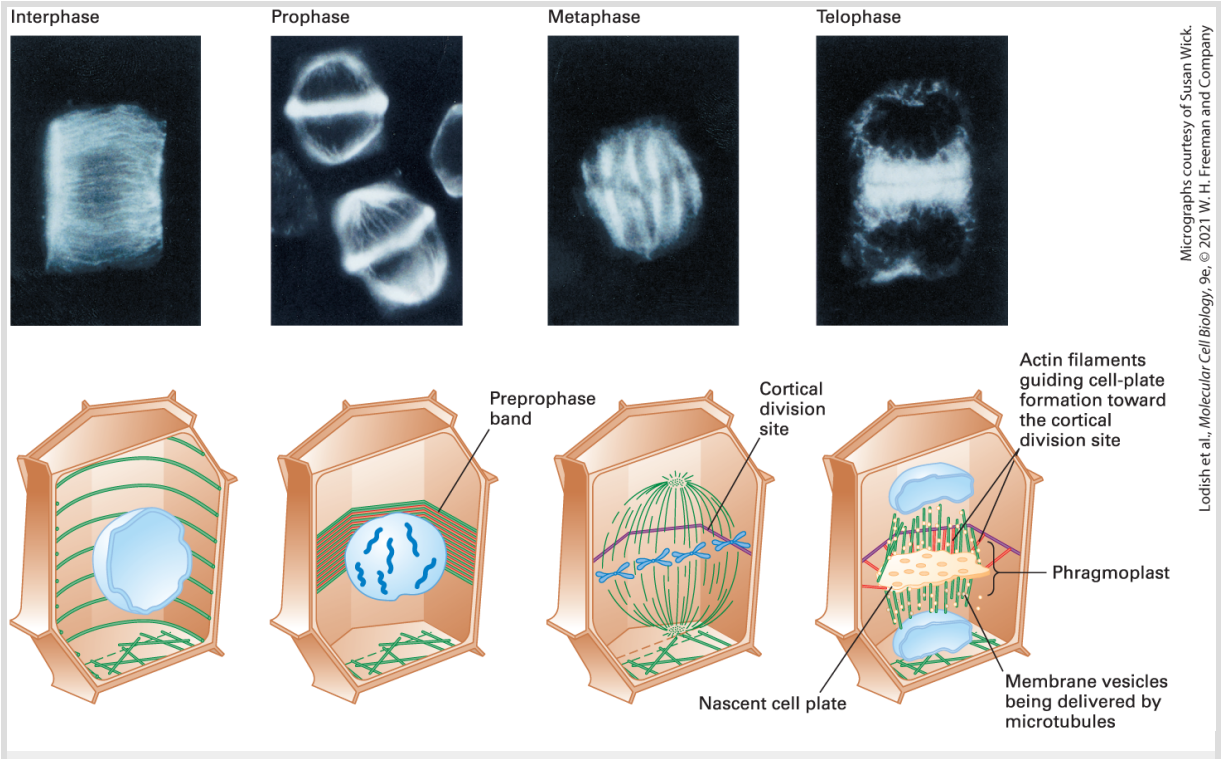


FIGURE 18-48 Mitosis in a flowering plant cell. Immunofluorescence micrographs (*top*) and corresponding diagrams (*bottom*) showing arrangement of microtubules in interphase and mitotic plant cells. A cortical array of microtubules girdles a cell during interphase. As the cell enters prophase, the microtubules (green), together with actin filaments (red), assemble under the cell cortex into a preprophase band, which marks the future cortical division site. As the cell enters prometaphase and metaphase, a spindle similar to that seen in animal cells forms. However, due to the cell wall, cytokinesis in plant cells is very different from that in animal cells. Microtubules deliver vesicles whose membranes are used to assemble a membrane network called a *phragmoplast*, whose organization is defined by actin filaments linked to the cortical division site. Eventually, the phragmoplast becomes part of the plasma membranes of the two daughter cells. Enzymes secreted from the vesicles then build a cell wall between the two daughter cells. See G. Jürgens, 2005, *Annu. Rev. Plant Biol.* **56**:281–299.

[Micrographs courtesy of Susan Wick.]

Although mitotic events in plant cells are generally similar to those in animal cells, formation of the spindle and cytokinesis has unique features in plants (Figure 18-48). Plant cells bundle their cortical microtubules and actin filaments into a *preprophase band* and reorganize them into a spindle at prophase without the aid of centrosomes. The site of the preprophase band defines the later cortical division site. At metaphase, the mitotic apparatus appears much the same in plant and animal cells. Because plants have cell walls, the division of the cell into two is quite different from animal cells and requires the assembly of a new wall between the daughter cells. Golgi-derived vesicles, which appear at telophase, are transported along microtubules to form the *nascent cell plate*. The cell plate expands and is guided toward the division site by actin filaments to form the **phragmoplast**, a membrane structure that replaces the animal-cell contractile ring. The membranes of the

vesicles forming the phragmoplast become the plasma membranes of the daughter cells. The contents of these vesicles, such as polysaccharide precursors of cellulose and pectin, form the early cell plate, which develops into the new cell wall between the daughter cells. ■

KEY CONCEPTS OF SECTION 18.6

Mitosis

- Mitosis — the accurate separation of duplicated chromosomes — involves a molecular machine comprising dynamic microtubules and microtubule-associated motors.
- The mitotic spindle has three classes of microtubules, all emanating from the spindle poles: kinetochore microtubules, which attach to chromosomes; polar microtubules, which extend from each spindle pole and overlap in the middle of the spindle; and astral microtubules, which extend to the cell cortex (see Figure 18-40).
- In the first stage of mitosis, prophase, the nuclear chromosomes condense and the spindle poles move to either side of the nucleus (see Figure 18-38).
- At prometaphase, the nuclear envelope breaks down and microtubules emanating from the spindle poles capture sister chromatid pairs at their kinetochores. The two kinetochores (one on each chromatid) become attached to opposite spindle poles (bi-oriented), which allows the chromosome to congress to the middle of the spindle.
- The chromosomal passenger complex (CPC) associated with the inner kinetochore keeps microtubule attachments weak by the activity of its kinase component Aurora B, which phosphorylates critical kinetochore proteins. When a chromosome is bi-oriented, tension is generated, and the Aurora B substrates are pulled away from the kinase (see Figure 18-44). Without phosphorylation of kinetochore proteins by Aurora B, the chromosome-kinetochore attachment becomes stable.
- At metaphase, chromosomes are aligned on the metaphase plate. The spindle assembly checkpoint pathway monitors unattached kinetochores and delays anaphase until all chromosomes are attached.
- At anaphase, duplicated chromosomes are separated and moved toward the spindle poles by shortening of the kinetochore microtubules at both the kinetochore and spindle pole (anaphase A). The spindle poles also move apart, pushed by bipolar kinesin-5 moving toward the (+) ends of the polar microtubules (anaphase B).

Spindle separation is also facilitated by cortically located dynein pulling on astral microtubules (see Figure 18-45).

- The spindle is centered and oriented by two signals — a Ran·GTP gradient from the chromosomes and a polo-like kinase gradient from the poles — that negatively regulate the presence or activity of cortical anchors for dynein/dynactin. Active dynein pulls on astral microtubules to center the spindle.
- The position of the contractile ring is directed by the CPC complex that localizes to the middle of the spindle during anaphase. The CPC recruits factors to generate active Rho·GTP at the adjacent plasma membrane to drive the assembly of actin filaments and activate myosin-II. The contractile ring then contracts to pinch the cell in two during cytokinesis (see Figure 18-47).
- In plants, cell division involves the delivery of membranes by microtubules to assemble the phragmoplast, which becomes the plasma membrane of the two daughter cells (see Figure 18-48).

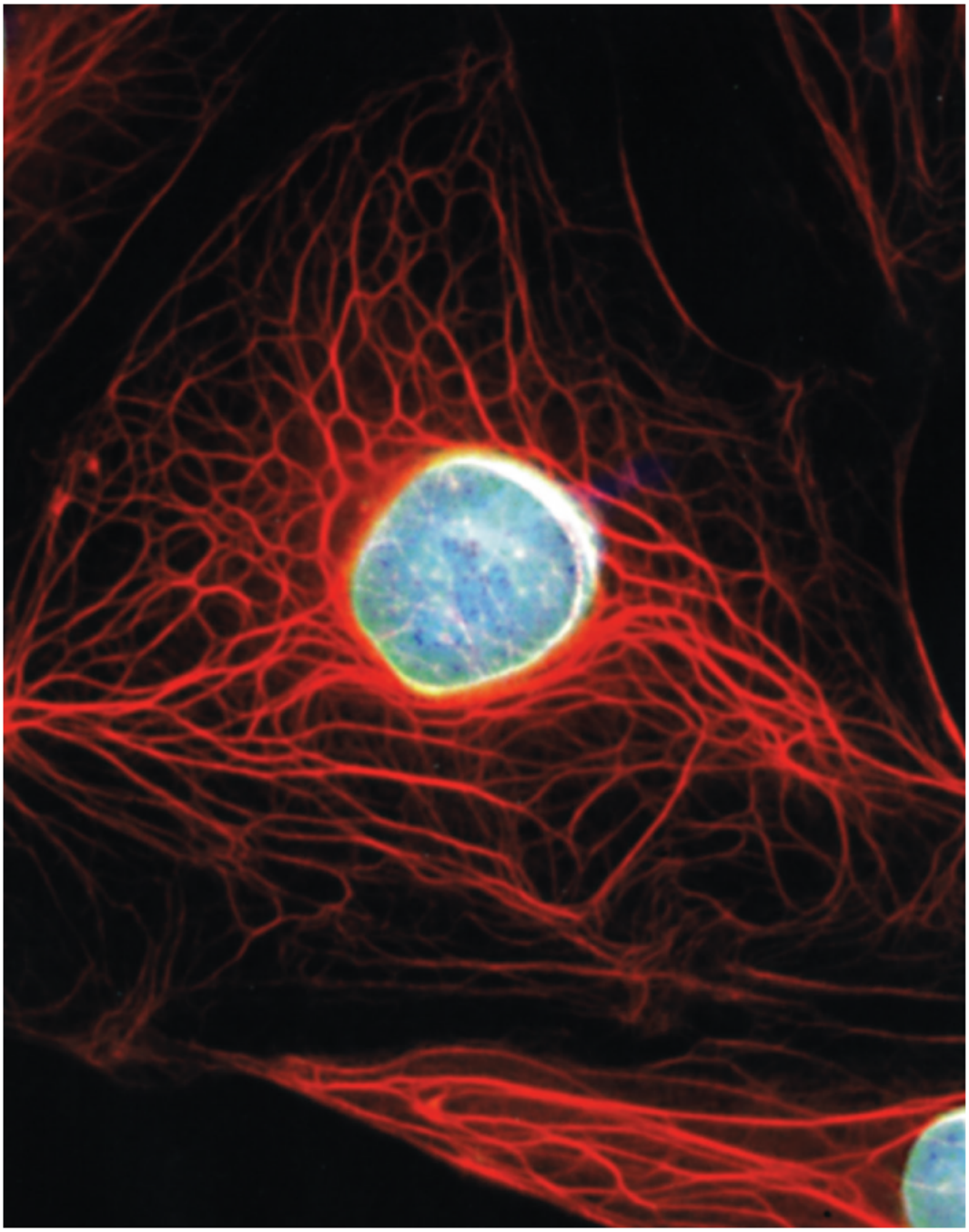
18.7 Intermediate Filaments

The third major cytoskeletal system of animals is the intermediate filaments (IFs). This name reflects their diameter of about 10 nm, which is intermediate between the 7–8-nm diameter of microfilaments and the 25-nm myosin thick filaments of skeletal muscle. IFs extend throughout the cytoplasm as well as lining the inner nuclear envelope of interphase animal cells (Figure 18-49). IFs have several unique properties that distinguish them from microfilaments and microtubules.

1. They are biochemically much more heterogeneous — that is, many different, but evolutionarily related, IF subunits exist — and are often expressed in a tissue-dependent manner.
2. They have great tensile strength, as is demonstrated by hair and nails, which consist primarily of the intermediate filaments of dead cells.
3. They do not have an intrinsic polarity like microfilaments and microtubules, and their constituent subunits do not bind a nucleotide.
4. Because they have no intrinsic polarity, it is not surprising that there are no motors that use them as tracks.
5. They are much more stable than microfilaments and microtubules, because the subunit exchange rate is much slower. Indeed, a standard way to purify intermediate filaments is to subject cells to harsh extraction conditions in a detergent so that all membranes, microfilaments, and microtubules are

solubilized, leaving a residue that is almost exclusively intermediate filaments.

6. Intermediate filaments are not found in all eukaryotes. Fungi and plants lack intermediate filaments, and insects have only one class, represented by two genes.



Courtesy of Robert D. Goldman.

EXPERIMENTAL FIGURE 18-49 Localization of two types of intermediate filaments in an epithelial cell. Immunofluorescence micrograph of an epithelial cell doubly stained with antibodies to keratin (red) and lamin (blue). A meshwork of lamin intermediate filaments can be seen underlying the nuclear membrane, whereas the keratin filaments extend from the nucleus to the plasma membrane.

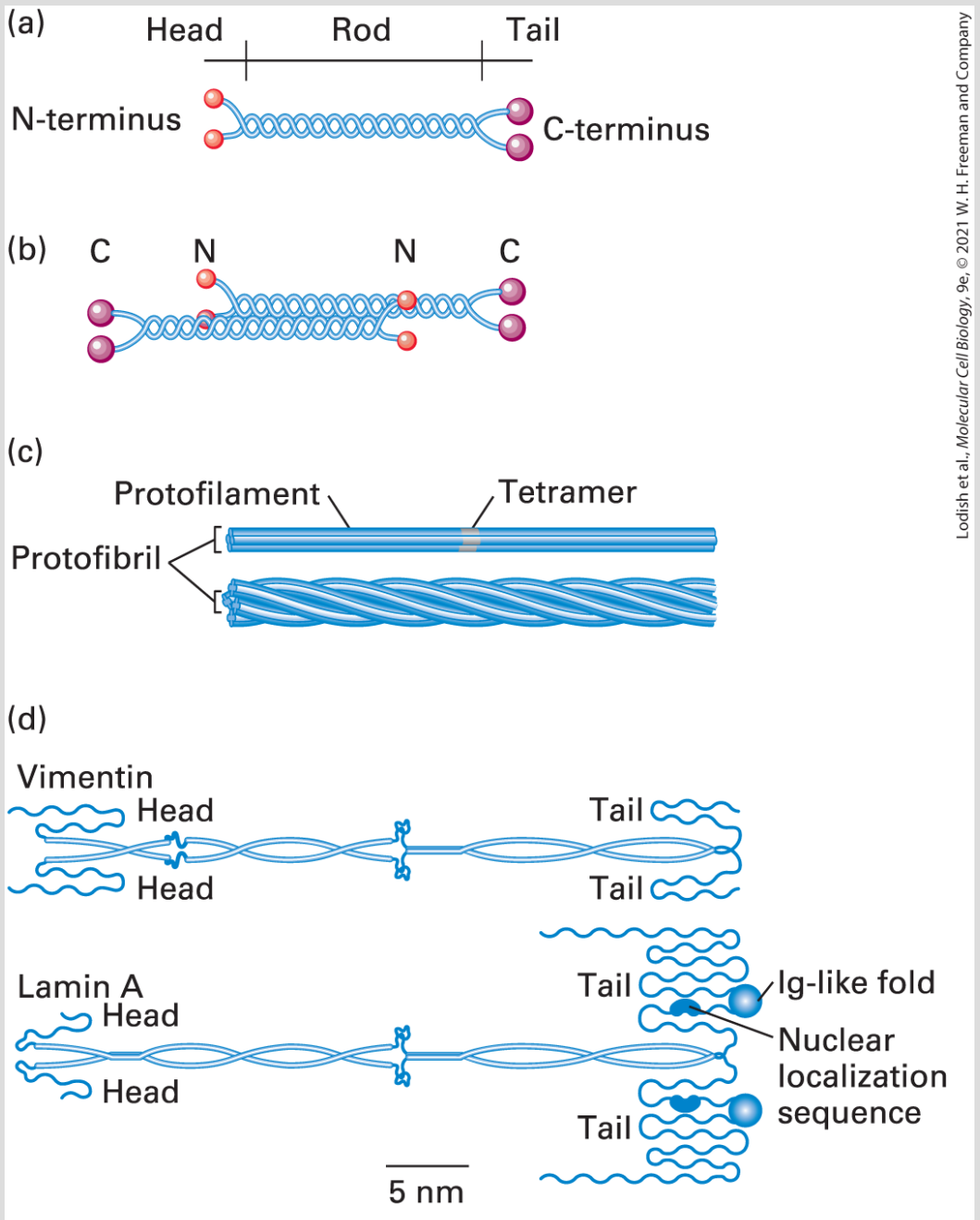
These properties make intermediate filaments unique and important structures of metazoans. Their importance is underscored by the identification of hundreds of clinical disorders, some of which are discussed here, associated with mutations in genes encoding IF proteins. To understand their contributions to cell and tissue structure, we first examine the structure of IF proteins and see how they assemble into filaments. Second, we discuss the dynamics of IFs, and third, we describe the different classes of intermediate filaments and the functions they perform.

Intermediate Filaments Are Assembled from Subunit Dimers

In the human genome intermediate filaments are encoded by 70 different genes in at least five classes. The defining feature of IF proteins is the presence of a conserved α-helical rod domain of about 310 residues that has the sequence features of a coiled-coil motif (see Figure 3-7a).

The primary building block of intermediate filaments is a dimer held together through the rod domains, which associate as a coiled coil (Figure 18-50a). Two dimers then associate in an offset fashion to make tetramers, in which the two dimers have opposite orientations (Figure 18-50b). Tetramers are assembled end to end and interlocked into long *protofilaments*. Four protofilaments

associate into a *protofibril*, and four protofibrils associate side to side to generate the 10-nm filament. Thus an intermediate filament has 16 protofilaments in it (Figure 18-50c). Flanking the rod domain of each dimer are nonhelical N- and C-terminal domains of different sizes, characteristic of each IF class (Figure 18-50d). Because the tetramer is symmetric, intermediate filaments have no polarity. This description of the filament is based on its structure rather than its mechanism of assembly. It is not clear how intermediate filaments are assembled in vivo. Unlike microfilaments and microtubules, there are no known intermediate filament nucleating, sequestering, capping, or filament-severing proteins.



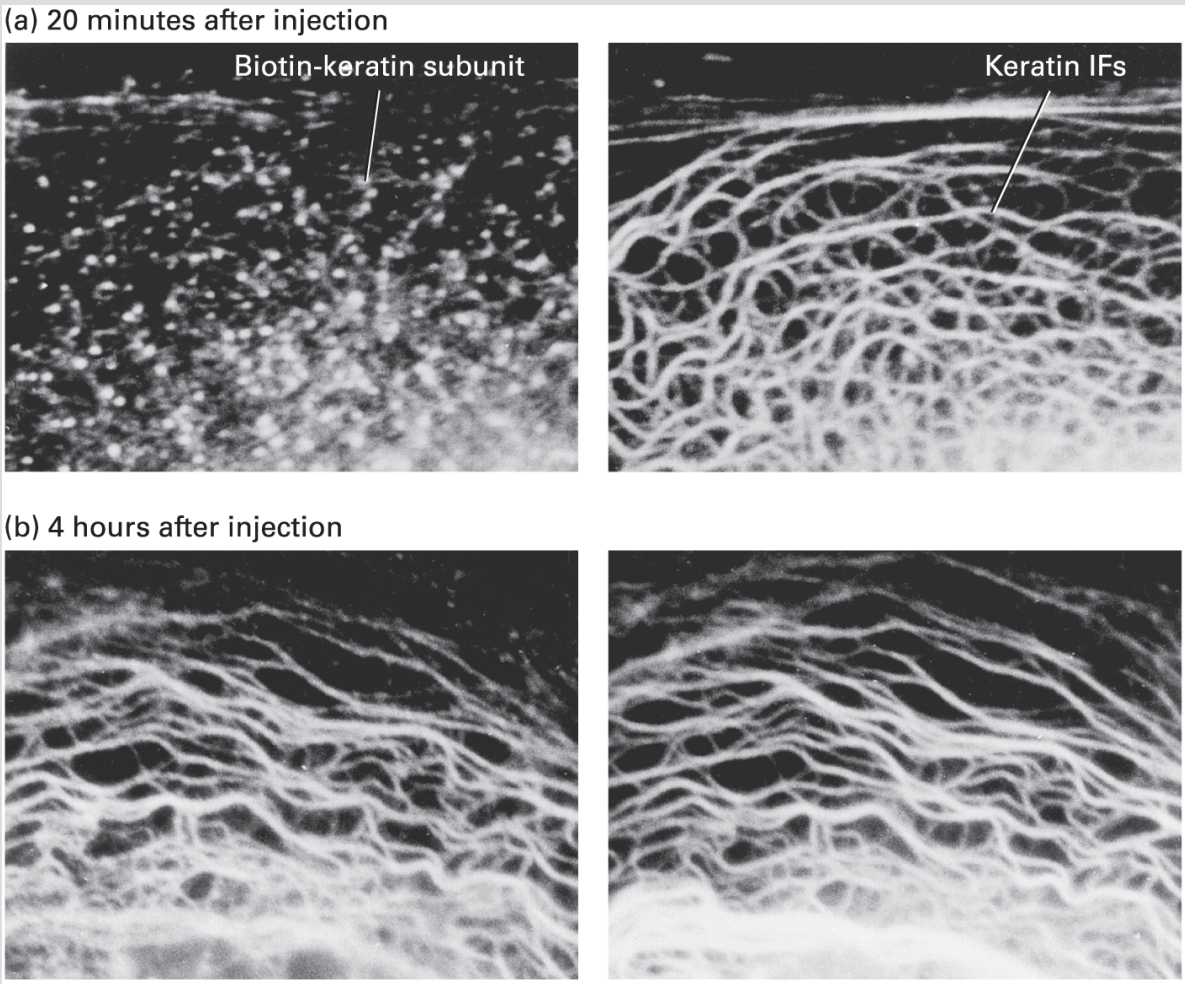
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 18-50 Structure and assembly of intermediate filaments. Drawings of IF protein dimers, tetramers, and mature intermediate filaments. (a) IF proteins form parallel dimers through a highly conserved coiled-coil core domain. The globular heads and tails are quite variable in length (as shown in more detail in panel d) and sequence among IF classes. (b) A tetramer is formed by antiparallel, staggered, side-by-side association of two identical dimers. (c) Tetramers associate end to end and laterally into a protofibril. In a mature filament, consisting of four protofibrils, the globular domains form beaded

clusters on the surface. (d) Comparison of the organization of dimers of vimentin and lamin A. Notice that the lamin protein has a nuclear localization sequence to target it to the nucleus. See H. Hermann et al., 2007, *Nat. Rev. Mol. Cell Biol.* **8**:562.

Intermediate Filaments Are Dynamic

Although intermediate filaments are much more stable than microtubules and microfilaments, IF protein subunits have been shown to be in dynamic equilibrium with the existing IF cytoskeleton. In one experiment, biotin-labeled keratin was injected into fibroblasts; within 2 hours, the labeled protein had been incorporated into the already existing keratin cytoskeleton (Figure 18-51). The results of this experiment and others demonstrate that IF subunits in a soluble pool are able to add themselves to preexisting filaments and that subunits are able to dissociate from intact filaments.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

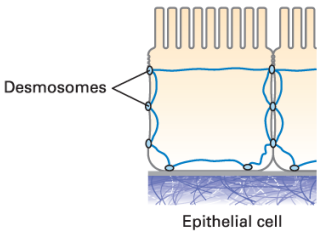
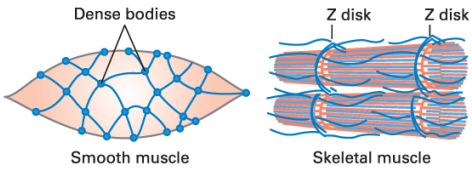
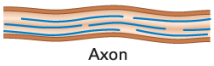
EXPERIMENTAL FIGURE 18-51 Keratin intermediate filaments are dynamic, as soluble keratin is incorporated into filaments. Monomeric type I keratin was purified, chemically labeled with biotin, and microinjected into living epithelial cells. The cells were then fixed at different times after injection and stained with an antibody to biotin and with antibodies to keratin. (a) At 20 minutes after injection, the injected biotin-labeled keratin is concentrated in small foci scattered throughout the cytoplasm (*left*) and has not been integrated into the endogenous keratin cytoskeleton (*right*). (b) After 4 hours, the biotin-labeled keratin (*left*) and the keratin filaments (*right*) display identical patterns, indicating that the microinjected protein has become incorporated into the existing cytoskeleton.

[© 1991 R. K. Miller, K. Vistrom, and R. D. Goldman, 1991, *J. Cell Biol.* **113**(4):843–855; https://doi.org/10.1083/jcb.113.4.843.]

Cytoplasmic Intermediate Filament Proteins Are Expressed in a Tissue-Specific Manner

Sequence analysis of IF proteins reveals that they fall into five homology classes, including four localized to the cytoplasm. These IF classes show a strong correspondence to the developmental origin of the cell type in which the IF protein is expressed (Table 18-2). We discuss the fifth class — the nuclear lamins — separately, as they perform functions distinct from the cytoplasmic intermediate filaments.

TABLE 18-2 • The Major Classes of Intermediate Filaments in Mammals

Class	Protein	Distribution	Proposed Function	
I	Acidic keratins	Epithelial cells	Tissue strength and integrity	
II	Basic keratins	Epithelial cells		
III	Desmin, GFAP, vimentin	Muscle, glial cells, mesenchymal cells	Sarcomere organization, integrity	
IV	Neurofilaments (NFL, NFM, and NFH)	Neurons	Axon organization	
V	Lamins	Nucleus	Nuclear structure and organization	

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

The **keratins** that make up IF protein classes I and II are found in epithelia; class III IF proteins are generally found in cells of mesodermal origin; and class IV IF proteins compose the **neurofilaments** found in neurons. The **lamins**, which make up class V, are found lining the nuclei of all animal tissues. Here we briefly summarize the four homology classes found in the cytoplasm and discuss their roles in specific tissues.

Keratins

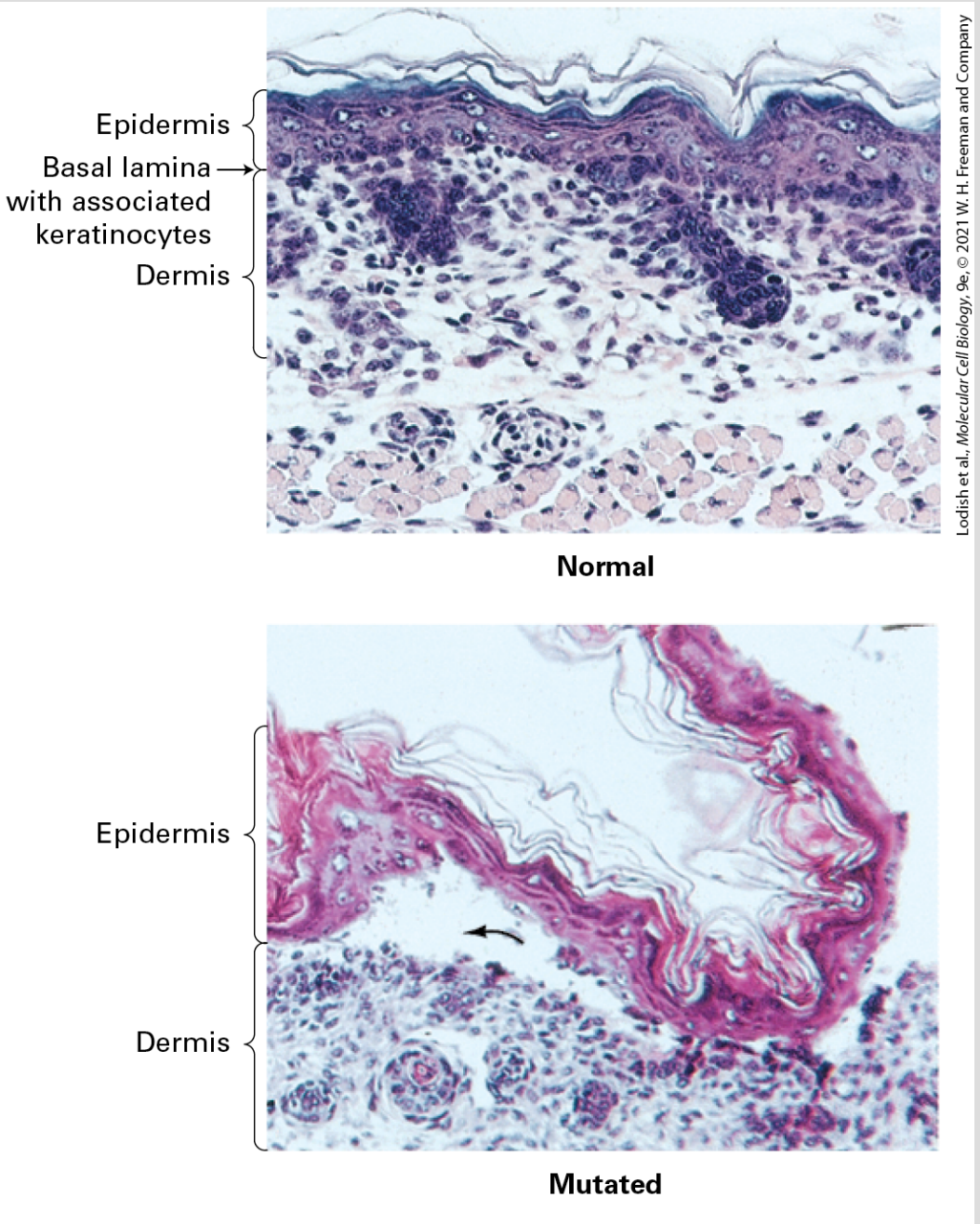
Keratins provide strength to epithelial cells. The first two IF protein homology classes are the so-called *acidic* and *basic keratins*. There are about 50 genes in the human genome encoding keratins, about evenly split between the acidic and basic classes. Dimers form consisting of one basic keratin chain and one acidic keratin chain.

These dimers then assemble into a filament as described in the previous section.

The keratins are by far the most diverse of the IF protein families. Basic and acidic keratin pairs exhibit different expression patterns both among the various types of epithelia as well as during cell differentiation. The hard keratins make up hair and nails. These keratins are rich in cysteine residues that become oxidized to form disulfide bridges, thereby strengthening the proteins. This property is exploited by hair stylists. If you do not like the shape of your hair, the disulfide bonds in your hair keratin can be reduced, the hair reshaped, and the disulfide bonds re-formed by oxidation — the result is so-called "permanent" hair curling or hair straightening.

The soft keratins, or *cytokeratins*, are found in epithelial cells. The epidermal-cell layers that make up the skin provide a good example of the function of these keratins (Figure 18-52). The lowest layer of cells, the *basal layer*, which is in contact with the basal lamina, proliferates constantly, giving rise to cells called *keratinocytes*. After they leave the basal layer, the keratinocytes differentiate and express abundant cytokeratins. The cytokeratins associate with specialized attachment sites between cells, creating sheets of cells that can withstand abrasion. The keratinocytes eventually die, leaving dead cells with their cytokeratin network but devoid of all cell organelles. This dead cell layer provides an essential barrier to water evaporation, without which we could not survive. The life of a

skin cell, from birth to its loss from the animal as a skin flake, is about one month.



EXPERIMENTAL FIGURE 18-52 Transgenic mice carrying a mutant keratin gene exhibit blistering similar to that in the human disease epidermolysis bullosa simplex. Histological sections through the skin of a normal mouse and a transgenic mouse carrying a mutant K14 keratin gene are shown. In the

normal mouse, the skin consists of a hard outer epidermal layer that covers and is in contact with the soft inner dermal layer. In the skin from the transgenic mouse, the two layers are separated (arrow) due to weakening of the cells at the base of the epidermis.

[Republished with permission from Elsevier, from P. A. Coulombe et al., 1991, "Point Mutations in Human Keratin 14 Genes of Epidermolysis Bullosa Simplex Patients," *Cell* **66**(6):1301–1311; permission conveyed through Copyright Clearance Center, Inc.]

In all epithelia, keratin filaments associate with desmosomes, which link adjacent cells together, and hemidesmosomes, which link cells to the extracellular matrix, thereby giving cells and tissues their strength. These structures are described in more detail in Chapter 20.

In addition to simply providing structural support, there is increasing evidence that keratin filaments provide some organization to organelles and participate in signal transduction pathways. For example, in response to tissue injury, rapid cell growth is induced. It has been shown that in epithelial cells, the growth signal requires an interaction between a cell-growth-signaling molecule and a specific keratin.

Desmin

The class III IF proteins include *vimentin*, found in mesenchymal cells; *GFAP (glial fibrillary acidic protein)*, found in glial cells; and

desmin, found in muscle cells. Desmin provides strength and organization to muscle cells (see cartoons in Table 18-2).

In smooth muscle, desmin filaments link cytoplasmic *dense bodies* to the plasma membrane to ensure that cells resist overstretching. In skeletal muscle, a lattice composed of a band of desmin filaments surrounds the sarcomere. The desmin filaments encircle the Z disk and are cross-linked to the plasma membrane. Longitudinal desmin filaments cross to neighboring Z disks within the myofibril, and connections between desmin filaments around Z disks in adjacent myofibrils serve to cross-link myofibrils into bundles within a muscle cell. The lattice is also attached to the sarcomere through interactions with myosin thick filaments. Because the desmin filaments lie outside the sarcomere, they do not actively participate in generating contractile forces. Rather, desmin plays an essential structural role in maintaining muscle integrity. In transgenic mice lacking desmin, for example, this supporting architecture is disrupted and Z disks are misaligned. The locations and morphology of mitochondria in these mice are also abnormal, suggesting that these intermediate filaments may also contribute to the organization of organelles.

Neurofilaments

Type IV intermediate filaments consist of the three related subunits — NF-L, NF-M, and NF-H (for NF light, medium, and heavy) — that

make up the neurofilaments found in the axons of neurons (see Figure 18-2). The three subunits differ mainly in the size of their C-terminal domains, and all form obligate heterodimers. Experiments with transgenic mice reveal that neurofilaments are necessary to establish the correct diameter of axons, which determines the rate at which nerve impulses are propagated down them.

The structural integrity of the skin is essential in order to withstand abrasion. In humans and mice, the K4 and K14 keratin isoforms form heterodimers that assemble into protofilaments. A mutant K14 with deletions in either the N- or the C-terminal domain can form heterodimers in vitro but does not assemble into protofilaments. The expression of such mutant keratin proteins in cells causes IF networks to break down into aggregates. Transgenic mice that express a mutant K14 protein in the basal stem cells of the epidermis display gross skin abnormalities, primarily blistering of the epidermis, that resemble the human skin disease *epidermolysis bullosa simplex*. Histological examination of the blistered area reveals a high incidence of dead basal cells. The death of these cells appears to be caused by mechanical trauma from rubbing of the skin during movement of the limbs. Without their normal bundles of keratin filaments, the mutant basal cells become fragile and easily damaged, causing the overlying epidermal layers to delaminate and blister (see Figure 18-52). Like the role of desmin filaments in supporting muscle tissue, the general role of keratin filaments appears to maintain the structural integrity of epithelial

tissues by mechanically reinforcing the connections between cells.

Lamins Line the Inner Nuclear Envelope to Provide Organization and Rigidity to the Nucleus

The most widespread IF proteins are the class V proteins, the lamins. Lamins are the progenitors of all IF proteins, from which the cytoplasmic IF proteins arose by gene duplication and mutation. They are the major components of a two-dimensional meshwork called the *nuclear lamina* that lies between the nuclear envelope and the chromatin of the nucleus (Figure 18-53). In humans, three genes encode lamins: one alternatively spliced gene encodes lamins A and C, and two other genes encode lamins B1 and B2. The B-type lamins appear to be the primordial lamin proteins and are expressed in essentially all cells, whereas lamins A and C are developmentally regulated. B lamins are post-translationally prenylated (see Section 10.2), which helps them associate with the inner nuclear envelope membrane. Lamin proteins contain the coiled-coil regions characteristic of intermediate filaments that are needed for dimerization but also have a nuclear localization sequence that targets them to the nucleus as well as a conserved immunoglobulin-like fold (see Figure 18-50d).

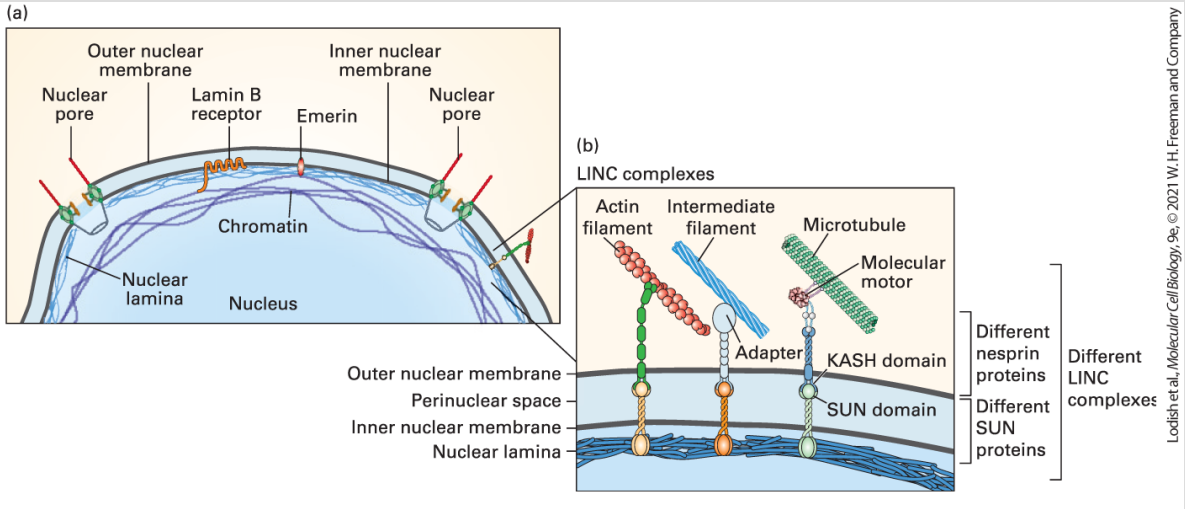


FIGURE 18-53 The nuclear lamina is attached to chromatin and through LINC complexes to the cytoskeleton. (a) Diagram of part of a nucleus, showing the association of the lamin-containing nuclear lamina with chromatin and, through the two membranes of the nucleus, with the cytoskeleton. Proteins such as the membrane-associated lamin B receptor and emerin tether the lamin intermediate filaments to the inner nuclear membrane. Lamins are also tethered to the inner nuclear membrane by the prenylation of lamin B (not shown). Diverse linkages, called *LINC complexes*, attach the lamins through the two nuclear membranes to the cytoskeleton. (b) A LINC (*l*inker of *n*ucleoskeleton and *c*ytoskeleton) complex consists of a SUN domain–containing protein that interacts with lamins and extends across the inner nuclear membrane, and a KASH domain–containing protein that interacts with a SUN domain–containing protein in the perinuclear space and crosses the outer nuclear membrane to interact with components of the cytoskeleton. See C. S. Janota et al., 2017, *Cell* **169**:970.

Cells that are subject to mechanical stress have a nuclear lamin meshwork to maintain the integrity of the nucleus. Cells with strong cell walls, such as plant and fungal cells, do not have such a meshwork. The lamin meshwork therefore provides both strength and support to the inner surface of the nuclear membrane. In fact, cells regulate the level of lamin A to match the stiffness of the tissue they are in. Thus neutrophils, for example, which need to move

through thin capillaries and migrate through tight interstitial spaces, have a highly lobulated nucleus as a result of very low levels of lamin A. If they had large, rigid nuclei that were resistant to deformation, they might have trouble squeezing through the small spaces found in the extracellular matrix.

To provide rigidity, the lamin meshwork is associated with chromatin on one side and attached to the cytoskeleton on the other. Some proteins embedded in the inner nuclear membrane, such as the lamin-B receptor and emerin, can bind both chromatin-associated proteins and lamins (see Figure 18-53a). Interestingly, transcriptionally silent regions of the genome are preferentially associated with lamins, and recent evidence suggests that lamins play a role in genome organization and DNA repair. Attachment to the cytoskeleton through both the inner and outer nuclear membranes involves proteins with so-called SUN and KASH domains. The SUN domain proteins are synthesized on the endoplasmic reticulum as transmembrane proteins with their SUN domain in the lumen of the endoplasmic reticulum, and sorting signals in their cytoplasmic domain target them to the outer nuclear membrane, which is continuous with the endoplasmic reticulum. They are then transported as membrane proteins through nuclear pores and, upon reaching the inner nuclear membrane, associate with the nuclear lamina (Figure 18-53b). Nesprins are KASH domain–containing transmembrane proteins of the outer nuclear membrane oriented so that the KASH domain can associate with the SUN domain of another protein in the perinuclear space. The

nesprins, in turn, associate, either directly or through adapters, with intermediate filaments, actin filaments, and microtubules, thereby linking the nucleus to the cytoskeleton (see Figure 18-53b). These attachments are used to move the nucleus to the correct location in a cell as well as transport it, for example, in the long processes of the vertebrate neuroepithelium.

Lamins Are Reversibly Disassembled by Phosphorylation During Mitosis

For the nuclear envelope to be broken down as cells go from prophase to prometaphase during mitosis, the nuclear lamina has to be disassembled. As we discuss in Chapter 19, protein kinases called *mitotic CDKs* drive cells into mitosis, and one of their substrates is the lamins. Phosphorylation of lamins A, B, and C results in disassembly of the intermediate filament lattice into lamin dimers. Due to their C-terminal prenylation, lamin B dimers remain associated with the nuclear membrane. As we saw when we discussed mitosis, depolymerization of the nuclear lamin filaments leads to the disintegration of the nuclear lamina meshwork and contributes to disassembly of the nuclear envelope. During telophase, de-phosphorylation by specific phosphatases promotes lamin reassembly, which is critical to re-formation of a nuclear envelope around the daughter chromosomes. The opposing actions of kinases and phosphatases thus provide a rapid mechanism for

controlling the assembly state of lamin intermediate filaments. Other intermediate filaments undergo similar disassembly and reassembly in other phases of the cell cycle.

There are over 600 mutations located throughout the human gene for lamin A that are known to cause diseases, collectively called *laminopathies*. These diseases include cardiomyopathies, muscular dystrophies, lipodystrophy, and aging-related progeria. Some of these mutations cause Emery-Dreifuss muscular dystrophy (EDMD), most likely because the fragile nuclei cannot stand the stress and strains in the muscle tissue, so these cells are the first to show symptoms. Other forms of EDMD have been traced to mutations in emerin, the lamin-binding membrane protein of the inner nuclear envelope, as well as a nesprin and a SUN protein. Yet other mutations in lamin A cause progeria — accelerated aging, such as Hutchison-Gilford progeria ("prematurely old"). It is still a mystery why different mutations in the same human gene can cause such a wide variety of diseases. ■

KEY CONCEPTS OF SECTION 18.7

Intermediate Filaments

- Intermediate filaments are the only nonpolar fibrous component of the cytoskeleton and are not associated with motor proteins. Intermediate filaments are built from coiled-coil dimers that associate in an antiparallel fashion into tetramers and then into protofilaments, 16 of which make up the filament (see Figure 18-50).
- There are five major classes of intermediate filament proteins, with the nuclear lamins (class V) being the most ancient and ubiquitous in animal cells. The other four classes

show tissue-specific expression (see Table 18-2).

- Keratins (IF classes I and II) are found in animal hair and nails, as well as in cytokeratin filaments that associate with desmosomes in epithelial cells to provide the cells and tissue with strength.
- The class III filaments include vimentin, GFAP, and desmin, which provide structure and order to muscle Z disks and restrain smooth muscle from overextension.
- The neurofilaments make up class IV and are important for the structure of axons.
- The lamins are major components of the nuclear lamina. They contribute to genome organization as well as to the rigidity of the nucleus through linkages to the cytoskeleton involving proteins with SUN and KASH domains (see Figure 18-53).
- Many diseases are associated with defects in intermediate filaments, especially laminopathies, which include a variety of conditions, and mutations in keratin genes, which can cause severe defects in skin (see Figure 18-52).

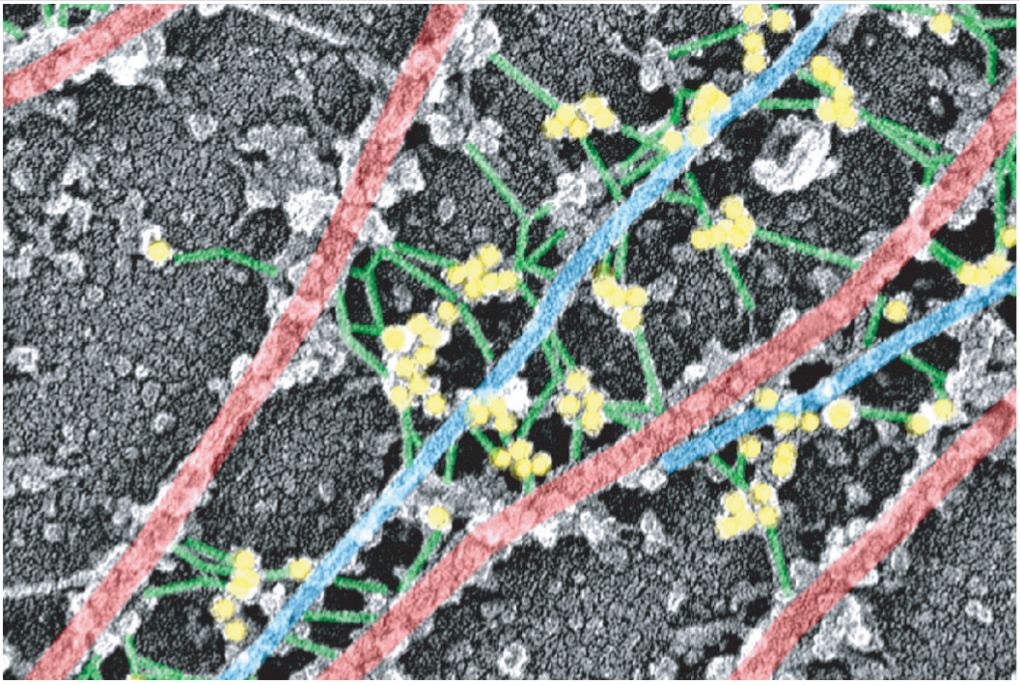
18.8 Coordination and Cooperation Between Cytoskeletal Elements

So far, we have discussed the three cytoskeletal filament classes — microfilaments, microtubules, and intermediate filaments — as though they mostly function independently of one another. But in reality that is not the case. We saw, for example, that the microtubule-based mitotic spindle determines the site of formation of the microfilament-based contractile ring. We close this chapter with some other examples of linkages, both physical and regulatory, between cytoskeletal elements and their integration into other aspects of cellular organization.

Intermediate Filament–Associated Proteins Contribute to Cellular Organization

There is a group of proteins collectively called *intermediate filament–associated proteins (IFAPs)* that co-purify with intermediate filaments. Among these IFAPs are members of the **plakin** family, which are involved in attaching intermediate filaments to other structures. Some plakins link keratin filaments to desmosomes, which are junctions between epithelial cells that provide stability to a tissue, and hemidesmosomes, which are located at regions of the plasma

membrane where intermediate filaments are linked to the extracellular matrix (these topics are covered in detail in Chapter 20). Other plakins, found bound along intermediate filaments, have binding sites for microfilaments and microtubules. One of these proteins, called *plectin*, can be seen by immunoelectron microscopy to provide connections between microtubules and intermediate filaments (Figure 18-54).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-54 Gold-labeled antibody identifies plectin cross-links between intermediate filaments and microtubules. In this immunoelectron micrograph of a fibroblast cell, microtubules are highlighted in red, intermediate filaments in blue, and the short connecting fibers between them in green. Staining with gold-labeled antibodies to plectin (yellow) reveals that these connecting fibers contain plectin.

[© 1996 T. M. Svitkina, A. B. Verkhhovsky, and G. G. Borisy, 1996, *J. Cell Biol.* **135**(4):991–1007; https://doi.org/10.1083/jcb.135.4.991.]

Microfilaments and Microtubules Cooperate to Transport Melanosomes

Studies of mutant mice with light-colored coats have uncovered a pathway in which microtubules and microfilaments cooperate to transport melanosomes. The pigment in mammalian hair is produced in cells called *melanocytes*, cells that are very similar to the fish and frog melanophores discussed earlier (see Figure 18-29). Melanocytes are found in the hair follicle at the base of the hair shaft and contain pigment-laden granules called *melanosomes*. Melanosomes are transported to the dendritic extensions of melanocytes for subsequent exocytosis to the surrounding epithelial cells. Transport to the cell periphery is mediated, just as in frog melanophores, by a kinesin family member. At the periphery, the melanosomes are then handed off to myosin V and delivered for exocytosis. If the myosin V system is defective, the melanosomes are not captured and stay in the melanocyte cell body. Thus microtubules are responsible for the long-range transport of melanosomes, whereas microfilament-based myosin V is responsible for their capture and delivery at the cell cortex. This type of division of labor — long-range transport by microtubules and short-range transport by microfilaments — has been found in many other systems, from transport in filamentous fungi to transport along axons.

Cdc42 Coordinates Microtubules and Microfilaments During Cell Migration

In Chapter 17, we discussed how the polarity of a migrating cell is regulated by small GTPase Cdc42, which results in the formation of an actin-based leading edge at the front of the cell and contraction at the back (see Figure 17-44 and Figure 18-55, step **1**). It turns out that Cdc42 activation at the cell front also leads to polarization of the microtubule cytoskeleton. This phenomenon was originally studied in wound-healing assays (see Figure 17-43), in which it was noticed that when the cells at the edge of a scratch are induced to polarize and move to fill in the empty space, the Golgi complex is moved to the front of the nucleus toward the cell front. Golgi localization at the front of the cell indicates that the centrosome moves to lie in front of the nucleus (recall that Golgi localization is dependent on the location of the MTOC; see Figures 18-1c and 18-28). Recent studies have suggested how this happens. Active Cdc42 at the front of the cell binds the polarity factor Par6, which results in the recruitment of the dynein-dynactin complex (Figure 18-55, step **2**). Cortically localized dynein-dynactin then interacts with microtubules, pulling on them to orient the centrosome and hence the whole radial array of microtubules (Figure 18-55, step **3**). This reorientation of the microtubule system leads to the reorganization of the secretory pathway to deliver secretory products, especially integrins to bind the extracellular matrix, to the front of the cell for

attachment to the substratum for cell migration (Figure 18-55, step 4).

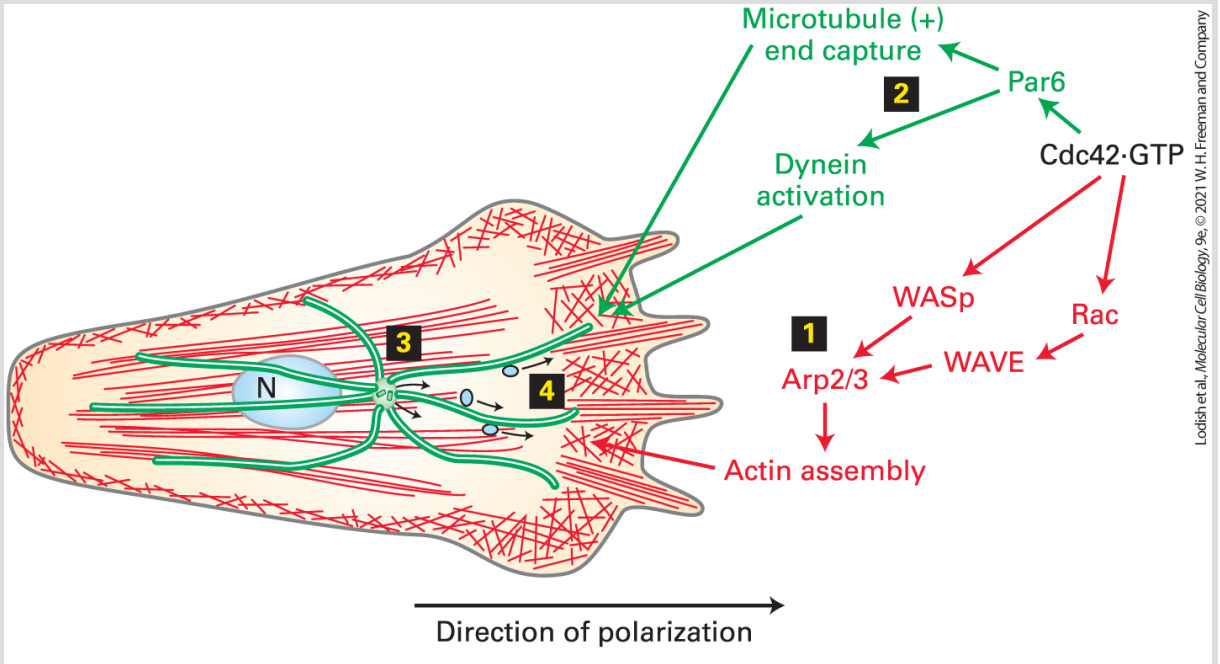
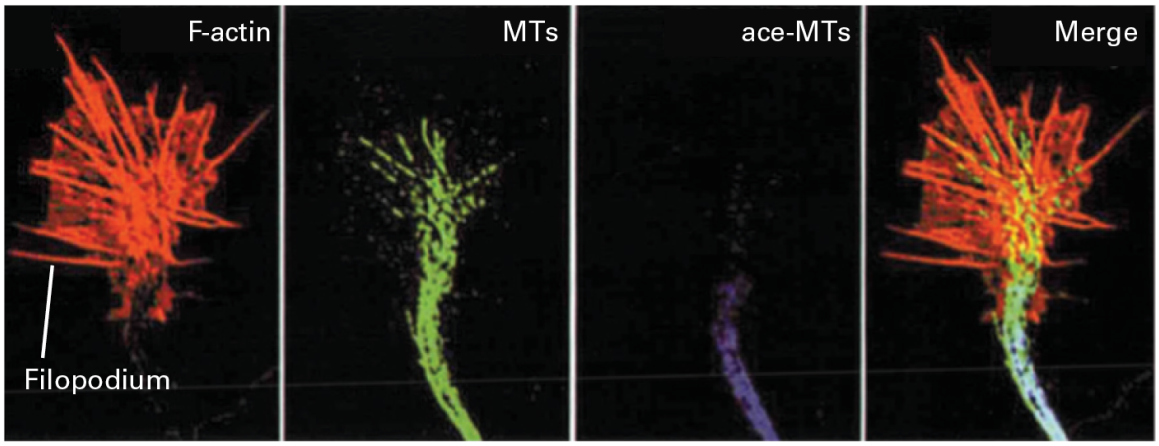


FIGURE 18-55 Independent Cdc42 regulation of microfilaments and microtubules to polarize a migrating cell. Active Cdc42·GTP at the front of the cell leads to Rac and WASp activation, which results in the assembly of a microfilament-based leading edge (step 1) (see Figure 17-44). Independently, Cdc42·GTP also leads to the capture of microtubule (+) ends and the activation of dynein (step 2). Dynein pulls on microtubules to orient the centrosome (step 3) toward the front of the cell. This reorientation reorganizes the secretory pathway for the delivery of adhesion molecules carried in secretory vesicles along microtubules to the front of the cell (step 4). See S. Etienne-Manneville et al., 2005, *J. Cell Biol.* **170**:895–901.

Advancement of Neural Growth Cones Is Coordinated by Microfilaments and Microtubules

The nervous system depends on the integration and transmission of signals by neurons. Neurons have specialized structures, called *dendrites*, that receive signals and a single axon that terminates in one or more synapses on a target cell or cells (see Figure 18-2). It is critical that neurons make the right connections; so how are growing axons guided to their correct destinations? As an axon extends, its terminal growth cone senses signals from the extracellular matrix and from other cells that guide it along the right path. Therefore, how the growth cone receives and interprets cues that direct axon growth is critical to the function of the nervous system. Growth cones are very rich in actin, and they typically have a broad lamellipodium and multiple filopodia. Microtubules are also essential for the guidance of growth cones. Recall that axons have microtubules of uniform polarity along which materials for growth of the axon move by axonal transport (see Figure 18-5e). These microtubules extend into the growth cone and, together with actin, which is necessary for the advancement of the growth cone, are involved in guiding its direction of advancement. Although the mechanisms involved have not been fully elucidated, it has been found that a local growth signal alters local actin dynamics, with the result that microtubules extend into the region of the signal. It has also been found that microtubules in the shaft of the axon have post-translational modifications, such as acetylation, that stabilize them, whereas the more dynamic microtubules in the growth cone often do not (Figure 18-56).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 18-56 Localization of actin (red), microtubules (green), and acetylated microtubules (blue) in a small growth cone. Notice how the stable acetylated microtubules are localized to the shaft of the axon and do not penetrate into the dynamic growth cone.

[Republished with permission from Elsevier, from E. W. Dent and F. B. Gertler, 2003, "Cytoskeletal Dynamics and Transport in Growth Cone Motility and Axon Guidance," *Neuron* **40**(2):209–227; permission conveyed through Copyright Clearance Center, Inc.]

KEY CONCEPTS OF SECTION 18.8

Coordination and Cooperation Between Cytoskeletal Elements

- Intermediate filaments are linked both to specific attachment sites on the plasma membrane (called *desmosomes* and *hemidesmosomes*) and to microfilaments and microtubules (see Figure 18-54).
- In animal cells, microtubules are generally used for the long-range delivery of organelles, whereas microfilaments handle their local delivery.
- The signaling molecule Cdc42 coordinately regulates microfilaments and microtubules during cell migration.
- The advancement of growth cones in neurons requires the interplay of microfilaments and microtubules.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

anaphase
aneuploidy
aster
axonal transport
axoneme
basal body
cell cycle
centriole
centromere
centrosome
cilium
cytokinesis

desmin
dynein
flagellum
intermediate filament
keratin
kinesin
kinetochore
lamin
metaphase
microtubule
microtubule-associated protein (MAP)
microtubule-organizing center (MTOC)
mitosis
mitotic spindle
motor protein
neurofilament
pericentriolar material
phragmoplast
plakin
polarity
primary cilium
prometaphase
prophase
sister chromatid
telophase
tubulin

Review the Concepts

1. Microtubules are polar filaments; that is, one end is different from the other. What is the basis for this polarity, how is polarity related to microtubule organization within the cell, and how is polarity related to the intracellular movements powered by microtubule-dependent motors?
2. Microtubules both in vitro and in vivo undergo dynamic instability, and this type of assembly is thought to be intrinsic to the microtubule. What is the current model that accounts for dynamic instability?
3. In cells, microtubule assembly depends on other proteins as well as tubulin concentration and temperature. What types of proteins influence microtubule assembly in vivo, and how does each type affect assembly?
4. Microtubules within a cell appear to be arranged in specific arrays. What cellular structure is responsible for determining the arrangement of microtubules within a cell? How many of these structures are found in a typical cell? Describe how such structures serve to nucleate microtubule assembly.
5. Many drugs that inhibit mitosis bind specifically to tubulin, microtubules, or both. What diseases are such drugs used to treat? Functionally speaking, these drugs can be divided into two groups based on their effect on microtubule assembly. What are the two mechanisms by which such drugs alter microtubule structure?

6. Kinesin-1 was the first member of the kinesin superfamily to be identified and therefore is perhaps the best-characterized superfamily member. What fundamental property of kinesin was used to purify it?
7. Certain cellular components appear to move bidirectionally on microtubules. Describe how this is possible given that microtubule orientation is fixed by the MTOC.
8. The movement of kinesin motor proteins involves both the motor domain and the linker domain. Describe the role of each domain in kinesin movement, direction of movement, or both. Could kinesin-1 with one inactive head efficiently move a vesicle along a microtubule?
9. What features of the dynactin complex enable cytoplasmic dynein to transport cargo toward the microtubule (−) end? What effect could inhibition of dynactin interaction with the +TIP EB1 have on spindle orientation?
10. Cell swimming depends on appendages containing microtubules. What is the underlying structure of these appendages, and how do these structures generate the force required to produce swimming?
11. What effect would dynein inactivation have on kinesin-2-dependent IFT?
12. The mitotic spindle is often described as a microtubule-based cellular machine. The microtubules that constitute the mitotic spindle can be classified into three distinct types. What are the three types of spindle microtubules, and what is the function of each?

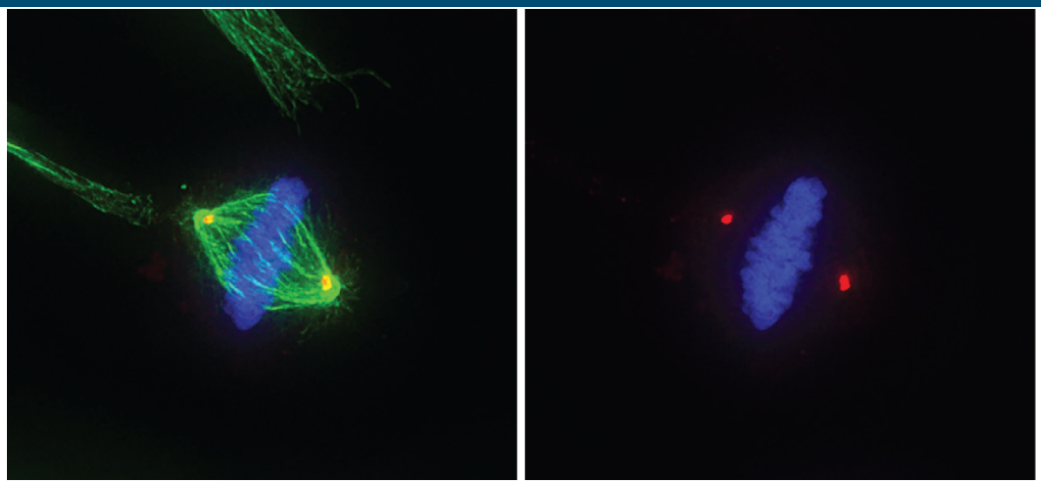
13. Mitotic spindle function relies heavily on microtubule motors. For each of the following motor proteins, predict the effect on spindle formation, function, or both of adding a drug that specifically inhibits only that motor: kinesin-5, kinesin-13, and kinesin-4.
14. The poleward movement of kinetochores, and hence chromatids, during anaphase A requires that kinetochores maintain a hold on the shortening microtubules. How does a kinetochore hold onto shortening microtubules?
15. Anaphase B involves the separation of spindle poles. What forces have been proposed to drive this separation? What underlying molecular mechanisms are thought to provide these forces?
16. Cytokinesis, the process of cytoplasmic division, occurs shortly after the separated sister chromatids have neared the opposite spindle poles. How is the plane of cytokinesis determined? What are the respective roles of microtubules and actin filaments in cytokinesis?
17. The best strategy for treating a specific type of human tumor can depend on identifying the type of cell that became cancerous to give rise to the tumor. For some tumors that have colonized a distant location (metastasized), identifying the parental cell type can be difficult. Because the type of IF protein expressed is cell-type-specific, using monoclonal antibodies that react with only one type of IF protein can help in this identification. What IF proteins would you produce monoclonal antibodies against to identify (a) a sarcoma of muscle cell

origin, (b) an epithelial cell carcinoma, and (c) an astrocytoma (glial cell tumor)?

18. Explain why there are no known motors that use intermediate filaments as tracks.
19. Growth cones are highly mobile regions of developing neurons. What prevents the growth cone from moving or collapsing back into the main cell body, as often occurs with lamellipodia?

CHAPTER 19

The Eukaryotic Cell Cycle



Jesse Patterson.

Micrograph on the left shows a human epithelial cell in metaphase. Following DNA replication, cells undergo mitosis to segregate their replicated chromosomes. The cell has aligned its chromosomes (blue) in the center of the mitotic spindle apparatus (green), prior to pulling them to opposite poles of the spindle, which will occur during anaphase. Thereafter, the cytoplasm of the cell is divided to produce two identical daughter cells. The micrograph on the right is the same cell without imaging the spindle apparatus in order to show the location of the centrosomes or spindle poles (red).

OUTLINE

19.1 Overview of the Cell Cycle
19.2 Model Organisms and Methods of Studying the Cell Cycle
19.3 Cell Cycle Progression and Control: Feedback Loops and Post-Translational Modification
19.4 The Transition from G_1 into S Phase and DNA Replication
19.5 The G_2/M Transition and the Irreversible Engine of Mitosis

19.6 The Mitotic Spindle, Chromosome Segregation, and Exit from Mitosis
19.7 Surveillance Mechanisms in Cell Cycle Regulation
19.8 Meiosis: A Special Type of Cell Division

The central tenet of life and the definition of living matter is inextricably linked to the concept of replication. The pioneering discoveries of Matthias Schleiden and Theodor Schwann, and those of Rudolf Virchow, that living matter was made of cells, and that cells arise only from the division of other cells, dictates that cells proliferate by going through repeated cycles of growth and division of a parent cell into two daughter cells. The daughter cells, in turn, repeat this process, giving rise to more cells. This behavior leads to the general concept of a *cell cycle* that incorporates each of the stages required for cells to replicate themselves: growth in size and nutrient content; duplication of subcellular organelles; DNA replication; chromosome segregation; and finally, separation of the cytoplasmic organelles, contents, and cell membrane between the two progeny. These steps in the cell cycle occur in distinct phases, and the transition from one phase to the next is abrupt and precise. Errors that occur within one phase must be detected and corrected before the cell transits into the next phase. Surveillance mechanisms prevent the transition from one phase of the cell cycle into the next phase until all of the steps of a phase have been completed properly. It is through this process of repeated cell cycles that a fertilized egg — one cell — can eventually give rise to a fully formed animal containing hundreds of millions of cells or more.

Proper control of cell division is vital to all organisms. First, cell division must be balanced with cell growth so that cell size is maintained. If several divisions occur before parent cells have reached the proper size, daughter cells eventually become too small to be viable. If cells grow too large before cell division, the cells function improperly and the number of cells increases too slowly for proper tissue, organ, and animal development during embryogenesis. Second, in developing multicellular organisms, the replication of each cell must be precisely controlled and timed with respect to the replication of other cells in order to faithfully and reproducibly complete the developmental program in every individual in order for complex organs (e.g., the brain, heart, intestine, and kidney) that are made of various cell types to develop properly. Finally, there is a balance between cell replication and cell differentiation, so that cells with high replicative potential are generally much less differentiated into specific cell types, while cells that have highly differentiated into specific lineages usually have quite limited replicative potential. Cells that have the potential to continuously replicate and have not yet differentiated (or have only partially differentiated) but have the ability to differentiate into one or more specific cell types at a later time are called *stem cells*. During embryogenesis in animals, for example, there is a period of rapid cell division after fertilization of the egg, resulting in a small number of cells that have nearly unlimited replicative potential but have not yet undergone any type of differentiation (see Chapter 22).

Their progeny, however, will later differentiate into every type of cell in the body.

In this chapter, we first present an overview of the cell cycle and then describe the various experimental systems that have contributed to our current understanding of it. We discuss the molecular mechanisms responsible for regulating cell cycle progression, particularly the roles of cyclin-dependent kinases (CDKs) as master controllers, along with other protein kinases, phosphatases, phospho-binding modules, and ubiquitin ligases that modulate their activity and function. Next we examine each cell cycle phase in greater detail, with an emphasis on how the events that take place in each phase are governed by protein post-translational modifications, CDK activity, and regulation of downstream targets. We discuss the checkpoint pathways that establish the order of the cell cycle and ensure that each cell cycle phase has been properly and accurately completed prior to transition into the next phase. The chapter concludes with a discussion of meiosis, a special type of cell division that generates haploid germ cells (eggs and sperm), and the molecular mechanisms that distinguish it from mitosis. Throughout the chapter, we emphasize the general principles governing cell cycle progression and use a species-spanning nomenclature when discussing examples of the factors controlling each cell cycle phase.

19.1 Overview of the Cell Cycle

The cell cycle in eukaryotes is a series of processes that occur in a specific sequence every time a cell divides. In cells that are continuously replicating, each process occurs on a schedule, as though the cells were following some type of molecular clock. This clock idea was part of the motivation that eventually led to the discovery of cyclin-CDK complexes as master regulators of the cell division cycle.

The cell cycle is divided into four major phases (Figure 19-1). Cycling (replicating) mammalian somatic cells grow in size and synthesize the RNAs and proteins required for DNA synthesis during the G_1 (first gap) phase. When cells have reached the appropriate size and have synthesized the required proteins, they enter the cell cycle by traversing a point in G_1 known as *START* in yeast and the *restriction point* in mammals. Once this point has been crossed, cells are committed to cell division. The first step toward successful cell division is entry into the S (synthesis) phase, the period in which cells replicate their chromosomes. After progressing through a second gap phase, the G_2 phase, cells begin the complicated process of mitosis, also called the *M (mitotic) phase*.

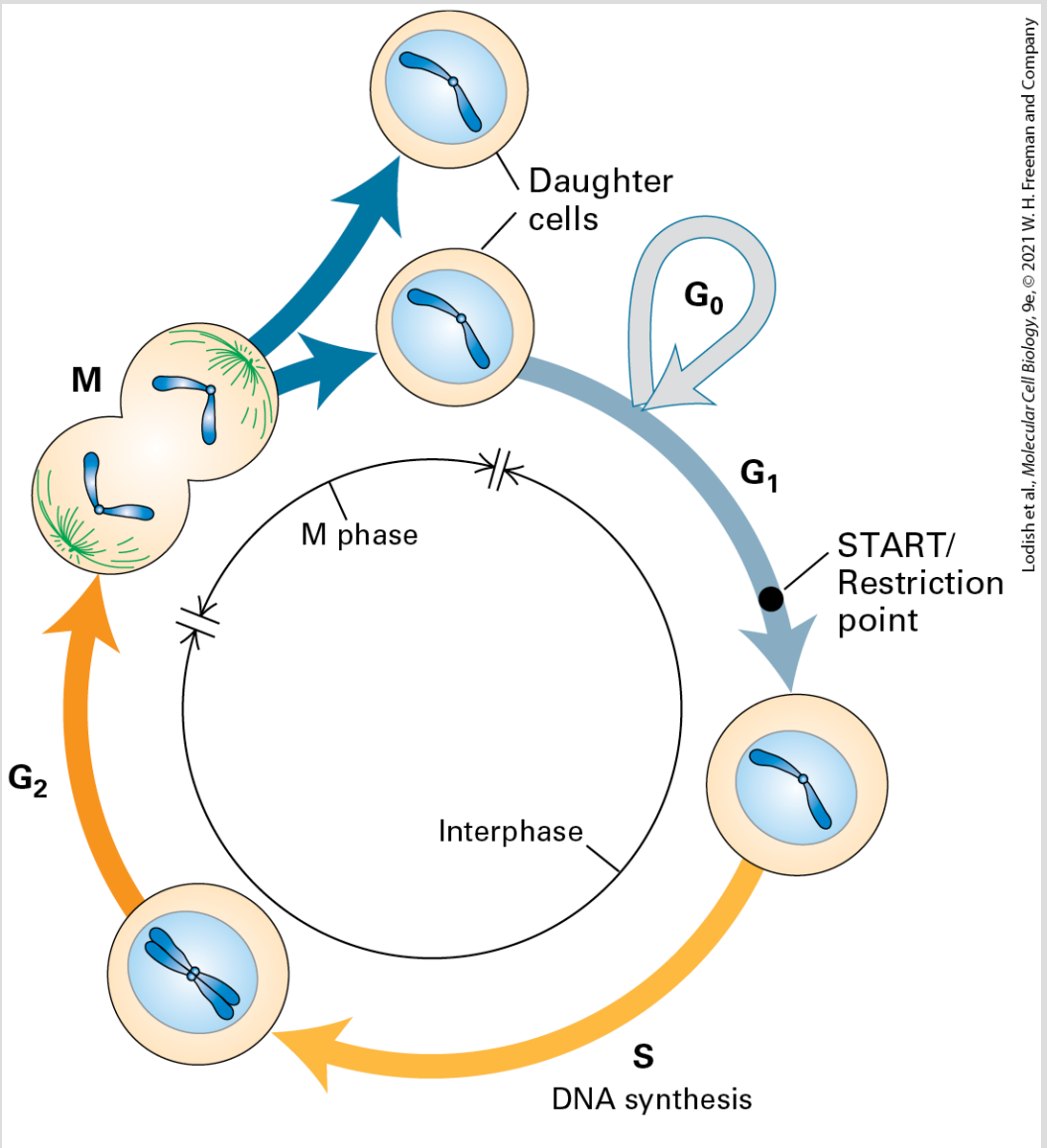


FIGURE 19-1 The eukaryotic cell cycle. The fate of a single chromosome during different stages of the cell cycle is depicted. Following mitosis (M), daughter cells contain 2*n* chromosomes in diploid organisms and 1*n* chromosomes in haploid organisms. In proliferating cells, G_1 is the period between the "birth" of a cell following mitosis and the initiation of DNA synthesis, which marks the beginning of the S phase. The point in late G_1 when cells become committed to enter S phase is called *START* or the *restriction point*. At the end of the S phase, cells enter G_2 containing twice the number of chromosomes they had as G_1 cells (4*n* in diploid organisms, 2*n* in haploid organisms). The end of G_2 is marked by the onset of mitosis, during which numerous events leading to cell division occur. The G_1, S, and G_2 phases are collectively referred to as

interphase: the period between one mitosis and the next. Most nonproliferating cells in vertebrates leave the cell cycle in G_1, entering the G_0 state. Although chromosomes condense only during mitosis, here they are shown in condensed form throughout the cell cycle to emphasize the number of chromosomes at each stage. The nuclear envelope surrounds the chromosomes during interphase in all cells but breaks down during mitosis in metazoan cells, but not in yeast.

Based on studying cell cycles in budding yeast, Lee Hartwell proposed that cells use specific surveillance mechanisms to ensure that each stage of the cell cycle is properly completed before cells enter the next stage. These surveillance programs establish *checkpoints* at the border between the different cell cycle phases. Cells must pass through the G_1-S checkpoint, for example before they can enter S phase and begin performing DNA synthesis, and they must pass through the G_2-M checkpoint before they can enter mitosis. The progression of cell cycle stages from G_1 to S to G_2 to M is the same for all eukaryotes, though the time it takes to complete one turn of the cycle varies considerably among organisms. Rapidly replicating human cells progress through the full cell cycle in about 24 hours: G_1 takes 9 hours; the S phase, 10 hours; G_2, 4.5 hours; and mitosis, 30 minutes. In contrast, the full cycle takes only 90 minutes in rapidly growing yeast cells. The cell divisions that take place during early embryonic development of the fruit fly *Drosophila melanogaster* are completed in as little as 8 minutes!

G_1 Controls Entry into S Phase

If the purpose of S phase is to duplicate the chromosomal DNA, and the purpose of M phase is to segregate the chromosomes to the daughter cells, what are G_1 and G_2 for? During G_1, a cell must evaluate its status and decide whether it is appropriate to commit to doubling and cell division. This assessment includes an evaluation of the cell's size, nutrient status, substrate attachment, density of neighboring cells, and the presence of extracellular growth factors and other chemicals that stimulate cell division. Growth factors and chemicals that promote cells to transit through the cell cycle and divide are generally referred to as *mitogens*. Specific signaling pathways monitor the state of the cell and its environment, and the outputs of these pathways are integrated to control the decision whether to replicate and divide. The TOR and Hippo signaling pathways, discussed in Chapter 21, monitor a cell's nutrient status, size, attachment status, and the density of surrounding cells. Growth factor signal transduction pathways, the phosphoinositol-3 kinase pathway, and the mitogen-activated protein kinase (MAPK) pathway discussed in Chapter 16, all report on whether mitogens are in the immediate vicinity of the cell.

Once cells commit to replicating their DNA, the decision is irreversible. That is, once cells enter S phase, they must complete the process of duplicating all of their chromosomes regardless of whether growth factors or nutrients are removed from the

extracellular environment. The START or restriction point where cells commit to entering S phase occurs late in G_1; in the case of mammalian cells this point is 2–3 hours prior to S phase onset. The observation that growth factors could be removed after mammalian cells had passed the restriction point (or that nutrients could be removed or mating pheromones that normally arrest yeast cells in G_1 could be added after yeast cells had passed START) and the cells would still progress into S phase led scientists to posit that some mysterious factor was accumulating during G_1, and that once enough of this labile "R-factor" had accumulated, the cells were destined to begin DNA replication. When S phase was nearly completed, this factor would be destroyed, and would have to be re-synthesized during G_1 of the next cell cycle. This cyclic accumulation and degradation of R-factor could explain why cells only synthesized DNA during S phase, and how cells controlled the timing of S phase entry by regulating the levels of R-factor during G_1, based on nutrients, mitogens, cell size, density, and substrate attachment. In mammalian cells, we now know that this cycling R-factor is a protein known as *cyclin D* (likely working in a feedback loop with another protein, cyclin E), and in yeast this factor turned out to be Cln3 working through a feedback loop involving Cln1, Cln2, and a transcriptional repressor called *Whi5*. The term *cyclin* or its abbreviated form, *Cln*, captures the central concept that this labile regulatory factor cycles in abundance as the cells go through the different stages of cell duplication. As we shall see, all of the cyclin proteins play critical roles in controlling the parts of the cell

cycle. Cyclin proteins function at the molecular level by binding to one or more specific protein kinases — the cyclin-dependent kinases (CDKs) mentioned in the introduction — in order to increase the kinase's activity, change the kinase substrate specificity, and localize the kinase to specific parts of the cell.

Once in S phase, complex replication machinery is assembled at specific sites on the chromosomal DNA, at the origins of replication and precise duplication of each chromosome ensues. The centrosome, the main microtubule organizing center in animal cells (see Chapter 18) also duplicates during S phase. Once DNA replication has been completed, cells enter G_2, the second gap phase.

G_2 Phase Readies the Cell for Mitosis and Cell Division

During the G_2 phase, the cell verifies that all of the DNA has been correctly duplicated and is sufficient for producing two cells, errors that were made during DNA copying are corrected, and any breaks in the DNA strands are repaired. In addition, the first signs of chromosome condensation occur, and there is some early reorganization of the cytoskeleton and microtubules as the cell prepares to enter mitosis.

How does the cell know during G_2 when everything is ready for mitosis? As with the transition from G_1 into S phase, another labile cyclin protein accumulates during G_2, which when present in sufficient amounts is able to drive the cells from G_2 into mitosis. In a classic experiment, Tim Hunt added radioactive methionine to fertilized sea urchin eggs and noticed the gradual accumulation of a radiolabeled protein that was precipitously destroyed in mitosis. Work done in a variety of different organisms — budding yeast, frog eggs, sea urchin eggs, sea star, and surf clam embryos — expanded on this observation, and eventually led scientists to clone and characterize this protein, which we now know as cyclin B. The binding of cyclin B to another cyclin-dependent kinase, called *CDK1*, forms an active protein kinase complex that drives the entry of G_2 cells into M phase. Surveillance mechanisms at the G_2-M checkpoint ensure that cells do not progress from G_2 into M phase in the presence of cell stress and damage, such as that caused by radiation, DNA-damaging agents, or drugs that interfere with microtubule polymerization.

Intriguingly, certain organisms, including budding yeast (but not fission yeast), lack a discernable G_2 phase altogether. In other model organisms that have been used to study the cell cycle, such as the frog *Xenopus laevis* and the fruit fly *Drosophila melanogaster*, both the G_1 and G_2 phases appear to be absent during several of the rapid early embryonic cell cycles that occur when the egg first begins to divide, prior to the onset of development. In these cases, the cells

actually divide without growth, making the cells progressively smaller in size. Nonetheless, later cell cycles in these organisms, and most cell cycles in adult organisms in general, require growth during G_1 and the establishment of a G_2 phase between S phase and M phase. Additional accumulation of cell mass including lipid and protein synthesis and mitochondrial division occurs in G_2, while Golgi complex duplication occurs in late G_1, S phase, and G_2.

Mitosis and Cytokinesis Occur During M-Phase

At the completion of S and G_2, the DNA of each chromosome has been duplicated through the process of DNA replication, generating two identical DNA molecules, each of which is also coated with histones and other chromosome-associated proteins (see Figure 8-35). Each of the identical chromosomes is now referred to as a **chromatid** (Figure 19-2). The two sister chromatids are initially held tightly together all along their length by a protein complex called *cohesin*. The chromatid attachment points on mitotic chromosomes are densest at a knobby constricted central region of each chromosome called the **centromere**, which contains many repeated copies of certain DNA sequences. The parts of the chromatid on either side of the centromere are referred to as *arms*. During mitosis, the cohesin links between the two sister chromatids are dissolved, and the sister chromatids are separated from each other

along a proteinaceous spindle made primarily from microtubules; the chromatids ultimately are distributed into each of the two daughter cells.

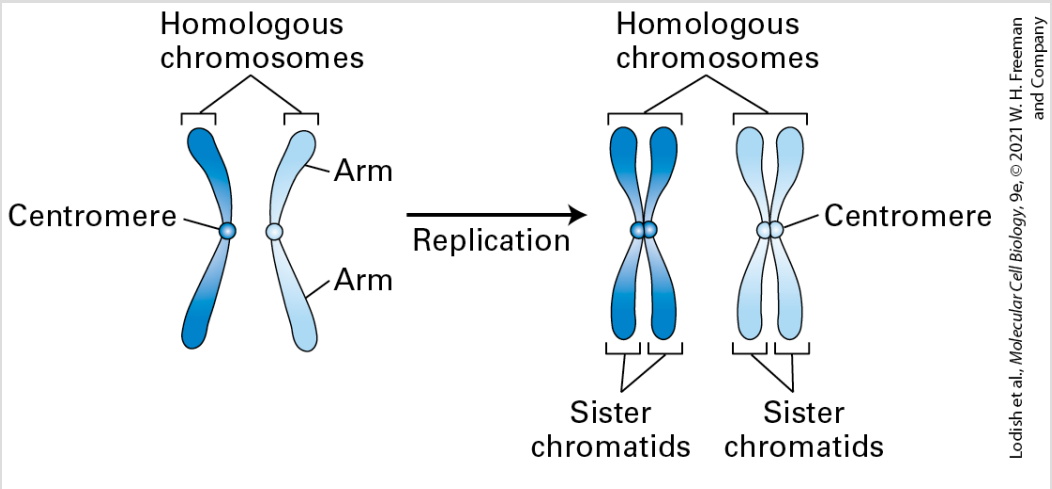


FIGURE 19-2 Duplication of chromosomes during S phase. Diploid cells contain two of each chromosome, referred to as *chromosome homologs*. At the end of S phase each homologous chromosome has two copies of itself, each copy of which is called a *sister chromatid*. The centromere and chromosome arms are indicated. For simplicity, the cohesin protein that holds the two sister chromatids together is not shown.

The process of separating the two copies of the chromosome is rather elaborate and occurs in a series of six stages that can be recognized when dividing cells are viewed under the microscope, as described in Chapter 18. These stages are summarized in Figure 19-3. To reiterate the main events that occur at each stage: during prophase the chromatin fibers in the nucleus condense into the discrete chromosomes that can be visualized under the light microscope. The cohesin protein that connects the arms of the

duplicated chromosomes to each other from S phase is degraded, leaving the two sister chromatids connected to each other only through cohesin bridges near the centromere. Nucleoli, the sites in the nucleus where ribosome synthesis occurs, disappear. The interphase array of microtubules breaks down, and the centrosomes (which are sometimes now called *mitotic asters* or *spindle poles*) move away from each other as a mitotic spindle begins to form between them. A complex of proteins called the *kinetochore*, which will anchor each chromosome to some of the spindle microtubules, assembles at the centromere of each chromosome.

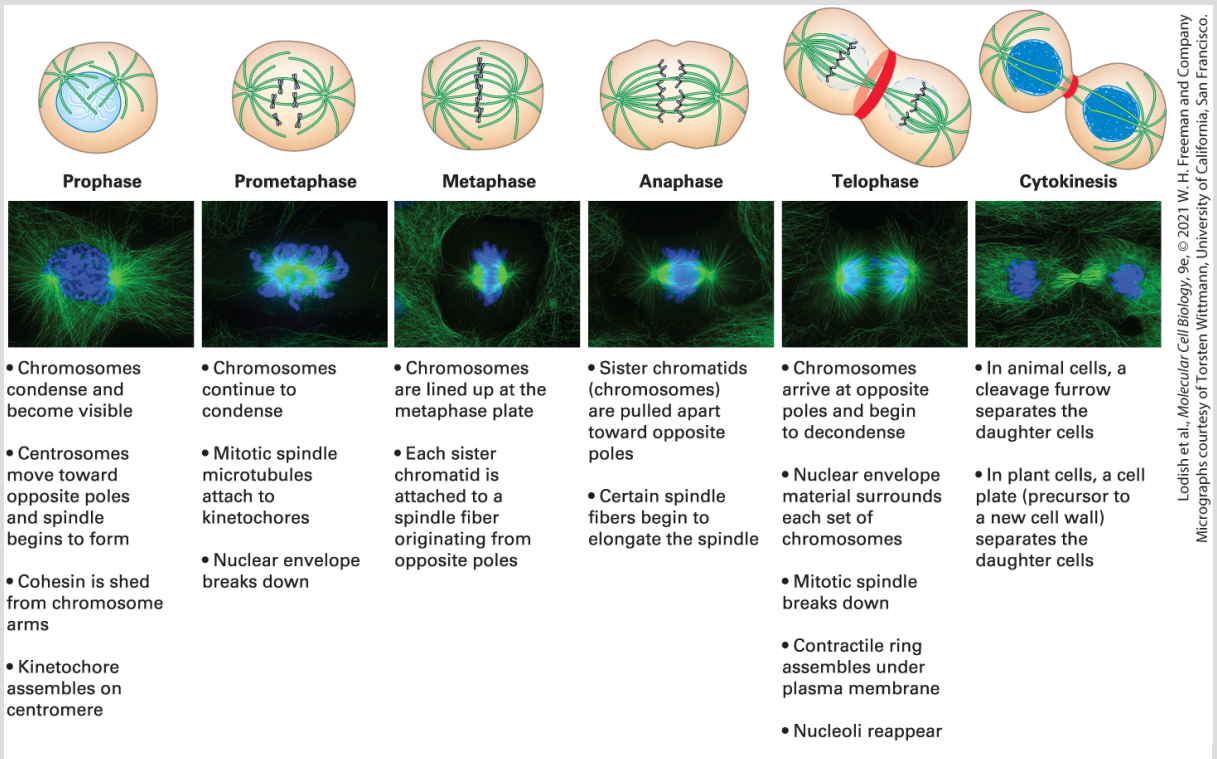


FIGURE 19-3 The stages of mitosis. The top panels show a diagrammatic representation of key events that occur in each stage of mitosis, the relatively short period of the cell cycle that follows the S and G_2 phases. The bottom panels are micrographs of dividing cells with

the DNA stained blue and the microtubules stained green. Below are listed the major events that occur in each stage.

Next, during prometaphase in higher eukaryotes, the nuclear membrane, the outer part of which was previously contiguous with the endoplasmic reticulum, begins to break down into vesicles and retract back into the endoplasmic reticulum, and the Golgi apparatus breaks down into vesicles. This allows the microtubules emanating from each centrosome to penetrate into the nuclear area and become captured by the kinetochores, becoming kinetochore microtubules. The chromosomes become even more condensed and now appear as the classic thick X-shaped mitotic chromosomes composed of paired sister chromatids (see Figure 1-17). Non-kinetochore microtubules reach across the cell to overlap with those arising from the other centrosome, while a second set of microtubules that face toward the cell membrane and away from the mitotic chromosomes becomes apparent. Following capture of the chromosomes by attachment of their kinetochores to microtubules, the chromosomes begin to move to and fro, with a net movement toward the center of the mitotic spindle.

When the centrosomes have moved to the opposite ends of the cell and the chromosomes have congressed to the middle of the spindle, the cell has entered metaphase. At this point the centromeres at the center of each chromosome have aligned along an imaginary plane midway between the spindle poles called the *metaphase plate*. The

kinetochores of the sister chromatids in each chromosome should now be attached to microtubules coming from opposite poles. Once this correct series of attachments has been confirmed, the cell enters anaphase. During anaphase A, the cohesin glue that held the sister chromatids together at the kinetochores is dissolved; the individual sister chromatids separate from one another and move toward opposite poles of the spindle as the kinetochore microtubules shorten at a rate of about 1 μm/min. During anaphase B, the non-kinetochore microtubules lengthen, causing the cell to elongate as the spindle poles separate further. Once the chromosomes have fully separated to opposite ends of the cell, the cell enters telophase. The chromosomes decondense, the nuclear envelope reforms, nucleoli reappear, any remaining spindle microtubules disassemble, and in animal cells, an actin/myosin–based contractile ring forms just under the plasma membrane in the region of the former spindle midzone. The next step in cell division is cytokinesis, when the cytoplasm becomes divided between the two daughter cells. Historically, cytokinesis was considered a separate process from mitosis, but in this book we follow the current convention to include cytokinesis as the terminal stage of mitosis. Cytokinesis begins in late anaphase or telophase. In animal cells, it involves the formation of a cleavage furrow by narrowing of the contractile ring and severing of any remaining microtubules in the region between where the two daughter cells will separate. The ring slowly closes until the plasma membrane pinches off, forming two discrete cells. The process is somewhat different in yeast, since they must also synthesize a new carbohydrate-based cell wall called

the *septum* at the same time that the actomyosin ring is contracting. In either case — animals or yeast, the end result of mitosis is the same — two daughter cells each containing a nucleus and cytoplasm with a share of the organelles, membranes, and other structures that were present in the mother cell.

All of the stages of mitosis are choreographed by the actions of mitotic protein kinases. As we shall see, cyclin B/CDK1 plays one of the most important roles in controlling the entry into mitosis and in orchestrating many of the molecular events that occur up until the onset of anaphase, when cyclin B is rapidly destroyed. A variety of other protein kinases, including Polo-like kinase, Aurora A and B kinases, and the Nek family of kinases, also play very important roles during the mitotic process. Once again, surveillance mechanisms ensure that each stage of mitosis has been completed properly before moving into the next stage. This is particularly evident at the transition from metaphase to anaphase, where the spindle assembly checkpoint monitors whether each of the chromosomes is properly positioned at the metaphase plate and whether the kinetochores of each sister chromatid are properly attached to microtubules arising from opposite spindle poles.

At the completion of mitosis, the daughter cells can continue cycling or can exit the cell cycle and proceed down a pathway of differentiation. In multicellular organisms, most differentiating cells exit the cell cycle and survive for days, weeks, or in some cases

(e.g., nerve cells and cells of the eye lens) even the lifetime of the organism without dividing again. Such *postmitotic* cells generally exit the cell cycle in G_1 and enter a phase called G_0 (see Figure 19-1). Some G_0 cells can return to the cell cycle and resume replicating; this re-entry is regulated, thereby providing control of cell proliferation.

KEY CONCEPTS OF SECTION 19.1

Overview of the Cell Cycle

- The eukaryotic cell cycle is divided into four phases: G_1 (the period between mitosis and the initiation of nuclear DNA replication), S (the period of nuclear DNA replication), G_2 (the period between the completion of nuclear DNA replication and mitosis), and M (mitosis and cytokinesis).
- Cells commit to a new cell division at a specific point in G_1 known as *START* or the *restriction point*. This commitment involves the integration of multiple signaling pathways that monitor cell size, nutrient status, and the conditions of the extracellular environment.
- Mitogens are growth factors and other molecules that induce cells to divide.
- Labile factors called *cyclins* drive the progression of cells through different stages of the cell cycle by binding to cyclin-dependent protein kinases (CDKs). The resulting cyclin-CDK complexes are present and active only in the cell cycle stage that they promote.
- Surveillance mechanisms, called *checkpoint pathways*, guarantee that each cell cycle step is completed correctly before the next one is initiated.

19.2 Model Organisms and Methods of Studying the Cell Cycle

Unraveling the molecular mechanisms governing cell cycle progression in eukaryotes was remarkably rapid and was fueled by a powerful combination of genetic and biochemical approaches. In this section, we discuss the contributions of several model systems to the discovery of the molecular mechanisms of cell division. Three of the most important systems employed to study the cell cycle are the single-celled yeasts *Saccharomyces cerevisiae* (budding yeast) and *Schizosaccharomyces pombe* (fission yeast) and the oocytes and early embryos of the frog *Xenopus laevis*. Studies using mammalian tissue culture cells led to the characterization of cell cycle control in mammals.

Studies of the cell cycle in many different experimental systems also led to two remarkable discoveries about general control of the cell cycle. First, complex molecular processes such as initiation of DNA replication and entry into mitosis are all regulated and coordinated by a small number of master cell cycle regulatory proteins. Second, these master regulators and the proteins that control them are highly conserved, so that cell cycle studies in fungi, sea urchins, insects, frogs, and other species are directly applicable to all eukaryotic cells, including human cells.

Budding and Fission Yeasts Are Powerful Systems for Genetic Analysis of the Cell Cycle

Budding and fission yeasts have proved to be extraordinarily valuable systems for the study of the cell cycle because of the ease with which they can be genetically manipulated. Both budding and fission yeast belong to the kingdom *Fungi*, but they diverged from each other over 350 million years ago and are therefore only distantly related. In fact, the evolutionary distance between these yeast is similar to the evolutionary distance between each of them and mammals. As a result, studies in both types of yeast reveal valuable insights that directly apply to metazoan animals. Both yeast organisms can exist in the haploid state, carrying only one copy of each chromosome and therefore single copies of each gene. Haploid yeast are an ideal genetic system in which to screen or select for mutants with defects in cell proliferation, because doing the same experiment in diploid organisms would require mutating both copies of a gene. In addition, it is relatively easy to manipulate the expression of individual yeast genes, and yeasts are easy to cultivate and manipulate so that cultures of cells progress through the cell cycle in a synchronous manner.

S. cerevisiae cells are ovoid in shape and divide by budding (Figure 19-4a). The bud, which is the future daughter cell, begins to form concomitant with the initiation of DNA replication and continues to

grow throughout the cell cycle (Figure 19-4b). Cell cycle stage can therefore be inferred from the size of the bud, which makes *S. cerevisiae* a useful system for identifying mutants that are blocked at specific steps in the cell cycle. As with mammalian cells, the budding yeast cell cycle has a long G_1 phase, and the study of the budding yeast cell cycle shaped our understanding of how the G_1–S phase transition is controlled.

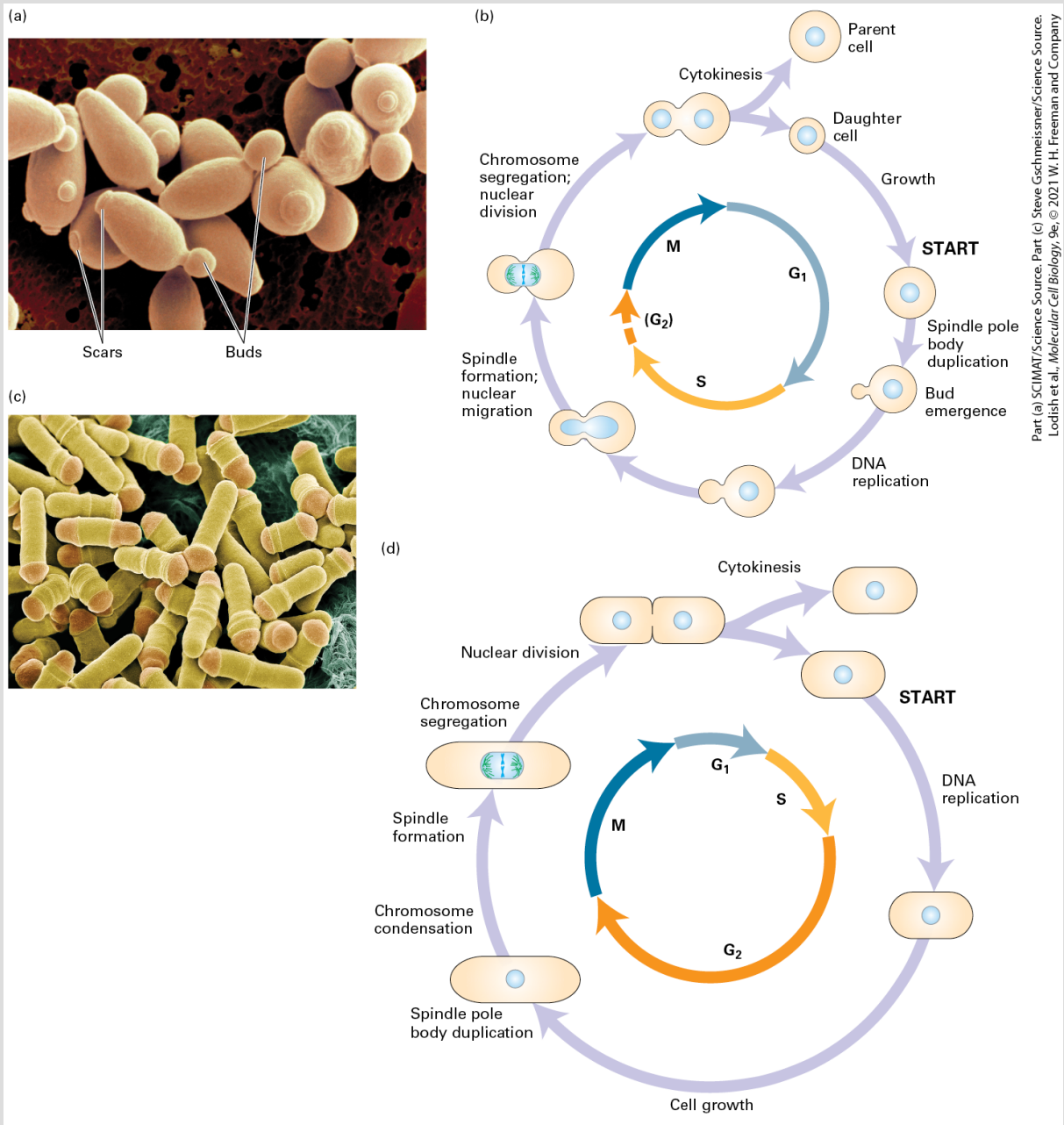


FIGURE 19-4 The cell cycle of budding yeast *S. cerevisiae* and fission yeast *S. pombe*. (a) Scanning electron micrograph of *S. cerevisiae* cells at various stages of the cell cycle. The larger the bud that emerges at the end of the G_1 phase, the farther along in the cycle the cell is. Bud scars are also evident at sites where previous daughter cells budded off. (b) Main events in the *S. cerevisiae* cell cycle. The daughter cells must grow to a certain size before they enter START. G_2 is not well defined in budding yeast and is therefore indicated by a dashed arrow. In contrast to higher eukaryotes, the nuclear envelope does not disassemble during mitosis in *S. cerevisiae* and other yeasts, and they undergo a “closed mitosis.”

Complicating matters further, the small *S. cerevisiae* chromosomes do not condense sufficiently to be visible by light microscopy. (c) Scanning electron micrograph of *S. pombe* cells at various stages of the cell cycle. Long cells are about to enter mitosis; short cells have just passed through cytokinesis. (d) Main events in the *S. pombe* cell cycle. This yeast has relatively short G_1 and S phases of the cell cycle and a prominent G_2 stage that is similar to that in metazoan cells. As in *S. cerevisiae*, the nuclear envelope does not break down during mitosis.

Fission yeast cells are rod-shaped, divide in the middle, and grow entirely by elongation at their ends (Figure 19-4c). Consequently, the cell cycle stage of this organism can be inferred from a simple measurement of its length. Cytokinesis occurs by the formation of a septum in the center of the cell, splitting it in two, hence the name *fission yeast* (Figure 19-4d). In contrast to budding yeast and mammalian cells, fission yeast have very short G_1 and S phases of the cell cycle, and instead spend most of the cell cycle in G_2. The molecular mechanisms governing G_2 and entry into mitosis in fission yeast and in metazoan cells are very similar; studies with this organism by Paul Nurse and colleagues revealed the molecular events surrounding the G_2/M phase transition.

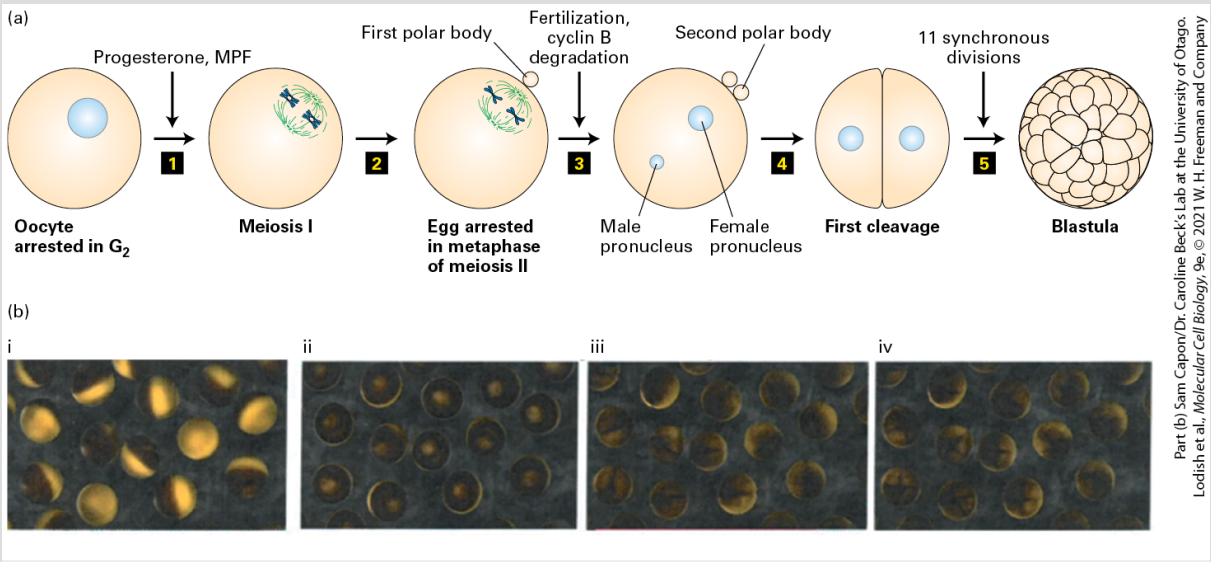
Studying yeast mutants that are blocked at specific steps in the cell cycle or that exhibit altered regulation of the cycle has been very useful. Because cell cycle progression is essential for viability, scientists isolated **temperature-sensitive mutants** whose genes encode proteins that are functional at one temperature but become inactive at a different, often higher, temperature (e.g., due to protein

misfolding at the nonpermissive temperature; see Figure 6-6). These mutants, when arrested at a particular cell cycle stage, are easily distinguished from normally dividing cells by microscopic examination. Such cells are called *cdc* (*c*ell *d*ivision *c*ycle) mutants. Identification of the genes mutated in these temperature-sensitive yeast strains provided a comprehensive list of genes critical for virtually all aspects of cell division.

Frog Oocytes and Early Embryos Facilitate Biochemical Characterization of the Cell Cycle Machinery

Although yeast are ideal for large-scale genetic screens, their small size and thick cell wall make them much less suitable for detailed biochemical analysis. To prepare cell extracts for biochemical studies of the cell cycle, the eggs and early embryos of amphibians and marine invertebrates have proven to be particularly suitable. These organisms typically have large eggs, and fertilization is followed by multiple synchronous cell cycles. By isolating large numbers of eggs from females and fertilizing them simultaneously by addition of sperm (or by treating them in ways that mimic fertilization), researchers can obtain extracts from cells at specific points in the cell cycle for analysis of proteins and enzymatic activities.

To understand how *X. laevis* oocytes and eggs can be used for the analysis of cell cycle progression, let's first look at oocyte maturation, which can be recapitulated in vitro. Oocytes undergo a meiotic division (see Figure 19-37 for an overview of meiosis). As oocytes develop in the frog ovary, they replicate their DNA and become arrested in G_2 for 8 months, during which time they grow in size to a diameter of 1 mm, stockpiling all the materials needed for the multiple cell divisions of the early embryo. When stimulated by a male, an adult female's ovarian cells secrete the steroid hormone progesterone, which induces the G_2-arrested oocytes to mature and enter meiosis. As we will see in Section 19.8, meiosis consists of two consecutive chromosome segregation phases known as *meiosis I* and *meiosis II*. Progesterone triggers oocytes to undergo meiosis I and progress to the second meiotic metaphase, where they arrest and await fertilization (Figure 19-5a). At this stage the cells are called *eggs*. When fertilized by sperm, the egg nucleus is released from its metaphase II arrest and completes meiosis. The resulting haploid egg nucleus then fuses with the haploid sperm nucleus, producing a diploid *zygote* nucleus. DNA replication follows, and the first mitotic division of embryogenesis begins (Figure 19-5b). The resulting embryonic cells proceed through 11 more rapid, synchronous cell cycles, generating a hollow sphere of cells called the *blastula*. Cell division then slows, and subsequent divisions are asynchronous, with cells at different positions in the blastula dividing at different times.



Part (b) Sam Capon/Dr. Caroline Beck's Lab at the University of Otago.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-5 Progesterone stimulates maturation of *Xenopus* oocytes. (a) Step **1**: Progesterone treatment of G_2-arrested *Xenopus* oocytes causes the oocytes to enter meiosis I. The same activity is seen if the oocytes are injected with MPF. Two pairs of synapsed homologous chromosomes (blue) connected to meiotic spindle microtubules (green) are shown schematically to represent cells in metaphase of meiosis I. Step **2**: Segregation of homologous chromosomes and a highly asymmetric cell division expels half the chromosomes into a small cell called the *first polar body*. The oocyte immediately commences meiosis II and arrests in metaphase II to yield an egg. Two chromosomes connected to spindle microtubules are shown schematically to represent egg cells arrested in metaphase of meiosis II. Step **3**: Fertilization by sperm releases eggs from their metaphase arrest, allowing them to proceed through anaphase of meiosis II and undergo a second highly asymmetric cell division that expels one chromatid of each chromosome into a second polar body. The resulting haploid female pronucleus fuses with the haploid sperm pronucleus to produce a diploid zygote. Step **4**: The zygote undergoes DNA replication and the first mitosis. Step **5**: The first mitosis is followed by 11 more synchronous divisions to form a blastula. (b) Micrographs of *Xenopus* eggs transitioning through the cell cycle. (i) G_2-arrested oocytes. (ii) Following progesterone or MPF treatment the oocytes transit from G_2 into meiosis. (iii) Upon fertilization, there is a rise in intracellular Ca^{2+} that triggers cyclin B destruction, leading to completion of meiosis and the first cell division. (iv) The first of the next 11 rapid synchronous cycles of cell division is shown.

The advantage of using *X. laevis* to study factors involved in mitosis is that large numbers of oocytes and eggs can be obtained, and following progesterone treatment and fertilization in the laboratory they will all proceed synchronously through the cell cycle in a manner that can be studied experimentally. Furthermore, the harvested eggs can be crushed by centrifugation in order to prepare reasonably large amounts of an extract that can be induced to undergo mitosis by the addition of sperm nuclei. This makes it possible to carry out biochemical experiments from cells that were all at the same point in the cell cycle. It was in this system that the cyclin-CDK complexes that trigger mitosis and the oscillatory nature of their activity were first discovered. This activity was called *maturation-promoting factor (MPF)* because of its ability to induce entry into meiosis and oocyte maturation when injected into G_2-arrested oocytes and later called *mitosis-promoting factor* when it was shown to induce mitosis in all eukaryotic cells.

The Study of Tissue Culture Cells Uncovers Cell Cycle Regulation in Mammals

Cell cycle regulation in human cells is more complex than in non-mammalian systems. To understand this increased level of complexity, and to understand the cell cycle alterations that are the cause of cancer, it is important to study the cell cycle not only in

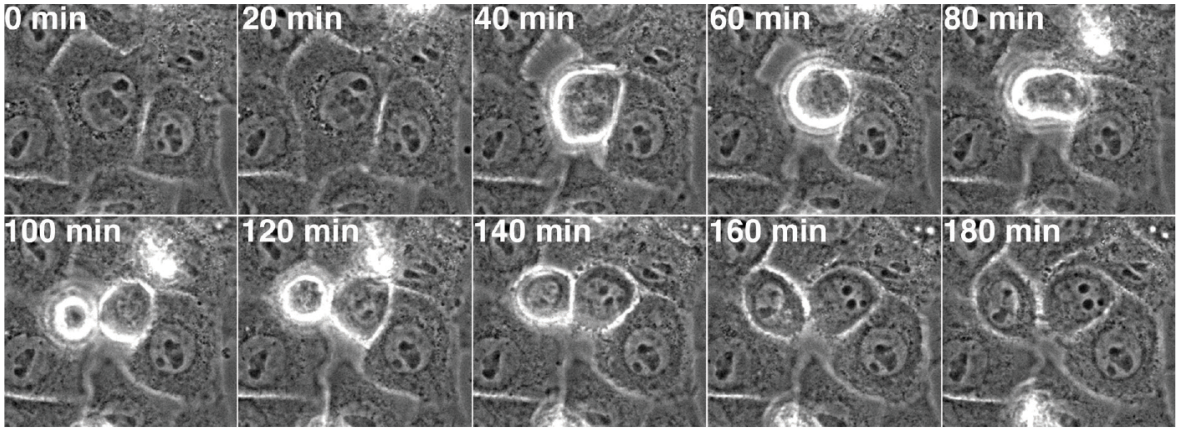
model organisms but also in human cells. To study the properties of the human cell cycle, researchers use normal or tumor cells grown in plastic dishes, a method called *tissue culture* or *cell culture* (see Chapter 4). However, note that many of the cell types used to study the human cell cycle have altered cell cycle properties due to genetic alterations that occurred during their culturing or because they were isolated from human tumors. Furthermore, in vitro culture conditions do not resemble those found in the organism and could lead to altered behavior of cells. Although some aspects of mammalian cell division are not recapitulated in cell culture conditions — such as tissue organization and developmental signals governing cell cycle control — cell culture systems nevertheless provide critical insights into the mammalian cell's intrinsic mechanisms governing cell division. Researchers are also working toward establishing culture systems that more closely resemble the cell architecture in tissues. For example, mixtures of different cell types can be grown together to form tissue-like architectures, while embryonic or adult stem cells, or patient-derived tumor cells can be grown in the absence of serum on purified extracellular matrix from tissues, resulting in the formation of self-organizing structures called *organoids* that more closely resemble real organs (see Figure 22-17). More chemically defined polymer lattices are currently being developed that also allow scientists to grow cells in three-dimensional culture mimicking cellular organization within a tissue.

Primary human cells and other mammalian cells have a finite life span when cultured in vitro. Normal human cells, for example, divide 25–50 times, but thereafter proliferation slows and eventually stops. This process is called *replicative senescence*. Some cells can escape this process and become immortalized, allowing researchers to establish cell lines. Although these cell lines harbor genetic alterations that affect some aspects of their proliferation, they are nevertheless a useful tool for studying cell cycle progression in human cells. These cell lines provide an inexhaustible supply of cells that, as we will see next, can be manipulated to progress through the cell cycle in a synchronous manner, allowing for the analysis of protein levels and enzymatic activity at different stages of the cell cycle.

Researchers Use Multiple Tools to Study the Cell Cycle

The experimental analysis of cell cycle properties requires that we be able to determine the cell cycle stage of individual cells. Light microscopy provides some estimate of cell cycle progression. For example, light microscopy allows a researcher to determine whether cultured mammalian cells are in interphase (G_1, S phase, and G_2) or in mitosis. Mammalian tissue culture cells are flat and adhere to their growth surface during interphase, but round up and form spherical structures as they undergo mitosis (Figure 19-6).

Fluorescence microscopy of cellular structures or analysis of specific cell cycle markers — that is, proteins that are present only in certain cell cycle stages — allows for a more accurate determination of cell cycle stage.

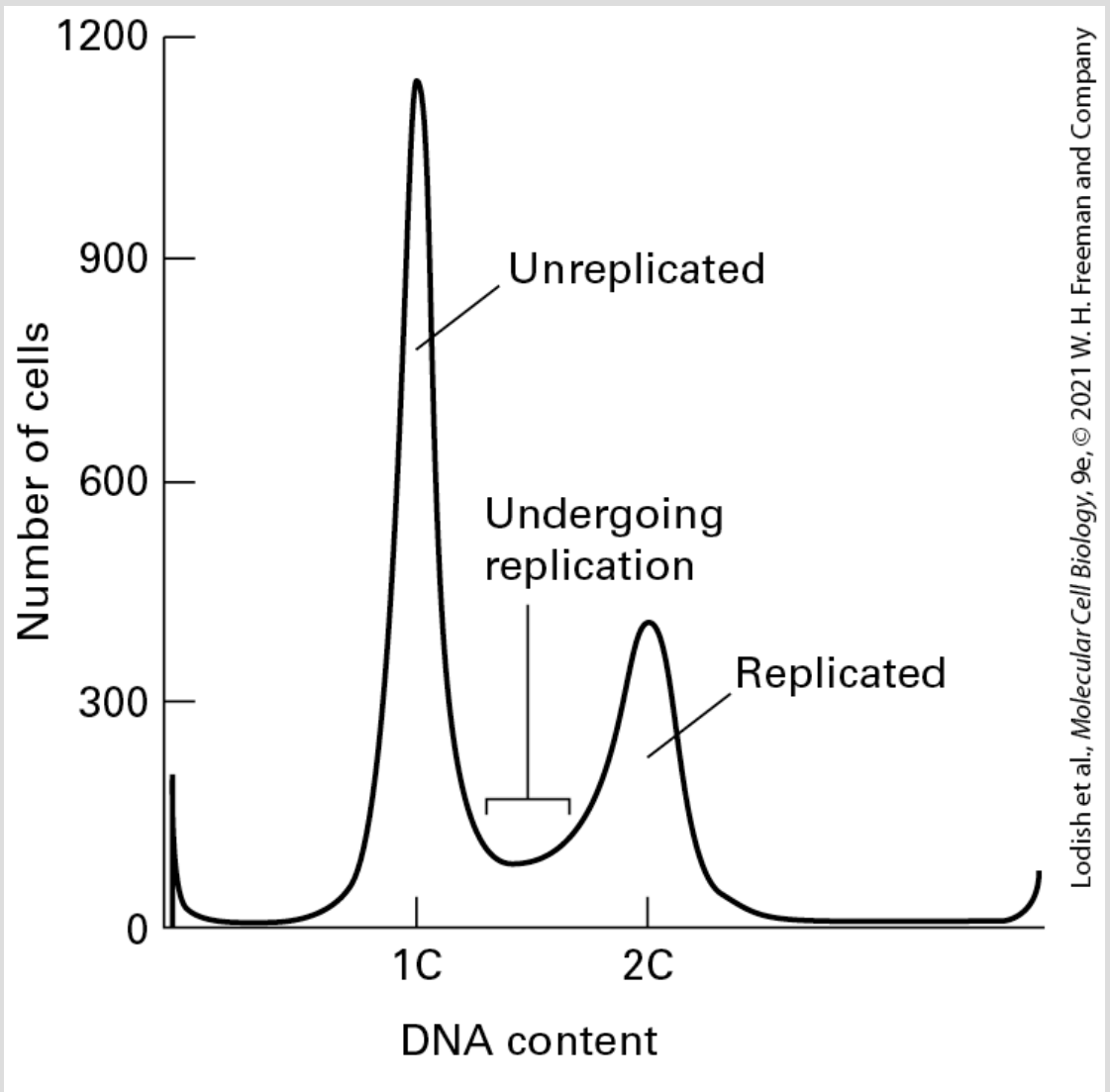


Courtesy of Sejal Vyas and Paul Chang, MIT.

FIGURE 19-6 Human cells undergoing mitosis. HeLa Kyoto cells seen under phase contrast microscopy were filmed as they underwent mitosis. The images shown were taken every 20 minutes. Cells are flat during interphase. During mitosis, the cells round up and divide, and acquire a bright and refractile appearance. Subsequently, they flatten out again.

In addition to microscopic tools, cell cycle researchers use flow cytometry to determine the DNA content of a cell population (Figure 19-7; see also Figure 4-1). Cells are treated with a DNA-binding fluorescent dye, and the amount of dye that is incorporated into the DNA of cells is quantitatively assessed using a flow cytometer. Cells are recorded by their DNA content, and the percentages of cells in G_1, S phase, and G_2 or mitosis can be assessed in this manner. Cells in G_1 have half as much DNA as cells in G_2 or mitosis. Cells

undergoing DNA synthesis in S phase have an intermediate amount of DNA.



EXPERIMENTAL FIGURE 19-7 Analysis of DNA content by flow cytometry. Haploid yeast cells were grown in culture and stained with propidium iodide, a fluorescent dye that is incorporated into DNA. The *x* axis shows DNA content, the *y* axis the number of cells. The DNA content analysis shows two predominant populations of cells: cells with unreplicated DNA (1C) and with replicated DNA (2C). The cells between the two peaks represent cells that are in the process of undergoing DNA replication.

To characterize different cell cycle events, it is essential to examine cell populations that progress through the cell cycle in unison. Researchers can generate such populations by *reversibly arresting* cells in a particular cell cycle stage. This cell cycle arrest is usually accomplished by restricting nutrients or growth factors, causing cells to arrest in G_1. In budding yeast, for example, cells treated with a mating pheromone arrest in G_1. When the pheromone is removed from the cells (usually by washing them extensively), the cells exit G_1 and progress through the cell cycle in a synchronous manner. In mammalian cells, removal of growth factors by removing serum from the culture medium (serum starvation) arrests cells in G_0. Re-addition of serum allows cells to re-enter the cell cycle.

Other methods of arresting cells involve blocking a certain cell cycle step with chemicals. Hydroxyurea inhibits DNA replication, leading to arrest in S phase. After removing the drug, cells resume DNA synthesis in unison. Nocodazole disrupts the mitotic spindle and halts the cell cycle in mitosis. Once the drug is washed away, cells resume progression through mitosis in a synchronous manner. In budding and fission yeasts, when temperature-sensitive cdc mutants are incubated at the nonpermissive temperature, they arrest in a particular cell cycle stage because they are defective in a certain key cell cycle protein. Returning cells to the permissive temperature allows them to continue with the cell division cycle in a synchronous fashion.

KEY CONCEPTS OF SECTION 19.2

Model Organisms and Methods of Studying the Cell Cycle

- The ability to isolate mutants of budding and fission yeasts and the use of powerful genetic tools led to the isolation of key genetic factors important for cell cycle regulation.
- Frog eggs and early embryos from synchronously fertilized eggs provide sources of extracts for biochemical studies of cell cycle events, which led to the identification of the oscillatory nature of cyclin-CDK complexes.
- Human tissue culture cells are used to study the properties of the mammalian cell cycle.
- The generation of synchronized cell populations by reversibly arresting cells in a particular cell cycle stage allows researchers to examine the behavior of proteins and cellular processes during the cell cycle.

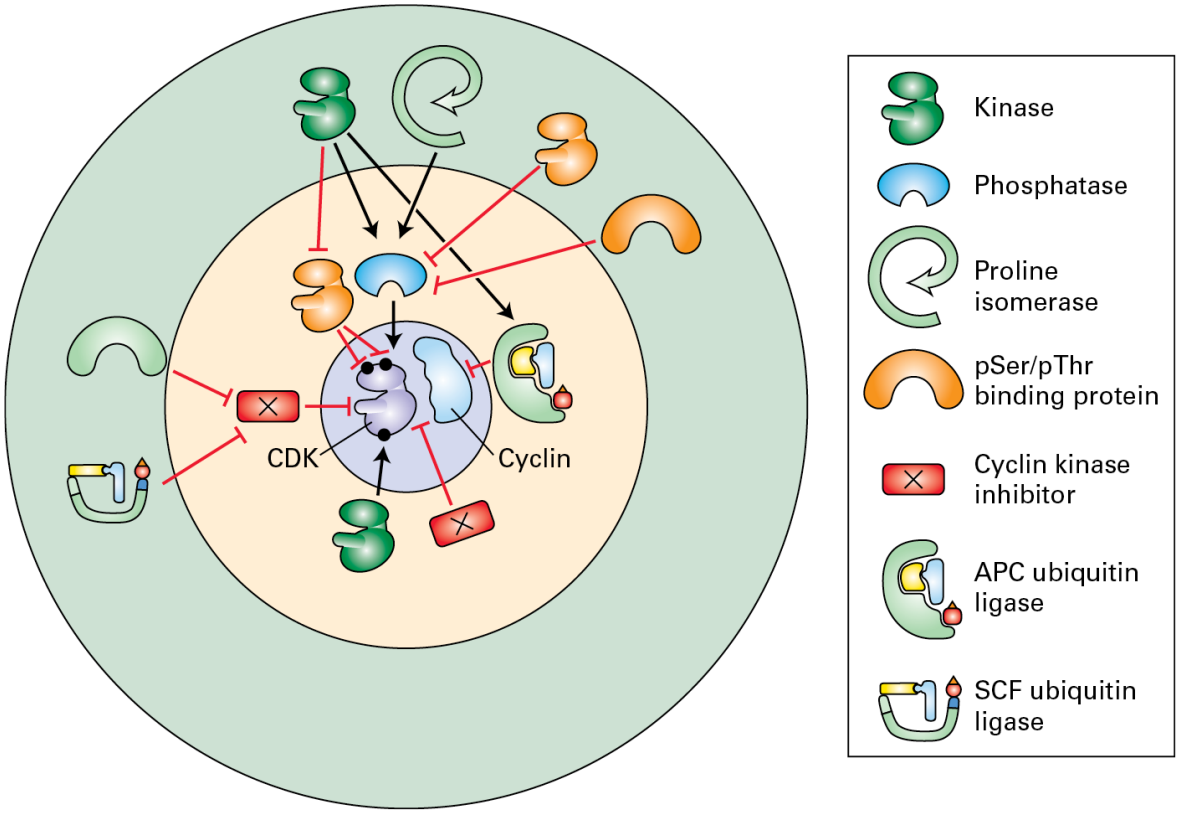
19.3 Cell Cycle Progression and Control: Feedback Loops and Post-Translational Modification

Progression through the cell cycle is driven, in large part, by the action of cyclin-dependent kinases (CDKs). These protein kinases act as master controllers of the cell cycle by phosphorylating substrates that when properly phosphorylated directly cause the events in each cell cycle stage to take place. In certain phases of the cell cycle, other kinases also play important roles, but the importance of CDKs in each phase of the cell cycle is paramount.

We know from many experimental observations that cells move irreversibly through the cell cycle, and they do so in only one direction (see Figure 19-1): from G_1 to S phase to G_2 to M phase. Cells have never been observed to go backward from G_2 into S phase, nor has anyone ever seen an S phase cell somehow unreplicate its DNA and revert back to a G_1-phase cell. This commitment once cells initiate each phase of the cell cycle means that the activity of these master controllers, the CDKs, must be regulated with extraordinary precision — at just the right time and place and in a manner that is essentially irreversible. Although details of how this is accomplished vary for different cell cycle stages, we will see in subsequent sections that the general molecular basis for CDK regulation is the same no matter which cell cycle stage

we are talking about. It always involves the use of cyclins — the accessory proteins that bind to the CDKs to control their activity — in combination with protein phosphorylation and ubiquitination (see Section 3.4). The control of CDKs can be thought of as occurring in several layers, represented by the concentric circles in Figure 19-8. At the center are the CDKs themselves, along with their activating accessory proteins, the cyclins. In the second layer are the direct regulators of the cyclin-CDK complexes, which includes protein kinases and phosphatases that activate or inhibit the CDKs, together with CDK inhibitor proteins that bind directly to the CDKs to block their activity, transcription factors that increase the rate of cyclin synthesis to create active cyclin-CDK complexes at specific points in the cell cycle and protein kinases and ubiquitin ligases that work together to promote the destruction of the cyclins once cells have completed a specific part of the cell cycle. In the outer layer are the indirect regulators of the cyclin-CDK complexes that function by regulating the direct regulators (Figure 19-8). Like the second layer of control, this third layer of indirect regulators is also composed of protein kinases and ubiquitin ligases, which work together with phosphoserine/threonine-binding domains and proline isomerases to control the levels, activity, and subcellular location of the direct regulators of cyclin-CDK activity. This multilayered network of control allows a cell to evaluate the massive amount of information about the state a cell finds itself in. A cell then integrates the information into all-or-none step-function-like decisions that control the onset of DNA replication or the initiation of mitosis, for example. Note that this information is reflected by gradual linear

changes in the activity of various signaling pathways. The combination of direct and indirect control afforded by this multilayered regulatory network allows a very small set of basic biochemical reactions — protein phosphorylation and de-phosphorylation, molecular recognition, and ubiquitin-mediated protein destruction — to build feed-forward and feedback loops that provide fine control of whether a cell cycle transition takes place. In addition, the outermost circle of indirect regulation provides the molecular basis for cell cycle checkpoint control, as explored in Section 19.7.



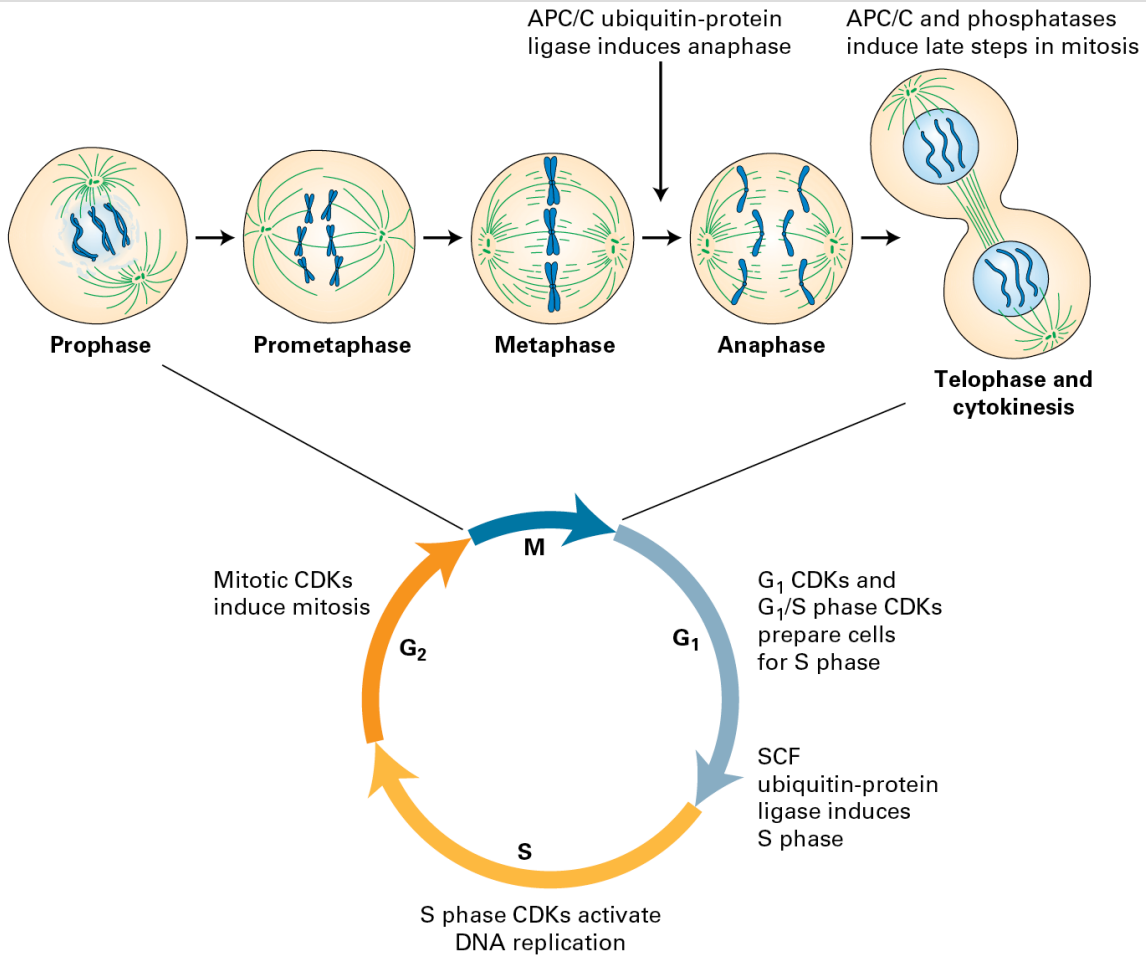
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-8 A multilayered network controls the activity of cyclin-dependent kinases. Concentric circles indicate the CDKs and their activating cyclins in the center, surrounded by their direct regulators, and then their indirect regulators. As shown in the key, the shapes

represent kinases, phosphatases, proline isomerases, phosphoserine/threonine-binding proteins and domains, cyclin kinase inhibitors, and ubiquitin ligases. Black lines generally denote activation of the indicated protein's target, while red lines denote inhibition. Feedback loops involving phosphorylation of the regulators by the cyclin-CDKs are not shown to maintain clarity. Note that there is substantial cross talk among the molecules and that the number of molecules and interactions are more complex than is shown here.

In the following sections, we describe the current model of eukaryotic cell cycle regulation, which is summarized in Figure 19-9. A key discovery in cell cycle studies was that cyclin-dependent kinases govern progression through the cell cycle. Three key features of these kinases are important to keep in mind throughout this chapter.

- Cyclin-dependent kinases (CDKs) are active only when bound to a regulatory cyclin subunit.
- Different types of cyclin-CDK complexes initiate different events. **G_1 cyclin-CDKs** and **G_1/S phase cyclin-CDKs** promote entry into the cell cycle, **S phase cyclin-CDKs** trigger S phase, and **mitotic cyclin-CDKs** initiate the events of mitosis (Figure 19-9).
- Multiple biochemical mechanisms ensure that the CDKs are active only in the stages of the cell cycle they trigger. These molecular mechanisms are organized into feedback loops that amplify the process of CDK activation and inhibition.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

	G_1 cyclin-CDKs	G_1/S cyclin-CDKs	S phase cyclin-CDKs	M phase cyclin-CDKs
Budding yeast	Cln3-Cdc28	Cln1/2-Cdc28	Clb5/6-Cdc28	Clb1/2/3/4-Cdc28
Humans	Cyclin D-CDK4 Cyclin D-CDK6	Cyclin E-CDK2	Cyclin A-CDK2	Cyclin A-CDK1 Cyclin B-CDK1

FIGURE 19-9 An overview of how CDKs regulate cell cycle progression. Cells harbor different types of CDKs that initiate different events of the cell cycle. Importantly, the CDKs are active only in the stages of the cell cycle that they trigger. G_1/S phase CDKs are active at the G_1–S phase transition to trigger entry into the cell cycle. S phase CDKs are active during S phase and trigger DNA replication. Mitotic CDKs are active during mitosis and trigger mitosis. The anaphase-promoting complex or cyclosome (APC/C) ubiquitin-protein ligase catalyzes two key cell cycle transitions by ubiquitinylating proteins, hence targeting them for degradation. APC/C initiates both anaphase and exit from mitosis. The names of the different CDKs and their associated cyclin subunits for yeast and humans are listed.

In this section, we discuss the properties of CDKs and investigate the structural basis of their activation and regulation. We then describe how cyclins activate CDKs and investigate the general regulatory mechanisms that restrict different cyclins to the appropriate cell cycle stage. We will see that protein degradation plays an essential part in this process, as do post-translational modifications to CDKs and inhibitory proteins that directly bind to cyclin-CDK complexes.

Cyclin-Dependent Kinases Are Small Protein Kinases That Require a Regulatory Cyclin Subunit for Their Activity

Cyclin-dependent kinases constitute a family of small (30–40 kDa) serine/threonine kinases that contain little more than the kinase domain. By themselves, monomeric CDKs are not active, and in general, their protein levels in the cell remain relatively constant throughout the cell cycle. The activity of these CDKs and the amounts of the substrates that they phosphorylate, however, change dramatically in the different cell cycle stages. These changes occur because the CDKs require the binding of an activating subunit — a cyclin — to become active protein kinases.

In budding and fission yeasts, a single CDK controls progression through the cell cycle. Its activity, and the specific substrates that it

phosphorylates, is specified by which of several cell-cycle-stage-specific cyclin subunits it is bound to. Mammalian cells, in contrast, contain as many as 20 different CDKs, with 4 of them, CDK1, CDK2, CDK4, and CDK6, having clear roles in regulating cell cycle progression. Each CDK binds to one or more cyclin subunits that control CDK activity and substrate preference. The different mammalian CDKs, together with their specific cyclins, promote different cell cycle transitions. Mammalian CDK4 and CDK6, for example are G_1 CDKs that promote entry into the cell cycle, CDK2 functions as a G_1/S phase and S phase CDK, and CDK1 is the mitotic CDK. For historical reasons, the names of various cyclin-dependent kinases from yeasts and vertebrates differ. Whenever possible, we will use the general terms G_1, G_1/S phase, S phase, and mitotic CDKs to describe CDKs instead of the species-specific terminology. The bottom of Figure 19-9 lists the names of the various mammalian and budding yeast cyclin-CDKs and indicates when in the cell cycle they are active.

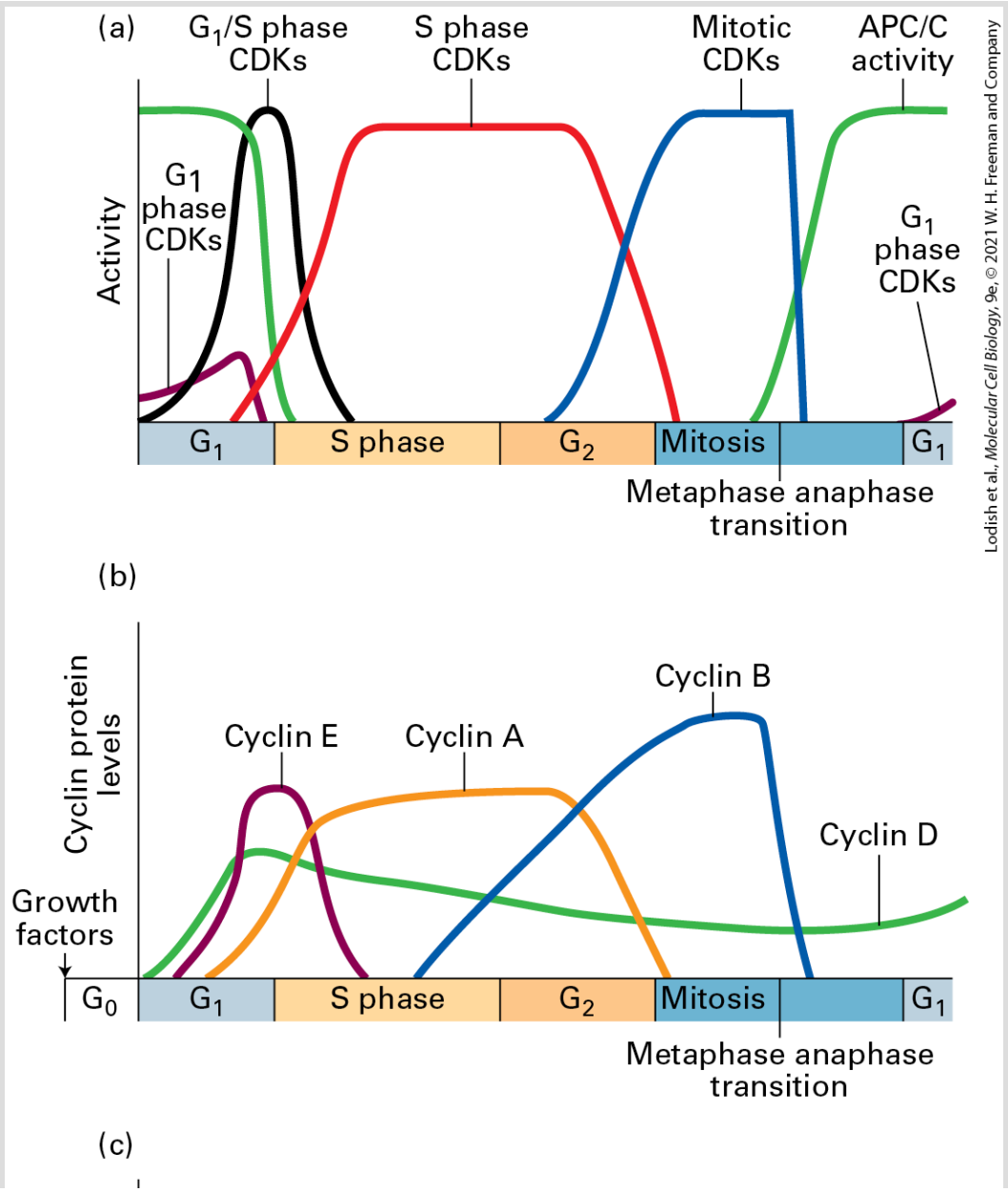
Cyclins Determine the Activity of CDKs

Cyclins are so named because their concentrations change during the cell cycle. They form a family of proteins that is defined by three features.

- Cyclins bind to and activate CDKs. The activity and substrate specificity of any given CDK is primarily defined by the

particular bound cyclin.

- Cyclins are present only during the cell cycle stage that they trigger and are absent from other cell cycle stages.
- In addition to triggering a particular cell cycle stage, cyclins also set in motion a series of events in preparation for the next cell cycle stage. In this way, they propel the cell cycle forward (Figure 19-10a).



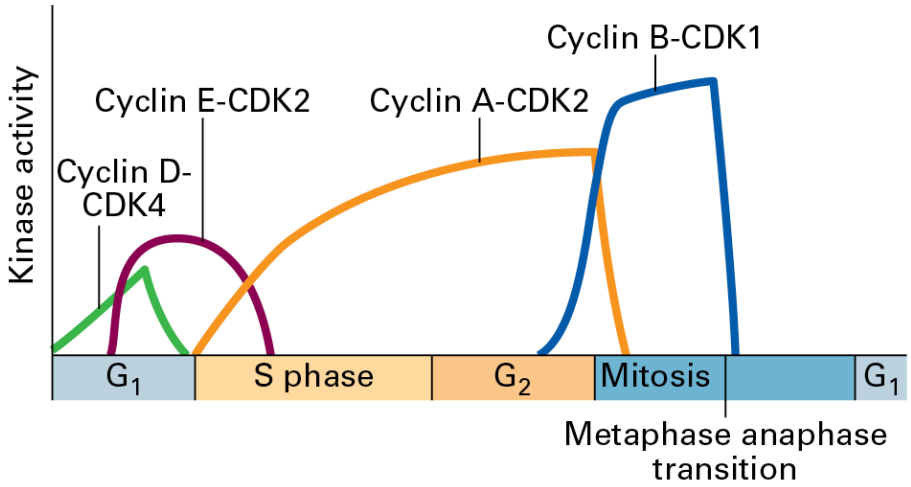


FIGURE 19-10 CDK activity during the cell cycle is regulated by binding to different cyclins. (a) The activity of different CDKs varies as cells transit through the stages of the cell cycle, as indicated in Figure 19-9. (b) The protein levels of different cyclin subunits that form the cell-cycle-stage-specific CDK complexes rise and fall in a cell-cycle-stage-specific manner. (c) The activity of the different cyclin-CDK complexes that are formed by the different cyclin-CDK complexes varies in a manner that is very similar, but not exactly identical to the accumulation of the cyclins. The difference between cyclin abundance in panel (b) and kinase activity shown in panel (c) indicates that there must be additional levels of regulation of cyclin-CDK activity, such as phosphorylation and de-phosphorylation in addition to cyclin binding.

Similar to the nomenclature used for CDKs, cyclins are divided into four classes defined by their presence and activity during specific phases of the cell cycle: G_1 cyclins, G_1/S phase cyclins, S phase cyclins, and mitotic cyclins (see Figure 19-9, *bottom*). Some cyclins can bind to and activate more than one CDK, while others can only bind to a single CDK. In mammalian cells, for example, cyclin A can form an activating complex with both CDK2 in S phase, and with

CDK1 at the very beginning of mitosis. In contrast, cyclin B, the main mitotic cyclin, can only form a complex with CDK1.

The G_1 cyclins are the linchpin in coordinating the cell cycle with extracellular events. Their activity is subject to regulation by signal transduction pathways that sense the presence of growth factors or cell proliferation inhibitory signals. In metazoans, G_1 cyclins are known as *cyclin Ds*, and they bind to CDK4 and CDK6. Unlike other cyclins, G_1 cyclins levels do not show strong, stage-dependent fluctuations. Instead, in response to macromolecule biosynthesis and extracellular signals, their levels gradually increase in G_1, drop a bit during S and G_2 phase, and rise again during late M phase (Figure 19-10b).

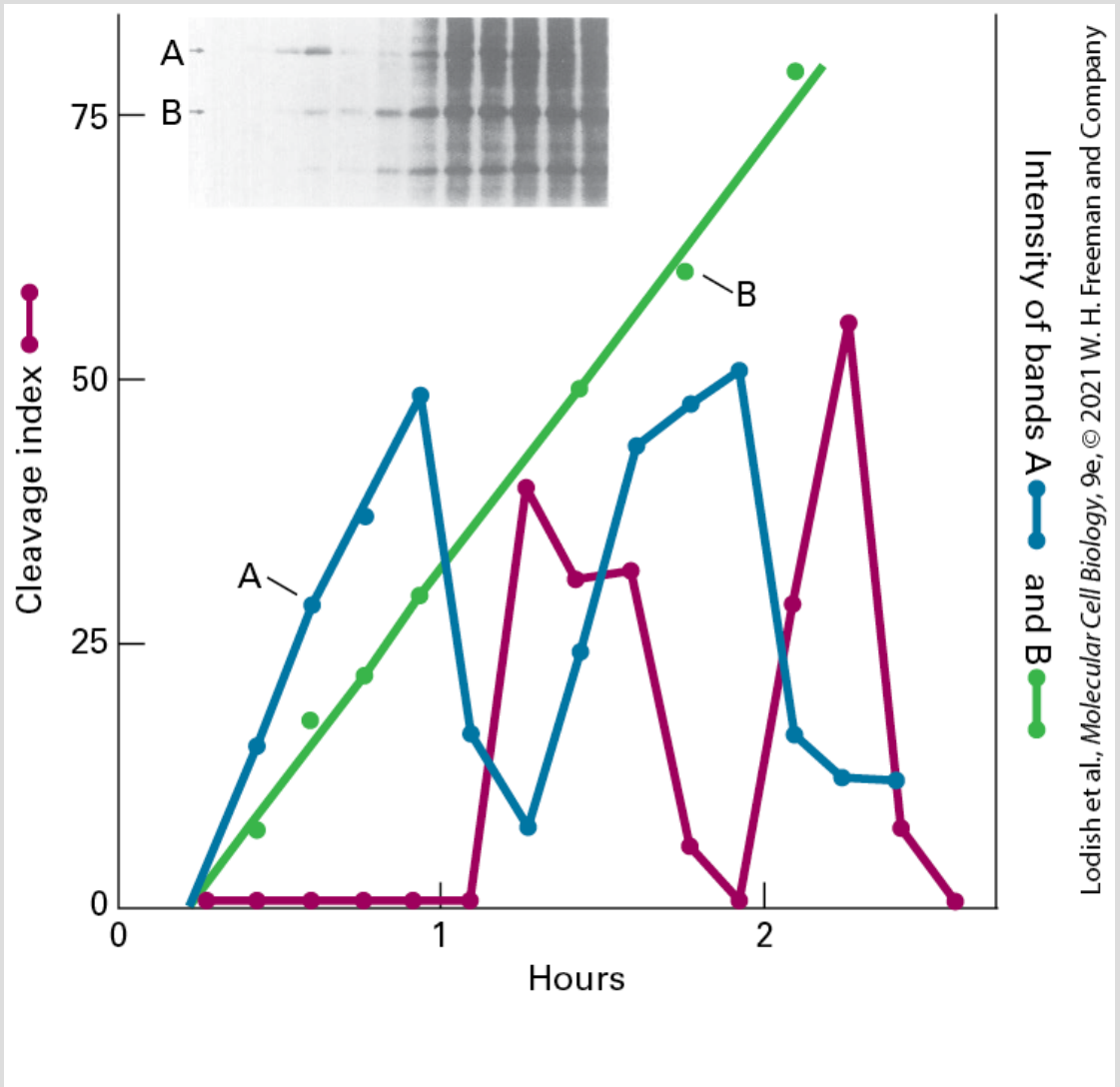
The G_1/S cyclins accumulate during late G_1, reach peak levels when cells enter S phase, and decline during S phase (Figure 19-10b). They are known as *cyclin E* in metazoans and bind to CDK2. The main function of cyclin E–CDK2 complexes, together with cyclin D–CDK4/6, is to trigger the G_1–S phase transition. This transition, which we defined in Section 19.1 as START in yeast and the restriction point in mammalian cells, is defined as the point at which cells are irreversibly committed to cell division and can no longer return to the G_1 state. In molecular terms, this means that cells initiate DNA replication as well as duplicating their centrosomes, which is the first step in the formation of the mitotic spindle that will be used during mitosis.

S phase cyclins are synthesized concomitantly with G_1 cyclins, but their levels remain high throughout S phase and do not decline until early mitosis. Two types of S phase cyclins trigger S phase in metazoans: cyclin E, which can also promote entry into the cell cycle and is therefore also a G_1/S cyclin, and cyclin A. Both cyclins bind CDK2 and are directly responsible for DNA synthesis. As we will see in Section 19.4, these protein kinases phosphorylate proteins that activate DNA helicases and load polymerases onto DNA.

Mitotic cyclins bind to CDK1 to promote entry into and progression through mitosis. The metazoan mitotic cyclins are cyclins A and cyclins B (note that cyclin A can also trigger S phase when bound to CDK2). Mitotic cyclin-CDK complexes are synthesized during S phase and G_2, but as we will see shortly, their activities are held in check until DNA synthesis is completed. In Section 19.5, we will see that once activated, mitotic CDKs promote entry into mitosis by phosphorylating and activating hundreds of proteins to promote chromosome condensation, nuclear envelope breakdown, mitotic spindle formation, and other aspects of mitosis. Their inactivation during anaphase prompts cells to exit mitosis, which involves the disassembly of the mitotic spindle, chromosome decondensation, the re-formation of the nuclear envelope, and eventually cytokinesis.

Mitotic cyclins were the first set of cyclins to be discovered, and it was their characterization that led to the discovery of the oscillatory nature of the activities that govern all of cell cycle progression. An impressive series of genetic and biochemical experiments in fission yeast by Paul Nurse and colleagues in the 1980s, building on related work from Lee Hartwell in budding yeast, had established that mitotic entry required the activity of a specific protein kinase, called *Cdc2* in fission yeast and *Cdc28* in budding yeast, and characterized some aspects of how this kinase was regulated. Independently, biochemical experiments in sea star embryos and eggs from frogs, surf clams, and sea urchins in the 1970s and 1980s had suggested that the maturation-promoting factor/mitosis-promoting factor (MPF) in those organisms, which we described in Section 19.2, might involve this same protein kinase. In a remarkable experiment in 1982, while studying protein synthesis in sea urchin eggs after fertilization, Tim Hunt unexpectedly observed a protein that showed cycles of synthesis followed by rapid destruction. The timing of this protein's destruction occurred just before the cells physically split into two, suggesting that its synthesis correlated with mitotic entry, and its destruction correlated with completion of mitosis (Figure 19-11). A similar pattern of cyclic protein synthesis and destruction was seen for two related proteins in clam eggs. These findings hinted that these proteins might be the mysterious labile factor(s) mentioned in Section 19.1 that control the switch from G_2 to mitosis. Cloning of their cDNAs further contributed to the identification of the proteins as sea urchin cyclin B and clam cyclin A. These cyclin

proteins were not the critical protein kinases that drove budding yeast, fission yeast, or vertebrate eggs and embryos into mitosis themselves, but cyclin immunoprecipitates (see Section 3.5) from the eggs had strong protein kinase activity. Furthermore, the addition of the sea urchin cyclin cDNA to RNase-treated frog egg extracts could drive the extracts through several rounds of cell cycles, indicating that these cyclin proteins were associating with and activating the protein kinases that were responsible for cell division.



EXPERIMENTAL FIGURE 19-11 How cyclins were discovered. The discovery of cyclins was based on the observation that the levels of a newly translated protein, labeled A in the autoradiogram, peaked in abundance and then rapidly declined each time before a sea urchin cell divided. In this experiment, the eggs were fertilized and incubated with radioactive methionine. Samples were then analyzed at 10-minute intervals using SDS-PAGE and autoradiography, and the eggs were also examined microscopically for cleavage. Compare the levels of one protein on the gel, labeled B, which progressively accumulates, with the levels of protein A that seems to rise and fall just before the cells undergo cleavage. The band intensities and cleavage fraction, as a function of time, are quantified in the graph. Protein A is the mitotic cyclin.

[Reprinted with permission from Elsevier, from Evans et al., 1983, "Cyclin: A Protein Specified by Maternal mRNA in Sea Urchin Eggs That Is Destroyed at Each Cleavage Division," *Cell* **33**:389–396; permission conveyed through Copyright Clearance Center, Inc.]

Ultimately, the composition of MPF from frog eggs was resolved by biochemical purification of MPF followed by protein analysis using SDS-PAGE. This revealed two bands, one at 32 kDa, similar in size to that of the yeast Cdc2 kinase, and a second band at 45 kDa that corresponded to cyclin B, indicating that frog egg MPF was, in fact, a complex between cyclin B and the mitotic CDK1 kinase. At around the same time, genetics experiments in fission yeast revealed a critical regulator for Cdc2, the yeast CDK, to be another cell division cycle gene, Cdc13. Cloning and sequencing of Cdc13 revealed that it was the fission yeast homolog of cyclin B. The levels of Cdc13 were shown to vary during the yeast cell cycle, just as cyclin B did during the sea urchin cell cycle. In addition, just as cyclin B was required for CDK1 kinase activity, Cdc13 was required for Cdc2 protein kinase

activity, and this complex, formed by a mitotic cyclin and a cyclin-dependent kinase (Cdc13 and Cdc2, respectively) was required for fission yeast to progress from G_2 into M phase. This combination of genetics and biochemical experiments performed in so many of the different organisms discussed in Section 19.2 ultimately revealed a universal mechanism that controls mitotic entry for all eukaryotic life forms: the formation of a complex between an unstable cyclin molecule and a small cyclin-dependent protein kinase whose activity is strictly dependent on the presence of the cyclin subunit. Subsequent studies showed that G_1/S phase cyclins had similar properties. Their expression is sufficient to promote entry of resting cells from G_1 into S phase, and subsequent movement through S phase.

How exactly does cyclin-binding convert CDKs from an inactive to an active state? The three-dimensional structure of CDKs and cyclin-CDK complexes, solved for cyclin B/CDK1 and cyclin A/CDK2 by x-ray crystallography, provides an answer (Figure 19-12). As discussed in Chapter 3, cyclin-dependent kinases, like all protein kinases, are bilobed structures, with an active site cleft where both ATP and substrates bind. In the absence of cyclin binding, however, several of the catalytic residues in the CDK are not in the proper position to bind ATP and catalyze the transfer of a phosphate group from ATP to the substrate proteins. In addition, the activation loop of the monomeric CDK blocks substrates from accessing the active site cleft. Cyclins are composed of two domains, each containing five α

helices. The overall structure is called a *cyclin fold*, and the first domain corresponds to the *cyclin box* region that is required for CDK binding and activation. In the case of CDK1 and CDK2, binding of the cyclin box causes an α helix in the CDK to move, reorganizing the amino acids in the ATP-binding pocket of the CDK into a more productive conformation for catalysis. In addition, cyclin binding structures and repositions the activation loop, as we discuss shortly. Besides stimulating CDK activity, the cyclin subunit can also interact directly with substrate proteins of the kinase, thereby contributing to the substrate specificity of the cyclin-CDK complex.

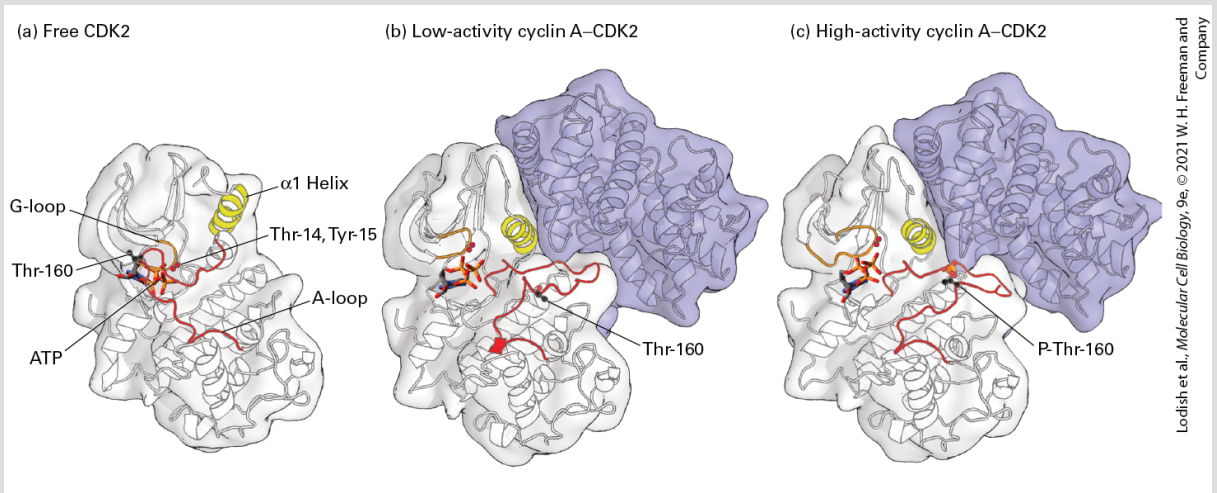


FIGURE 19-12 Structural models of human CDK2. (a) Free, inactive CDK2 not bound to its cyclin subunit, cyclin A. In free CDK2, the A-loop blocks access of protein substrates to the γ phosphate of the bound ATP, shown as a stick model. The conformations of the A-loop and the region highlighted in yellow (α1 helix) are altered when CDK is bound to cyclin A. (b) Nonphosphorylated, low-activity cyclin A-CDK2 complex. Conformational changes induced by binding of a domain of cyclin A (blue) cause the A-loop to pull away from the active site of CDK2 so that substrate proteins can bind. The α1 helix in CDK2, which interacts extensively with cyclin A, moves several angstroms into the catalytic cleft, repositioning key catalytic side chains required for the phosphotransfer reaction. The black ball marks the position of the threonine (Thr-160) whose phosphorylation by CAK activates CDKs. The red balls mark

the positions of phosphorylated Thr-14 and Tyr-15, which inhibit CDK activity. (c) Phosphorylated, high-activity cyclin A-CDK2 complex. The conformational changes induced by phosphorylation of the activating threonine (black ball) alter the shape of the substrate-binding surface, greatly increasing the affinity for protein substrates. Removal of the phosphate groups from Thr-14 and Tyr-15 enhances ATP binding and catalytic activity. See P. D. Jeffrey et al., 1995, *Nature* **376**:313–320.

[Data from A. A. Russo, P. D. Jeffrey, and N. P. Pavletich, 1996, *Nat. Struct. Biol.* **3**:696, PDB ID 1jst.]

CDKs Are Regulated by Activating and Inhibitory Phosphorylation

As we will discuss in detail below, in addition to cyclin-binding, a variety of other mechanisms ensure that CDKs are only active at the right stage of the cell cycle. Table 19-1 lists many of these additional key regulators of CDKs.

TABLE 19-1 • Regulators of Cyclin-CDK Activity

Type of Regulator	Function
Kinases and Phosphatases	
CAK kinase	Activates CDKs in all organisms
Wee1 kinase	Inhibits CDKs in all organisms
Cdc25 phosphatase	Activates CDKs in fission yeast

Cdc14 phosphatase	Activates Cdh1 to degrade mitotic cyclins in budding yeast
Cdc25A phosphatase	Activates vertebrate S phase CDKs
Cdc25B and C phosphatases	Activate vertebrate mitotic CDKs
Inhibitory Proteins	
Sic1	Binds and inhibits S phase CDKs in budding yeast
CKIs $p27^{KIP1}$, $p57^{KIP2}$, and $p21^{CIP}$	Bind and inhibit CDKs in animal cells
INK4	Binds and inhibits G_1 CDKs in animal cells
Rb	Binds E2Fs, preventing transcription of multiple cell cycle genes in animal cells
Ubiquitin-Protein Ligases	
SCF	Degradation of phosphorylated Sic1 or $p27^{KIP1}$ to activate S phase CDKs
APC/C^{Cdc20}	Degradation of securin, initiating anaphase. Induces degradation of B-type cyclins
APC/C^{Cdh1}	Degradation of B-type cyclins in G_1 and geminin in metazoans to allow loading of replicative helicases on DNA replication origins
Phosphoserine/Threonine-Binding Proteins and Modular Domains	
WD40 domains, leucine-rich	Recognize phosphorylated substrates and target them for degradation by SCF ligases

repeats	
Cks subunits	Target cyclin-CDKs for processive substrate phosphorylation
14-3-3 proteins	Bind and inactivate regulators of cyclin-CDKs to establish cell cycle checkpoints. Also regulates cytokinesis.
Pin1	Isomerizes phosphorylated substrates of CDKs to control mitotic entry and exit
FHA and BRCT domains	Form phospho-dependent molecular assemblies necessary for cell cycle surveillance mechanisms in all stages of the cell cycle, and regulates cytokinesis.
Polo-box domains	Target Polo kinases to sites of CDK phosphorylation to drive mitotic entry and progression

If cyclin-binding was the only thing required for CDKs to become fully active, then one would expect the activity of each CDK to perfectly parallel the accumulation of its cyclin partner. However, that is not what is observed for most cyclin-CDK pairs. Figure 19-13a shows the accumulation of cyclin B, and the activity of MPF (the cyclin B-CDK1 complex in *Xenopus* oocytes). The abrupt rise of MPF activity observed upon entering mitosis, despite the slow progressive rise in the amount of cyclin B present, indicates that additional types of regulation must be present besides cyclin binding. Indeed, at least two additional mechanisms are involved in controlling the activity of CDKs: activating and inhibitory phosphorylations on the CDK subunit, and the action of inhibitor proteins, called *CKIs (c*yclin *k*inase *i*nhibitors) that bind to CDK to

block its activity. These three regulatory events — cyclin binding, phosphorylation, and CKIs — together with cyclin synthesis and destruction, ensure that CDKs are active only at the appropriate time and cell cycle stage.

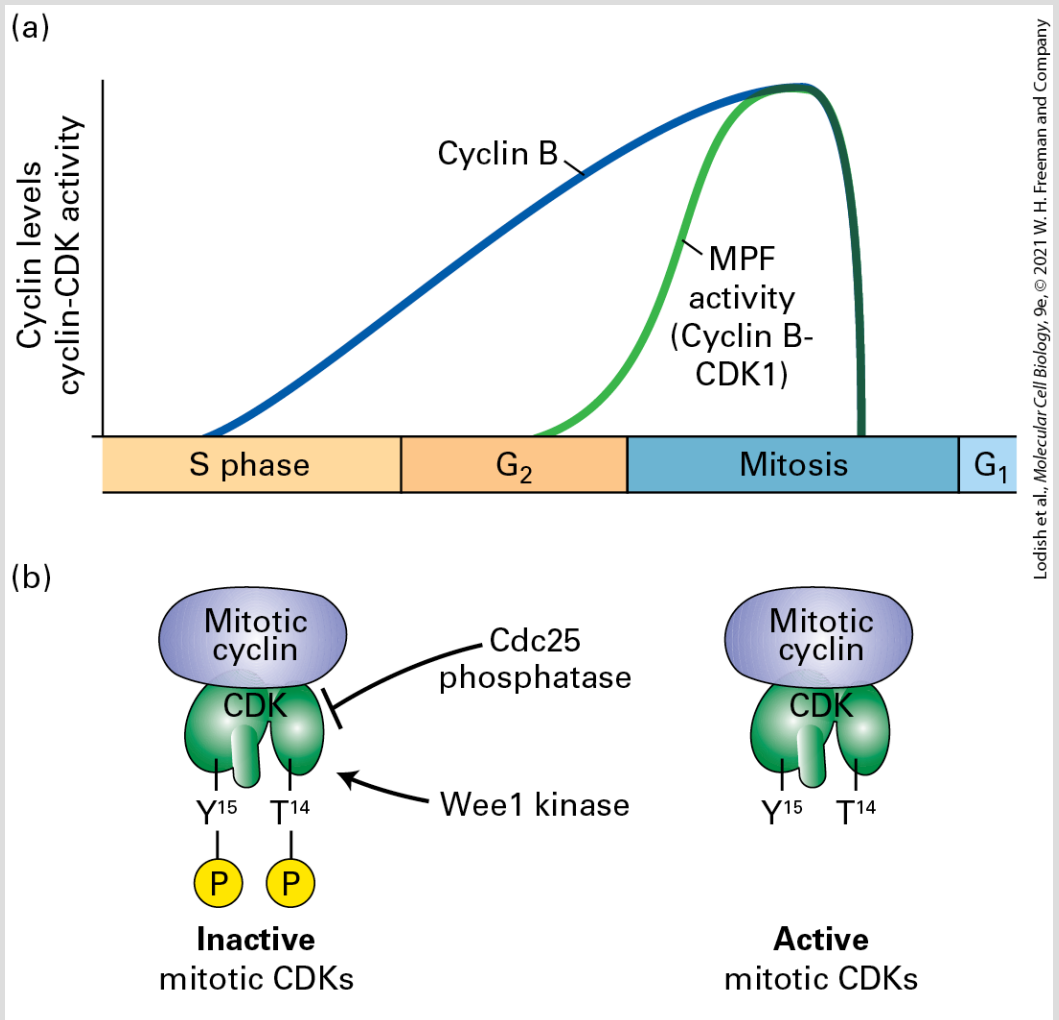


FIGURE 19-13 Phosphorylation of the CDK subunit restrains mitotic CDK activity during S phase and G_2. (a) Mitotic cyclin B is synthesized during S phase and G_2 and binds to CDK1, but the complex does not become active until aound the time of mitotic entry. (b) The cyclin-CDK complex is not active because threonine 14 and tyrosine 15 of the CDK1 subunit are phosphorylated by the protein kinase Wee1, which inhibits the accumulation of the active form. Once DNA replication has been completed, the protein phosphatase Cdc25 is

activated and de-phosphorylates CDK1, promoting its activity. A positive feedback loop, indicated by the black curved arrows and discussed in Section 19.5, controls the activation of Cdc25 and inhibits Wee1 in animal cells.

Nonphosphorylated CDK bound to one of its cyclin partners has minimal but detectable protein kinase activity in vitro, although it may be essentially inactive in vivo. This is because the activation loop, though partially structured by cyclin binding, still largely blocks the catalytic cleft (see Figure 19-12b). Phosphorylation of a critical threonine residue in the activation loop, Thr-160, is needed to pull the activation loop out of the way by tethering it to several basic residues in the CDK, allowing substrates to access the active site of the CDK and increasing the activity of the phosphorylated cyclin-CDK complex by over a hundredfold compared to the nonphosphorylated complex (see Figure 19-12c). This threonine phosphorylation is mediated by **CDK-activating kinase (CAK)**. Curiously, CAK activity is constant throughout the cell cycle and the activating phosphorylation of the CDK occurs either before or as soon as a cyclin-CDK complex is formed, so it is clear that CAK phosphorylation of the CDK is not a rate-limiting step in CDK activation. Why, then, does cyclin B/CDK1 (MPF) activity rise so rapidly when cells go from G_2 into M phase (Figure 19-13), if the threonine-phosphorylated cyclin B–bound kinase is already present? This behavior arises because the CDK also contains inhibitory phosphorylations.

The two inhibitory phosphorylations on CDK play a critical role in controlling CDK activity. A highly conserved tyrosine (Y15 in human CDKs) and an adjacent threonine (T14 in humans) are subject to regulated phosphorylation (see Figures 19-12b and 19-13b). Importantly, both of these residues are situated in the G-loop of the CDK, containing the consensus sequence GXGXXGXV, that drapes over the phosphate groups in ATP (see Figures 19-13 and 3-38). In the case of CDKs, the G-loop sequence is $GEGT_{14}Y_{15}GVV$. Phosphorylation of T14 and Y15 in this G-loop introduces significant negative charge that electrostatically interferes with the binding and positioning of ATP in the catalytic pocket. Changes in the phosphorylation of these sites are essential for the regulation of mitotic CDKs and have also been implicated in the control of G_1/S and S phase CDKs. As we will see in Section 19.5, a highly conserved kinase called *Wee1* brings about these inhibitory phosphorylations, and a family of highly conserved phosphatases called *Cdc25*, mediates their de-phosphorylation (Figure 19-13b).

CDK Inhibitors Provide Additional Control of Cyclin-CDK Activity

In addition to cyclin binding and reversible phosphorylations on the CDK itself, there is one more level of control of cyclin-CDK activity. A family of proteins known as *CDK inhibitors* or *CKIs* bind directly to CDKs and cyclin-CDK complexes to inhibit their activity. Different

classes of CKIs act in G_1, at the G_1–S boundary, or in S and G_2 phases of the cell cycle to prevent premature activation of G_1, S phase, and M phase CDKs.

An important class of CKIs acting in G_1 are members of the *INK4s* (*in*hibitors of *k*inase *4*) family. This family of proteins includes several small closely related proteins that bind to CDK4 and CDK6 monomers, blocking their interaction with cyclin D and thus rendering these G_1 CDKs inactive. As we will see in Section 19.4, these G_1 CDKs play an especially important role in the regulation of the G_1–S phase transition, and its integration with extracellular signals.

A second class of CKIs found in metazoan cells is the Cip/Kip family of CKIs. This family consists of three proteins — $p21^{CIP1}$, $p27^{KIP1}$, and $p57^{KIP2}$. The Cip/Kip family members bind to cyclin-bound CDKs and block the activity of cyclin D-, E-, A-, and B-containing complexes. Cip/Kip CKIs must be sequestered away from their respective cyclin/CDK targets or degraded in order for DNA replication or mitosis to commence. Each of the three proteins appears to play a unique physiological role. The $p21^{Cip1}$ protein, encoded by the *cdkn1a* gene, is important in establishing G_1 and G_2 arrest in metazoan cells in response to DNA damage, and its transcription is controlled by the important tumor suppressor protein p53. The $p27^{Kip1}$ protein, encoded by the *cdkn1b* gene, is critical for blocking progression from G_1 into S phase in response to

nutrient and growth factor limitation in order to limit organ and body size. Mice in whom the $p27^{Kip1}$ gene has been knocked out display increased body size and multiple organ hyperplasia (i.e., an excessive number of cells in their organs). The $p57^{Kip2}$ protein, encoded by the *cdkn1c* gene, plays an important role in regulating the cell cycle during embryonic development. Mouse embryos lacking *cdkn1c* display organ hyperplasia and impaired tissue differentiation, resulting in death soon after birth. Based on the crystal structure of $p27^{Kip1}$ bound to cyclin A-CDK2, the Cip/Kip proteins are thought to function both by blocking the site on the cyclin that interacts with cyclin-CDK substrates and occluding the active site cleft to block ATP binding and catalytic activity (see Figure 19-12). These proteins have other functions beyond cell cycle control, including effects on gene transcription, cell fate determination, cell migration, and apoptosis. Nonetheless, the Cip/Kip family members play critical roles in proper cell and tissue homeostasis.

Cyclin Levels Are Regulated by Transcriptional Activation and Ubiquitin-Mediated Protein Degradation

The expression and degradation of cyclins are carefully regulated so that the appropriate cyclins are present in the cell cycle stage at which they are needed. In this section, we discuss how cyclin levels are regulated.

Control of cyclin gene transcription is one mechanism that ensures proper temporal expression of the cyclins. In somatic cells and yeast, waves of transcription factor activity help establish waves of cyclin activity. A general principle is that an earlier wave of CDK phosphorylation-dependent transcriptional activity helps produce the factors essential to generate a subsequent CDK phosphorylation-dependent transcriptional wave. As we will see in Section 19.4, transcription of G_1/S phase cyclins is promoted by the E2F transcription factor complex that becomes activated at this point in the cell cycle.

A very important mechanism that restricts cyclins to the appropriate cell cycle stage is ubiquitin-mediated, proteasome-dependent protein degradation. Because protein degradation is an irreversible process, in the sense that the protein can be replenished only through de novo protein synthesis, this regulatory mechanism is ideal to ensure that the cell cycle engine is driven forward and that cells cannot go backward in the cell cycle.

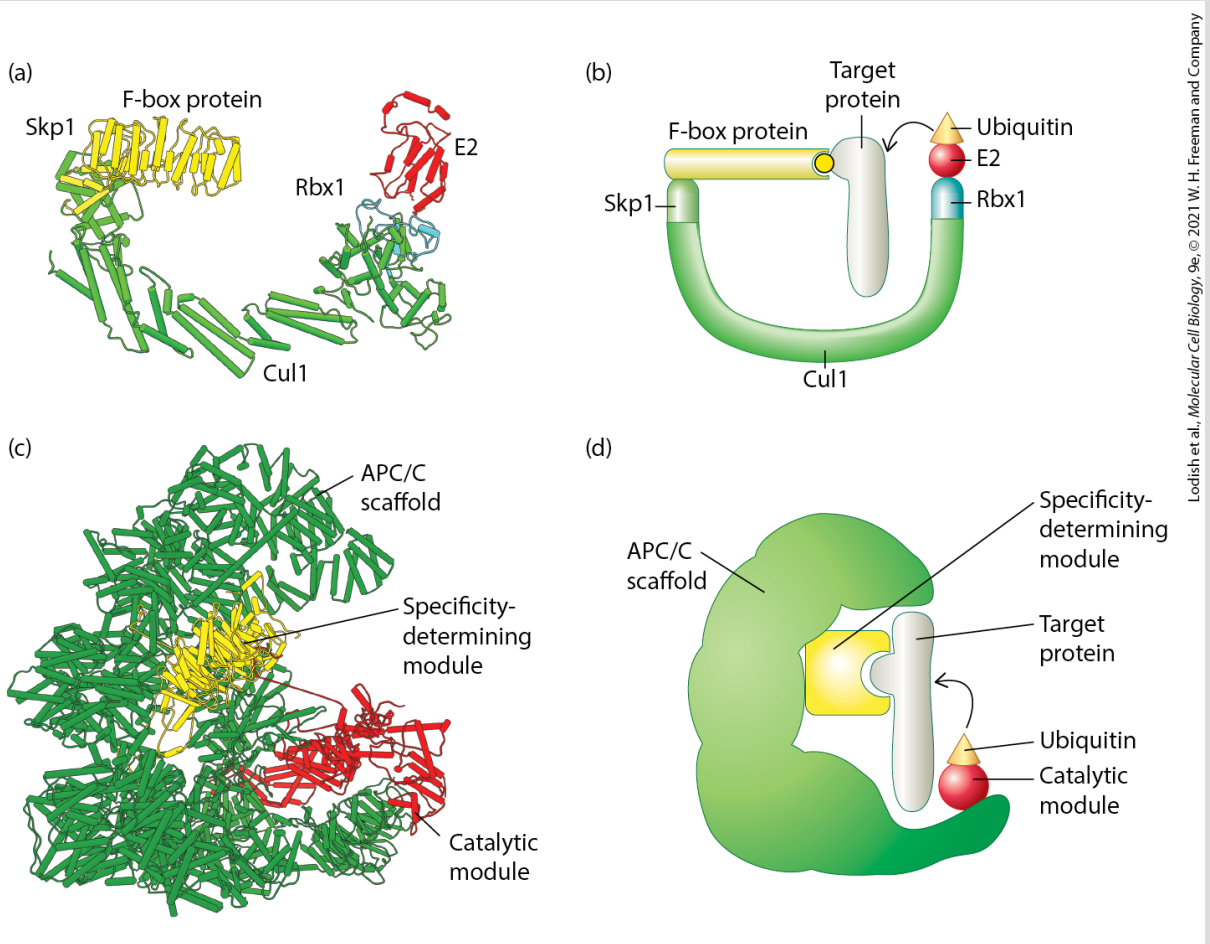
Recall from Chapter 3 that ubiquitin-protein ligases, also called *E3-ligases*, transfer ubiquitin from an E2 ubiquitin-conjugating enzyme

to one or more lysine residues on the E3-ligase's substrate proteins, making long polyubiquitin chains that mark these proteins for degradation by proteasomes (see Figure 3-32). Cyclins are degraded through the action of two different E3-ligases, called **SCF** (named after the first letters of its constituent proteins, *S*kp1, *C*ullin, and *F*-box proteins) and the **anaphase-promoting complex** or **cyclosome** (abbreviated as **APC/C** in this chapter). The mechanism of regulation of these two E3 ligases that are so critical for cell cycle control is strikingly different, though both involve protein phosphorylation. In the case of the SCF complex, it is the substrates that are phosphorylated to target them for degradation. In the case of the APC/C, it is the APC/C itself that is phosphorylated, controlling its activity.

SCF controls the G_1/S transition by targeting the G_1 and G_1/S phase cyclins for degradation. SCF controls the G_2/M transition by targeting the CDK inhibitory kinase Wee1, along with the $p21^{Cip1}$ and $p27^{Kip1}$ CDK inhibitory proteins for degradation, as we shall see in more detail shortly. The APC/C, in contrast, becomes activated at the onset of mitosis and remains active through the next G_1 phase of the cell cycle. The APC/C targets both S phase and mitotic cyclins for degradation. Activation of the APC/C during mitosis controls the onset of chromosome segregation at the metaphase-anaphase transition by degrading an anaphase inhibitory protein (as discussed in Section 19.6) and promotes mitotic exit. APC/C activity during G_1 inhibits mitotic and S phase CDKs, which allows replication origins

to be reset so that a new round of replication can take place during the subsequent S phase (as discussed in Section 19.4).

SCF and APC/C are multisubunit ubiquitin-protein ligases that belong to the RING finger family of ubiquitin-protein ligases. Despite the fact that SCF and APC/C belong to the same ubiquitin-protein ligase family, their structures and substrate targeting regulation are quite different. The structures of both complexes are shown in Figure 19-14. These molecular assemblies resemble U- or C-shaped clamps. Human SCF complexes contain three core components and a substrate-targeting F-box protein. As shown in Figure 19-14a and b, the Cul1 subunit forms a central structural scaffold that links RBX1 and Skp1. RBX1 is a RING finger–containing protein that binds the E2 ubiquitin-conjugating enzyme. Skp1 is a bridging protein that binds the F-box protein. A number of different interchangeable F-box proteins bind to Skp1, and each one recruits a distinct set of substrate proteins to the SCF, where they are ubiquitinated and targeted for destruction.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-14 E3 ubiquitin ligases control cell cycle progression. (a) SCF ligases are built from a central Cullin scaffold (Cul1) with a RING finger–containing Rbx1 protein and an E2 ubiquitin-conjugating enzyme at one end, and a Skp1 protein bound to an F-box substrate-targeting protein at the other. In this example, the F-box protein is Skp2, which uses a series of leucine-rich repeats to bind to target proteins only when they are phosphorylated. (b) This general architecture is conserved in all SCF ligases, but different phosphoserine/threonine-binding F-box proteins determine the substrate specificity. (c) The APC/C E3 ligases are assembled from 13 or 14 subunits, which are organized like SCF complexes into scaffolding (green), substrate-binding (yellow), and catalytic (red) functions. (d) The APC/C is inactive until it binds to one of the substrate-targeting co-activators, either Cdc20 or Cdh1. This binding activates the catalytic components and positions the Cdc20- or Cdh1-bound target protein for conjugation to ubiquitin (orange triangles).

The anaphase promoting complex or cyclosome (APC/C) consists of 14 core subunits in metazoans and 13 core subunits in yeast, together with one of two related substrate-targeting factors called *Cdc20* and *Cdh1* (Figure 19-14c and d). The activity of APC/C is entirely dependent on binding of the core complex to one of these two substrate-targeting subunits. At the metaphase to anaphase transition, APC/C bound to Cdc20 targets cyclins A and B for degradation, while during anaphase, APC/C bound to Cdc20 ubiquitinylates proteins that inhibit chromosome segregation. During late anaphase, telophase and G_1, APC/C bound to Cdh1 targets different substrates for degradation. Like SCF complexes, the APC/C core is organized into three modules (Figure 19-14c): a scaffold module that structurally organizes the entire complex, a catalytic module that accepts activated ubiquitin from an E2 enzyme, and a substrate recognition module that determines which proteins the APC/C will ubiquitinate.

In the absence of phosphorylation, a loop on one of the APC/C core subunits blocks the APC/C binding site for the substrate-targeting subunit Cdc20. Phosphorylation of this loop by mitotic CDKs shifts its position in a similar manner to how phosphorylation activates protein kinases (Section 3.4). This allows Cdc20 to bind and results in the formation of an active $\mathrm{APC/C}^{\mathrm{Cdc20}}$ complex that can ubiquitinate and target cyclins A and B for degradation, thereby shutting off mitotic CDK activity. Mitotic CDKs also phosphorylate the Cdh1 substrate-targeting subunit, which prevents it from

competing with the Cdc20 subunit for binding during metaphase and early anaphase. After $\mathrm{APC/C}^{\mathrm{Cdc20}}$ targets the mitotic cyclins for degradation, the APC/C and Cdh1 become dephosphorylated through the action of protein phosphatases, particularly PP2A in vertebrate cells or Cdc14 in yeast, preventing Cdc20 binding but instead allowing Cdh1 to bind and form an active $\mathrm{APC/C}^{\mathrm{Cdh1}}$ E3-ligase complex in late anaphase that persists until late G_1 of the next cell cycle.

Substrate-targeting subunits of the APC/C recognize specific sequence motifs on the proteins that the APC/C ubiquitinates. One motif, the destruction box or D-box, has the consensus sequence RXXLX[D/E]ΦΦΦXN[N/S], where X stands for any amino acid, Φ means any hydrophobic amino acid, and the brackets indicate that either of the two amino acids is found at that position in the sequence. Experimentally mutating or deleting the destruction box in mitotic cyclins prevented their ubiquitination and blocked cells from exiting mitosis, demonstrating that mitotic exit requires degradation of mitotic cyclins.

Phosphoserine/Threonine-Binding Domains Build Feedback Loops That Coordinate CDK Activation and Cell Cycle Progression

The final molecular components that we need to discuss in order to understand cyclin-CDK regulation are phosphoserine/threonine-binding domains. We have seen that phosphorylation by protein kinases plays an inordinately important role in controlling cyclin abundance, cyclin degradation, and CDK activity. While some of this regulation is a direct consequence of allosteric changes in the shapes of the phosphorylated proteins themselves, another important way that protein phosphorylation regulates the cell cycle is through the creation of short phosphoserine- or phosphothreonine-containing sequence motifs that are targeted for recognition by phospho-binding proteins or protein domains. The functions of these phosphoserine/threonine-binding domains are summarized in Table 19-1.

These phosphoserine/threonine-binding proteins and modules orchestrate nearly all aspects of cell cycle control. In the most general sense, binding of phosphoserine/threonine-binding domains to CDK substrates or to proteins that control CDK activation works by either targeting the bound proteins to specific subcellular locations, altering their three-dimensional conformation, changing their activity, or targeting them for specific modifications such as additional phosphorylation or ubiquitin-mediated destruction. One example of this is found in the F-box proteins that comprise the substrate-targeting subunits of SCF ligases mentioned previously. Many F-box proteins contain modular phosphoserine/threonine-binding domains that bind to a substrate only if the substrate contains a specific phosphoserine/threonine-containing sequence.

Thus the substrate protein is targeted for ubiquitination and degradation only if it has been phosphorylated previously by a protein kinase. As we will see in the next section, CDK-mediated phosphorylation of a CDK inhibitor can target it for SCF-mediated degradation, leading to more CDK activity, followed by additional CKI degradation, creating a feed-forward loop that amplifies total CDK activity. Other phosphoserine/threonine-binding molecules, including 14-3-3 proteins that regulate the cell cycle, Cks proteins that promote CDK phosphorylation, a proline isomerase called *Pin1* that changes the activity of phosphorylated cell cycle regulators, Polo-box domains that target Polo kinase 1 to its substrates in mitosis, and FHA and BRCT domains that provide surveillance for DNA damage and control cytokinesis. These phospho-binding modules are necessary to form the positive and negative feedback loops that control progression through the cell cycle (Section 19-6) and for surveillance mechanisms that stop the cell cycle in response to catastrophic events (Section 19-7). The importance of being able to specifically recognize short sequences created by protein kinases as a mechanism of cell cycle control is evident from the fact that many different types of protein structures are capable of this type of phospho-specific binding, indicating that the phospho-binding function likely arose in these different proteins by convergent evolution.

Mass Spectrometry Studies and Genetically Engineered CDKs Led to the Discovery of New CDK Substrates and Functions

Different CDKs initiate different cell cycle phases by phosphorylating specific proteins. It is now clear that CDKs phosphorylate a myriad of substrates, thereby directly initiating all aspects of a given cell cycle phase.

In recent years, there have been systematic efforts to identify all of the CDK substrates. The challenge in identifying the substrates of a particular kinase is to distinguish one kinase's phosphorylation events from those carried out by other kinases. Examination of mapped phosphorylation sites on known substrates of cyclin-CDKs and studies using synthetic peptides have revealed that cyclin-CDK complexes preferentially phosphorylate serine and threonine residues when they are found within the consensus sequence [S/T]PX[R/K]K where [S/T] (labeled the P0 position) is the site of CDK phosphorylation, and the sequence contains either an arginine or lysine in the P3 position, a lysine in the P4 position, or both.

A breakthrough in understanding which proteins are targets of CDKs was facilitated by the engineering of a CDK mutant that uses an ATP analog that is not bound by other kinases. The ATP analog has a

bulky benzyl group attached to N_6 of the adenine, which makes the analog too large to fit into the ATP-binding pocket of wild-type protein kinases. However, the ATP-binding pocket of the mutant CDK accommodates this N_6-benzyl ATP analog. Consequently, only the mutant CDK can use this ATP analog as a substrate for transferring its γ phosphate to a protein side chain. When the N_6-benzyl ATP analog with a thiol-labeled γ phosphate was incubated with cell extracts containing a recombinant mitotic CDK with the altered ATP-binding pocket, multiple proteins were labeled and then purified and identified by mass spectrometry (Figure 19-15). In yeast, this procedure identified most of the known CDK substrates plus more than 150 additional yeast proteins.

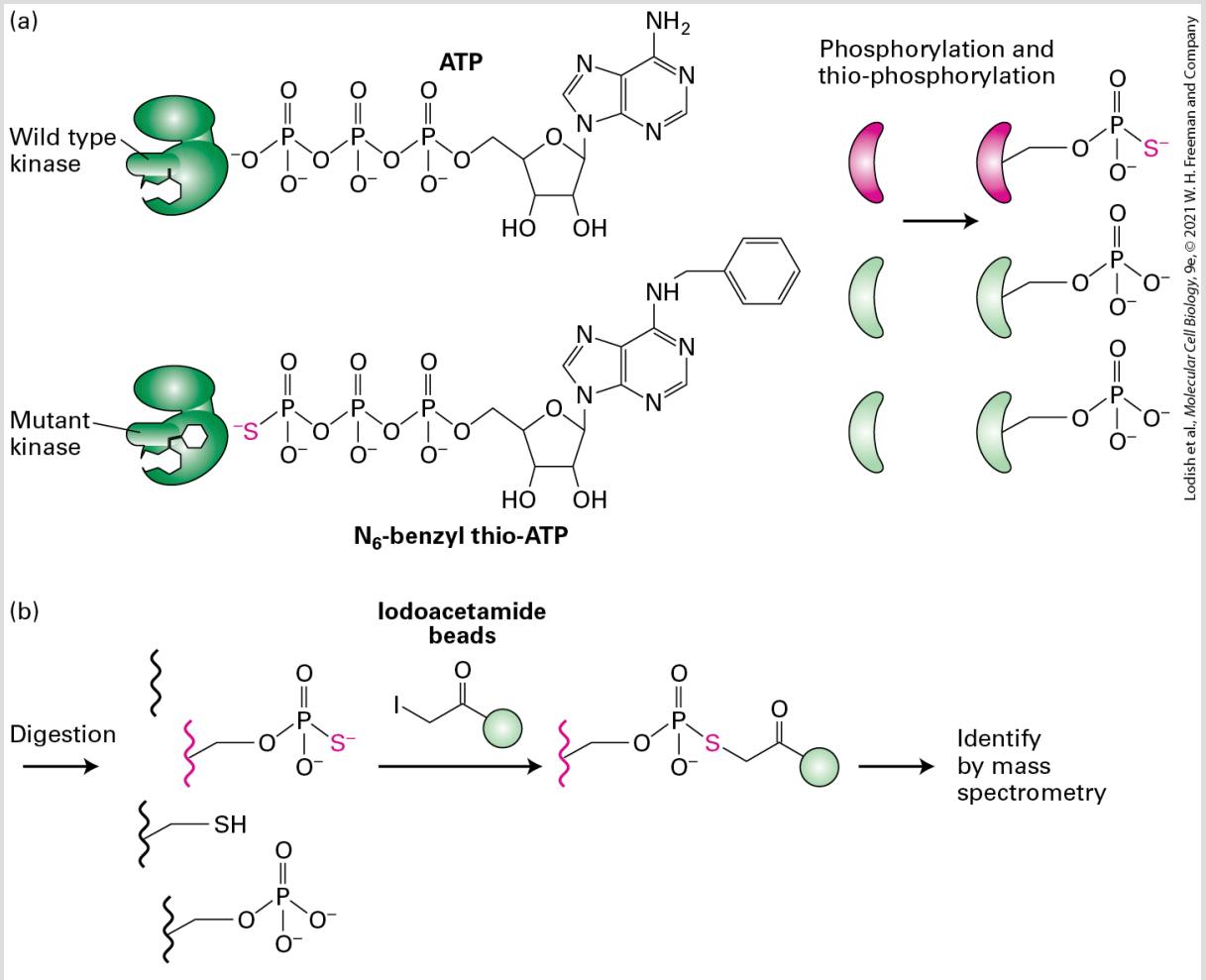


FIGURE 19-15 Identification of CDK substrates using genetically engineered CDK mutants. (a) The ATP-binding pocket of CDKs was engineered to accept a bulky N_6-benzyl ATP analog also containing a sulfur atom in place of a γ phosphate oxygen atom. Substrates of the mutant kinase that have been labeled with the ATP analog will contain a sulfur atom. (b) After protease digestion of cell proteins, the sulfur-containing peptides are purified using iodoacetamide beads (which bind sulfur), and the peptides are identified using mass spectrometry.

We have now discussed all of the core components necessary for the cell to circumnavigate around the cell cycle: cyclins, CDKs, the phosphoserine/threonine motifs that CDKs create on their substrates, transcription factors, other protein kinases,

phosphatases, CKIs, E3 ligases and the sequence motifs that they recognize, and phospho-binding proteins and domains. In the following sections, we see how these components work together to enable the cell to transition from one stage of the cell cycle to the next. We also describe the critical events that occur during each stage of the cell cycle.

KEY CONCEPTS OF SECTION 19.3

Cell Cycle Progression and Control: Feedback Loops and Post-Translational Modification

- Cyclin-dependent kinases are activated by cyclin subunits. Their activity is controlled at multiple levels.
- Different cyclin subunits activate CDKs at different cell cycle stages. Cyclins are present only in the cell cycle stages that they promote.
- Activating and inhibitory phosphorylation of the CDK subunit contributes to the regulation of CDK activity.
- CDK inhibitors (CKIs) inhibit CDK activity by binding directly to the cyclin-CDK complex.
- Protein degradation is the key mechanism responsible for restricting cyclins to the appropriate cell cycle stage. This degradation is mediated by the ubiquitin-proteasome system and the ubiquitin-protein ligases APC/C and SCF.
- Phosphoserine/threonine-binding proteins and modular domains are used to construct positive feed-forward and negative feedback loops that allow rapid activation and inactivation of CDK activity by binding to CDK regulators and substrates.
- CDKs initiate every aspect of each cell cycle stage by phosphorylating many different target proteins. Systematic efforts using protein kinases engineered to bind only modified forms of ATP have led to the identification of many of these substrates.

19.4 The Transition from G_1 into S Phase and DNA Replication

The previous section described the multiple mechanisms that control the different cyclin-CDK complexes. In this and the following two sections, we examine each cell cycle stage carefully and discuss how it is induced and controlled. We examine how cells initiate DNA replication and mitosis and how chromosomes are segregated. We focus on how cyclin-CDK complexes and other key cell cycle regulators affect each cell cycle phase, and we examine the mechanisms that coordinate their activities. This section investigates how cells decide whether or not to undergo cell division, the molecular basis for the transition from G_1 into S phase, and how DNA replication is initiated. We begin with yeast.

The G_1/S Transition in Budding Yeast Is Controlled by Cyclin-CDK Complexes

The process of cell cycle entry is particularly well understood in budding yeast, and it was in this organism that the molecular mechanisms underlying this cell cycle transition were initially elucidated.

Recall from Section 19.1 that the decision to enter S phase occurs late in G_1 and requires the accumulation of a labile R-factor that

drives the cells through the restriction point, called *START* in yeast. Elegant genetic experiments identified temperature-sensitive mutants in budding yeast that arrest in G_1, failing to form a bud and to initiate DNA replication. One group of these mutants was in the gene encoding the yeast CDK. (As mentioned earlier, budding yeast has only one CDK that triggers all cell cycle transitions and is known as *CDC28*.) This observation proved that CDK activity is essential for entry into S phase. A second set of mutants involved the gene for the labile R-factor, which turned out to be the yeast G_1 cyclin known as *CLN3*. We now know that a CDK cascade triggers further progress from G_1 through S phase in yeast (Figure 19-16a). The G_1 cyclin-CDK complex (Cln3-CDK in yeast) stimulates the formation of G_1/S phase cyclin-CDKs, which initiate bud formation, centrosome duplication, and DNA replication. The mRNA for the yeast G_1 cyclin gene *CLN3* is produced at a nearly constant level throughout the cell cycle, but its translation is regulated in response to nutrient levels and, as we will see shortly, it is a linchpin in coupling cell cycle entry to nutrient signals. Once sufficient Cln3 is synthesized from its mRNA, Cln3-CDK complexes phosphorylate and inactivate the transcriptional repressor Whi5. Phosphorylation of Whi5 promotes its export out of the nucleus, allowing the transcription factor complex SBF to induce transcription of the G_1/S phase cyclin genes *CLN1* and *CLN2* as well as other genes important for DNA replication. Once produced, Cln1/2-CDKs contribute to further Whi5 phosphorylation. This positive feedback loop ensures the rapid accumulation of G_1/S phase cyclin-CDKs. The point in the cell cycle at which 50 percent of

Whi5 has exited the nucleus marks the point when cells are irreversibly committed to division and provides a molecular definition of START. Cln1/2-CDKs then trigger bud formation, entry into S phase, and the duplication of the centrosome (also known as the *spindle pole body* in yeast), which will organize the mitotic spindle.

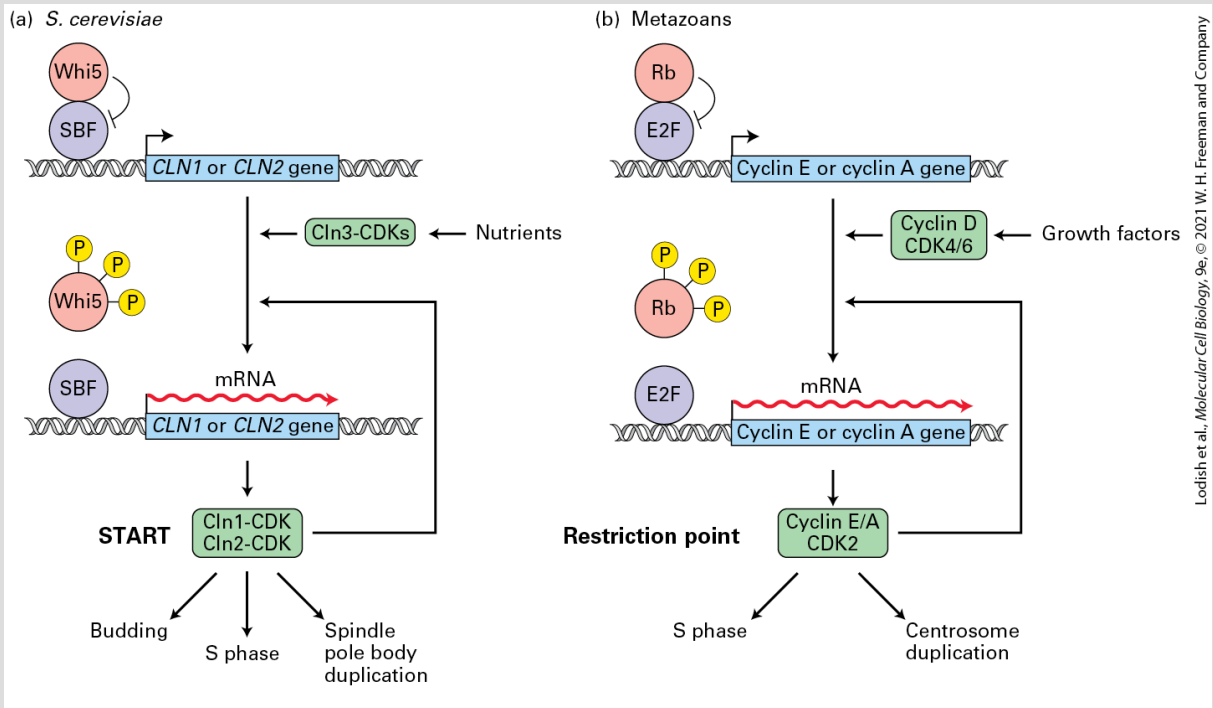


FIGURE 19-16 Control of the G_1–S phase transition. (a) In budding yeast, activity of the G_1 cyclin–CDK Cln3-CDK rises during G_1 and is controlled by nutrient availability. Once sufficiently active, the kinase phosphorylates the transcriptional repressor Whi5, promoting its export from the nucleus. This releases the transcription factor complex SBF, allowing it to induce the transcription of the G_1/S phase cyclin genes *CLN1* and *CLN2* and of other genes whose products are needed for DNA replication. G_1/S phase CDKs further phosphorylate Whi5, promoting further *CLN1* and *CLN2* transcription. Once sufficiently high levels of G_1/S phase CDKs have been produced, START is traversed. Cells irreversibly enter the cell cycle: they initiate DNA replication, bud formation, and spindle pole body duplication. (b) In vertebrates, G_1 CDK activity rises during G_1 and is stimulated by the presence of growth factors. When signaling from growth factors is sustained, the resulting cyclin D–CDK4/6

complexes begin phosphorylating Rb, releasing some E2F, which stimulates transcription of the genes encoding cyclin E, CDK2, and E2F itself. The cyclin E–CDK2 complexes further phosphorylate Rb, resulting in a positive feedback loop that leads to a rapid rise in the expression and activity of both E2F and cyclin E–CDK2. Once G_1/S phase CDKs are sufficiently high, cells pass through the restriction point; they irreversibly commence DNA replication and centrosome duplication.

The G_1–S Phase Transition in Metazoans Involves Cyclin-CDK Control of the E2F Transcription Factor Through Its Regulator Rb

The sequence of events we described in yeast is similar in mammalian — and in fact all metazoan — cells. Growth factor signaling through the mitogen-activated protein kinase (MAPK) pathway triggers increased transcription of the G_1 cyclins, members of the cyclin D family. As long as levels of the Ink4 CKI family members are low, cyclin D binds to the G_1 CDKs, CDK4 and CDK6, to form active G_1 cyclin-CDK complexes. In turn, these G_1 CDKs activate members of a small family of related transcription factors, referred to collectively as **E2F transcription factors (E2Fs)**. During G_1, the E2Fs are maintained in an inactive state through tight association with the retinoblastoma protein (Rb; Figure 19-16b). When G_1 CDKs become activated, they phosphorylate Rb, causing it to dissociate from the E2Fs. Released E2Fs then activate genes

encoding many of the proteins involved in DNA synthesis. Thus the E2Fs have a function in late G_1 that is similar to that of the *S. cerevisiae* transcription factor complex SBF [compare (a) and (b) in Figure 19-16]. The E2Fs also stimulate transcription of genes encoding the G_1/S phase cyclin, cyclin E, and the S phase cyclin, cyclin A. Cyclin E forms a complex with the G_1/S CDK, CDK2, which then further phosphorylates Rb in a positive feedback loop, much like how the Cln1/2-CDK complex further phosphorylates Whi5 in yeast (Figure 19-16), to reinforce S phase entry.

As noted above, key to the regulation of E2F function is the Rb protein. When E2Fs are bound to Rb, not only are they transcriptionally inactive, but they also function to repress transcription. This is because Rb recruits chromatin-modifying enzymes that promote deacetylation and methylation of specific histone lysines, causing chromatin to assume a condensed, transcriptionally inactive form (see Section 8.5). *RB* was initially identified as the gene mutated in retinoblastoma, a childhood cancer of the retina. Subsequent studies found Rb to be inactivated in many cancers, either by mutations in both alleles of *RB* or by abnormal regulation of Rb phosphorylation (see Chapter 25).

Rb protein regulation by G_1 CDKs in mammalian cells is analogous to Whi5 regulation by Cln3-CDK in yeast. Phosphorylation on multiple sites by G_1 CDKs blocks Rb from associating with E2F and promotes its export out of the nucleus. This allows E2F to activate

the transcription of genes required for entry into S phase. Once the expression of genes coding for the G_1/S cyclins and CDK has been induced by phosphorylation of some of the Rb molecules, the resulting G_1/S phase CDK complexes further phosphorylate Rb in late G_1. This is one of the principal biochemical events responsible for passage through the restriction point (Figure 19-16b). Since E2F stimulates its own expression as well as that of the G_1/S cyclin-CDKs, cross-regulation of E2F and G_1/S cyclin-CDKs through a positive feedback loop produces a rapid rise of both activities in late G_1.

As they accumulate, S phase CDKs as well as mitotic CDKs maintain Rb protein in the phosphorylated state throughout the S, G_2, and early M phases. After cells complete anaphase and enter early G_1 or G_0, a fall in all cyclin-CDK activities leads to the de-phosphorylation of Rb. As a consequence, poorly phosphorylated Rb is available to inhibit E2F activity during early G_1 of the next cycle and in G_0-arrested cells. Thus G_1/S phase CDK activity remains low until cells decide to enter a new cell cycle and G_1 CDKs break the inhibitory grip of Rb on E2F.

Extracellular Signals Govern Cell Cycle Entry

Whether or not cells enter the cell cycle is influenced by extracellular as well as intracellular signals. Unicellular organisms such as yeasts, for example, enter the cell cycle only when they have reached an appropriate size, known as the **critical cell size**. This critical size, in turn, is controlled by nutrients available in the environment. Here we restrict our discussion to the fact that G_1 cyclin synthesis is responsive to the rate of protein synthesis, which is in turn controlled by pathways that are regulated by nutrients in the environment. This link between the macromolecule biosynthesis machinery and the cell cycle machinery is well understood in budding yeast. In this organism, the G_1 cyclin transcript *CLN3* contains a short upstream open reading frame that inhibits translation initiation when nutrients are limited. This inhibition is diminished when nutrients are in abundance. In the presence of sufficient nutrients, the TOR signaling pathway, which senses nutrients and growth factor signals, is active and stimulates translational activity (see Figure 21-3b). Since Cln3 is an unstable protein, its concentration fluctuates with the translation rate of its mRNA. Consequently, the amount and activity of Cln3-CDK complexes, which depend on the concentration of Cln3 protein, are regulated by nutrient levels.

In multicellular organisms, cells are surrounded by nutrients, and as such, nutrients do not usually limit the rate of cell proliferation. Rather, cell proliferation is controlled by the presence of growth-promoting factors (**mitogens**) and growth-inhibiting factors (*anti-*

mitogens) in the cell surroundings. As discussed in Chapter 16, addition of mitogens to G_0-arrested mammalian cells induces receptor tyrosine kinase–linked signal transduction pathways that initiate signal transduction cascades. These signaling cascades ultimately influence transcription and cell cycle control. They do so in multiple ways.

Mitogens activate the transcription of multiple genes. Most of these genes fall into one of two classes — *early response* or *delayed response* genes — depending on how soon their encoded mRNAs appear. Transcription of early response genes is induced within a few minutes after addition of growth factors by signal transduction cascades that activate preexisting transcription factors in the cytosol or nucleus (see Chapter 16). Many of the early response genes encode transcription factors, such as c-Fos and c-Jun, that stimulate transcription of the delayed response genes. The early response transcription factor Myc induces the transcription of G_1 cyclin and CDK genes. In addition to being controlled by transcription, G_1 CDKs are regulated by CKIs. The CKI $p15^{INK4b}$ is a potent CDK inhibitor. In some tissues, mitogens inhibit the production of this CKI by inhibiting its transcription.

In many tissues, cell proliferation is also regulated by anti-mitogens, which prevent entry into the cell cycle. Similarly, during differentiation, cells cease to divide and enter G_0. Some differentiated cells (e.g., fibroblasts and lymphocytes) can be

stimulated to re-enter the cell cycle and replicate. Many postmitotic differentiated cells, however, never re-enter the cell cycle to replicate again. Anti-mitogens and differentiation pathways prevent the accumulation of G_1 CDKs. They antagonize the production of G_1 cyclins and induce the production of CKIs. Transforming growth factor β (TGF-β) is an important anti-mitogen. This hormone induces a signaling cascade that brings about G_1 arrest by inducing the expression of $p15^{INK4b}$. As we will see in Chapter 25, the signaling pathways that regulate G_1 CDKs are often mutated in many human cancers.

Degradation of an S Phase CDK Inhibitor Triggers DNA Replication

Entry into S phase is defined by the unwinding of origins of DNA replication. G_1/S phase CDKs play an essential role in this process by turning off the machinery that degrades S phase cyclins during exit from mitosis and G_1. In addition, they induce the degradation of a CKI that inhibits S phase CDKs, initiating a rapid transition from low to high S phase CDK activity.

One of the important substrates of the G_1/S phase cyclin-CDK complexes is Cdh1. During late anaphase, Cdh1 directs APC/C to ubiquitinylate substrate proteins, including S phase and mitotic cyclins, marking them for proteolysis by proteasomes (see Figure 19-

14d). This ensures that when cells exit mitosis, they arrive into a G_1 state of low cyclin-CDK activity, allowing them to assess their internal and external environments prior to committing to another round of cell division. The APC/C^{Cdh1} complex remains active throughout G_1, preventing the premature accumulation of S phase and mitotic cyclins. If conditions for another round of cell division are favorable, then a G_1/S cyclin-CDK induced signaling pathway results in phosphorylation of Cdh1, causing it to dissociate from the APC/C complex, inhibiting further ubiquitinylation of S phase and mitotic cyclins during late G_1 and allowing them to progressively accumulate (Figure 19-17). This block in cyclin degradation, combined with the induced transcription of S phase cyclins during late G_1, allows S phase cyclins to accumulate as G_1/S cyclin-CDK levels rise. Later in the cell cycle, S phase and mitotic CDKs take over to maintain Cdh1 in the phosphorylated, and hence inactive, state. Only as mitotic CDKs decline and a protein phosphatase in yeast known as *Cdc14* becomes activated are these inhibitory phosphates removed from Cdh1, leading to its reactivation, again targeting mitotic proteins including M and S phase cyclins for degradation. In mammalian cells, similar mechanisms are responsible for stabilizing S phase and mitotic cyclins, but the phosphatase(s) involved in de-phosphorylation of Cdh1 have not been identified.

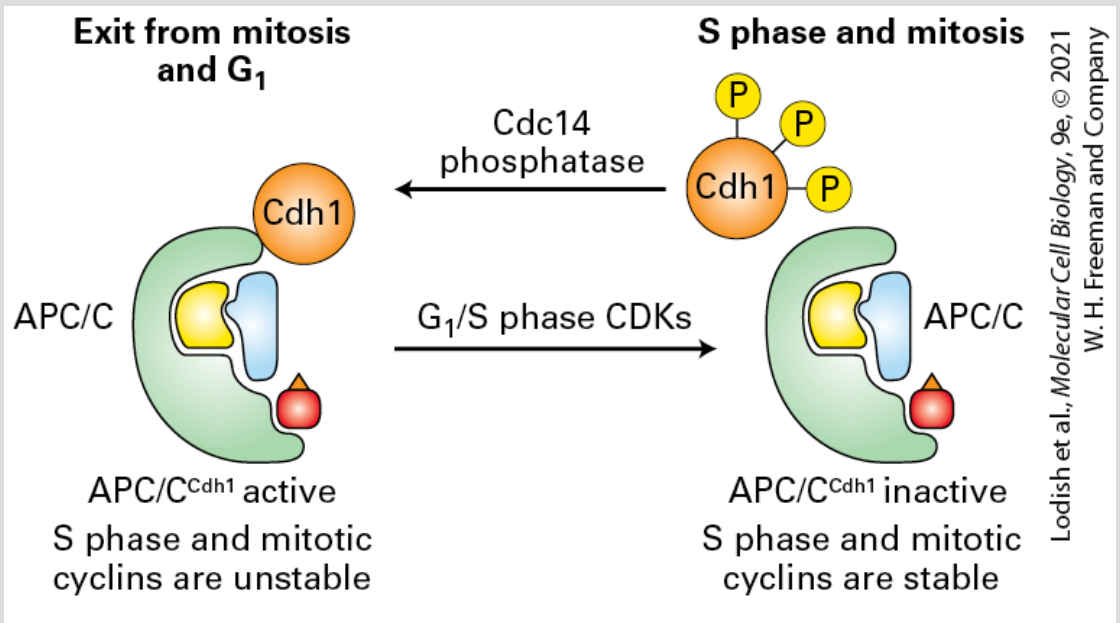


FIGURE 19-17 Regulation of S phase and mitotic cyclin levels in budding yeast. In late anaphase, the anaphase-promoting complex (APC/C) ubiquitinylates S phase and mitotic cyclins. The activity of this ubiquitin-protein ligase is directed toward mitotic cyclins by a specificity factor called *Cdh1*. Cdh1 activity is regulated by phosphorylation. During exit from mitosis and G_1, Cdh1 is de-phosphorylated and APC/C^{Cdh1} is active. During S phase and mitosis, Cdh1 is phosphorylated, Cdh1 dissociates from APC/C, and APC/C becomes inactive. The G_1/S phase CDKs, which themselves are not APC/C^{Cdh1} substrates, phosphorylate Cdh1 at the G_1–S phase transition. A specific yeast phosphatase called *Cdc14* removes the regulatory phosphate from the specificity factor late in anaphase.

In *S. cerevisiae*, as S phase cyclin-CDK accumulates in late G_1, they are immediately inactivated by binding of a CKI called *Sic1* that is expressed late in mitosis and early in G_1 (Figure 19-18a). Because Sic1 specifically inhibits S phase and M phase CDK complexes but has no effect on the G_1 CDK and G_1/S phase CDK complexes, it functions as an S phase inhibitor. Initiation of DNA replication occurs when the Sic1 inhibitor is precipitously degraded following

its ubiquitinylation by the SCF ubiquitin-protein ligase (see Figure 19-14a).

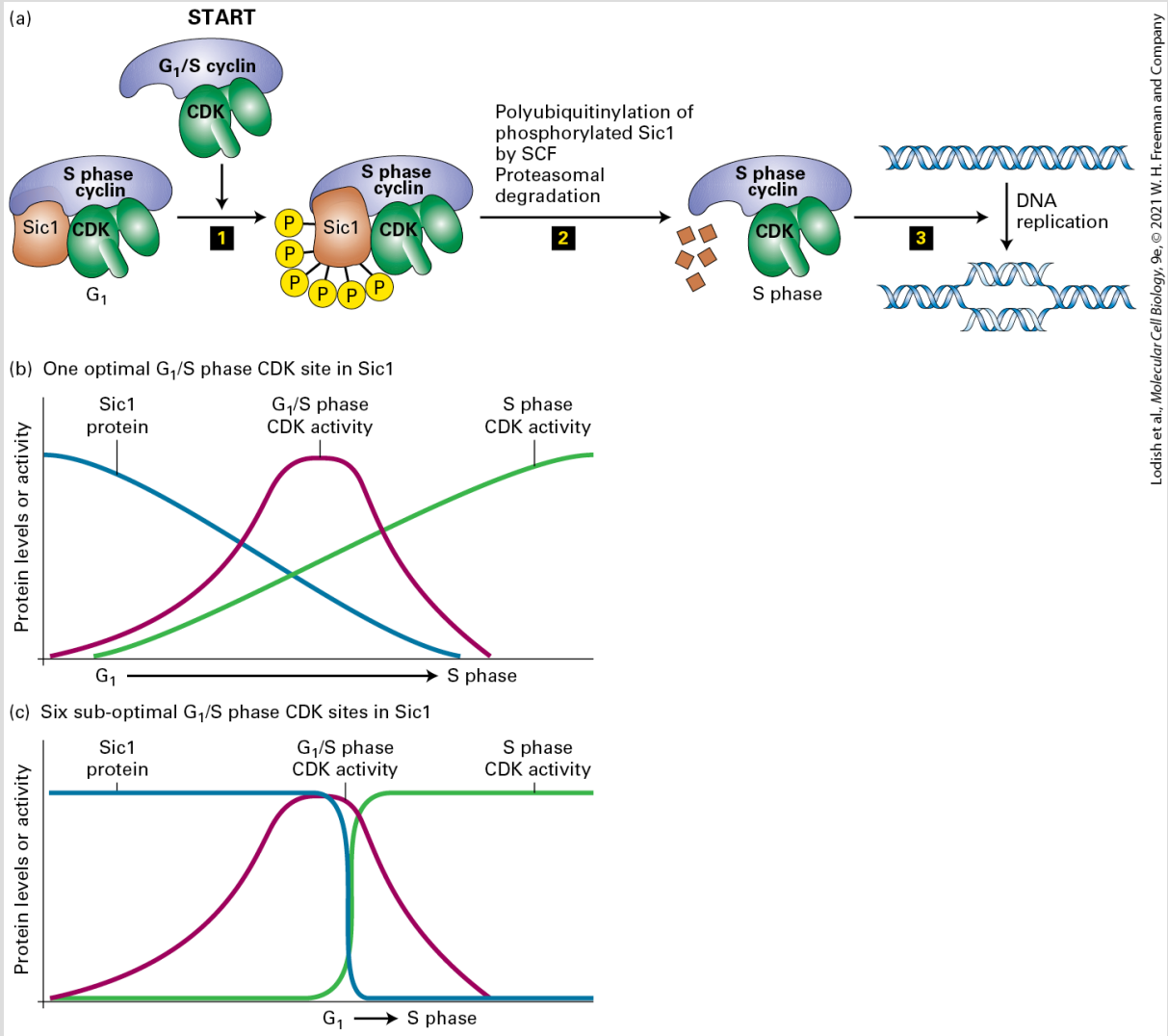


FIGURE 19-18 Switch-like control of S phase onset in *S. cerevisiae* by phospho-regulated proteolysis of the S phase inhibitor Sic1. (a) The S phase cyclin-CDK complexes begin to accumulate in G_1 but are inhibited by Sic1. This inhibition prevents initiation of DNA replication until the cells have completed all G_1 events. G_1/S phase CDKs assembled in late G_1 phosphorylate Sic1 at multiple sites (step **1**), marking it for phospho-dependent binding and ubiquitinylation by the SCF ubiquitin-protein ligase and subsequent proteasomal degradation (step **2**). The active S phase CDKs then trigger initiation of DNA synthesis (step **3**) by phosphorylating and recruiting MCM helicase activators to DNA

replication origins (not shown). (b) A single optimal G_1/S phase CDK phosphorylation site in Sic1 that was strongly recognized by the phospho-binding F-box in SCF ubiquitin-protein ligase would result in a sluggish G_1–S phase transition. As G_1/S phase CDKs accumulate during G_1, Sic1 would be progressively degraded. As a result, S phase CDKs would slowly rise, and initiation of S phase would be a drawn-out event. (c) Six suboptimal phosphorylation sites in Sic1 create a switch-like cell cycle entry. Only when all of the suboptimal sites in Sic1 are fully phosphorylated is the protein recognized by the phospho-binding F box in SCF. This can only occur when G_1/S phase CDKs have reached high levels. This ensures that Sic1 degradation occurs rapidly, and only when G_1/S phase CDKs have accomplished all their other G_1 tasks. See P. Nash et al., 2001, *Nature* **414**:514–521; and D. O. Morgan, 2006.

Degradation of Sic1 is induced by its phosphorylation by G_1/S phase CDKs (Figure 19-18a). It must be phosphorylated at no fewer than six sites, each of which is a relatively poor substrate for the G_1/S phase CDKs, before it is bound sufficiently well by SCF to be ubiquitinylated. In contrast, if Sic1 were inactivated following the phosphorylation of a single site, Sic1 molecules would start to get phosphorylated and degraded as soon as the levels of G_1/S phase CDK activity begin to rise, leading to a gradual decrease in Sic1 levels (Figure 19-18b) and poor control over S phase entry. Because six sites need to be phosphorylated, at low levels of G_1/S phase CDK activity Sic1 is only partially phosphorylated and Sic1 is not destroyed. Only when G_1/S phase CDK levels are high is Sic1 sufficiently phosphorylated to target it for degradation. Having multiple suboptimal phosphorylation sites therefore leads to an ultrasensitive, switch-like response in Sic1 degradation and hence precipitous activation of S phase CDKs (Figure 19-18c). Sic1

degradation occurs only when G_1/S phase CDK activity has reached its peak and virtually all the other G_1/S phase CDK substrates have been phosphorylated.

An obvious advantage of proteolysis for controlling passage through this critical point in the cell cycle is that protein degradation is an *irreversible process*, ensuring that cells proceed in one direction through the cycle.

Entry into S phase in metazoan cells is regulated by a switch-like mechanism similar to that in budding yeast. Like Sic1, the CKI p27 binds to and prevents the premature activation of S phase CDKs during G_1. Unlike Sic1, however, this CKI inhibits both S phase CDKs and G_1/S phase CDKs. Like Sic1 in yeast, p27 is removed from cyclin-CDK complexes by ubiquitin-dependent protein degradation and a pathway, analogous to the one operating on Sic1, targets p27 for degradation at the G_1–S phase transition. As G_1/S phase CDKs and S phase CDKs reach high levels during late G_1 and early S phase, they begin to phosphorylate p27, targeting it for ubiquitinylation by SCF^{SKP2}. Degradation of p27 causes activation of G_1/S phase and S phase CDKs. These kinases then initiate S phase by phosphorylating proteins important for the initiation of DNA replication.

Replication at Each Origin Is Initiated Once and Only Once During the Cell

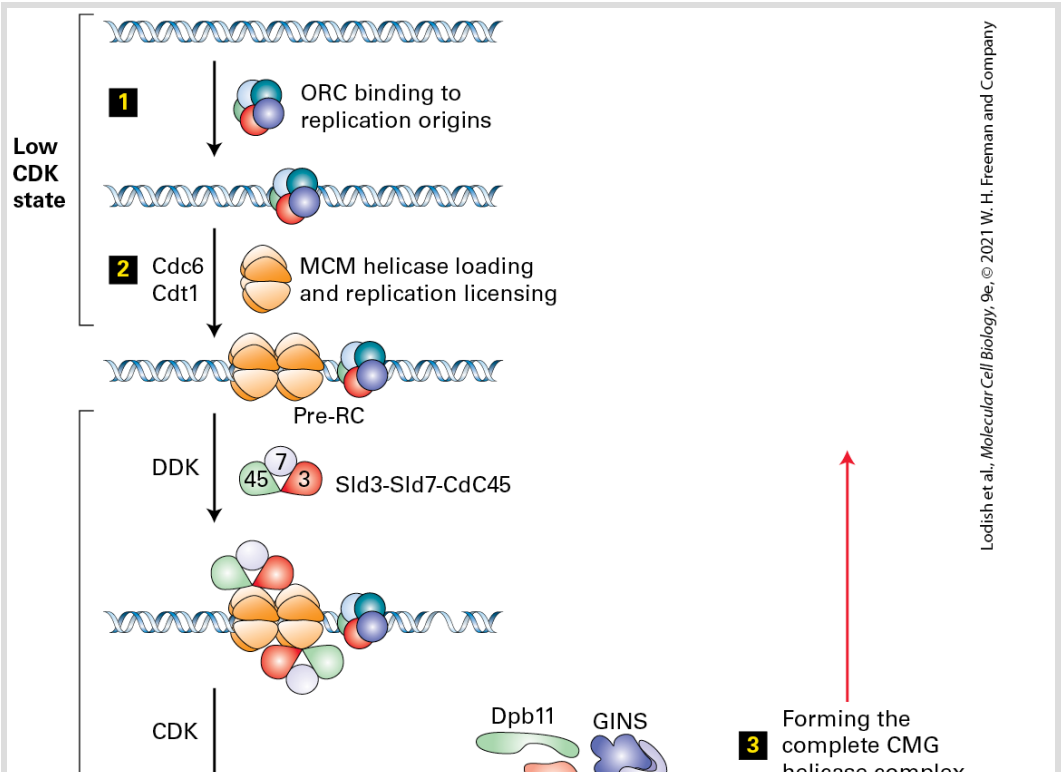
Cycle

As discussed in Chapter 5, eukaryotic chromosomes are replicated from multiple replication origins, specific sites on the chromosomes where the replication machinery is assembled, and where the copying of the two strands of DNA begins. Replication of all of the chromosomal DNA — starting from these origins — occurs during S phase. Each eukaryotic origin fires only once per S phase to initiate replication from that site, and S phase continues until replication from each of the multiple origins that lie along the length of each chromosome has resulted in complete replication of the entire chromosome. The DNA that is copied from a single origin is called a *replicon*. This process of creating and fusing the replicons ensures that each DNA strand is copied only once, maintaining the correct gene copy number each time a cell replicates.

S phase CDKs play an essential role regulating DNA replication. As we shall see in more detail shortly, these kinases allow DNA replication to be initiated only at the G_1–S phase transition and prevent re-initiation from origins that have already fired. We first discuss how initiation of DNA replication is controlled, and the role of S phase CDKs in the process, before turning to the mechanisms these kinases use to prevent re-initiation.

The powerful genetic methods available for experimental manipulation of budding yeast have led to the identification of specific mutants defective in DNA replication. Consequently, we

have a fairly comprehensive understanding of the mechanisms by which yeast initiates and implements DNA replication. Although many details differ between yeast and mammalian cells, the basic processes are essentially the same, so we will focus our discussion first on how DNA replication is accomplished in budding yeast. It is important to remember that the five basic steps of genome replication are essentially the same in all eukaryotic species (Figure 19-19). The steps are: (1) recognizing the replication origins; (2) assembling replication initiation factors that then load a replicative helicase in order to form a so-called prereplication complexes, or pre-RCs; (3) recruiting critical helicase cofactors; (4) activation of the helicase to unwind the two DNA strands; and (5) recruitment of the replicative DNA polymerase machinery (i.e., the "replisome") that then copies each strand.



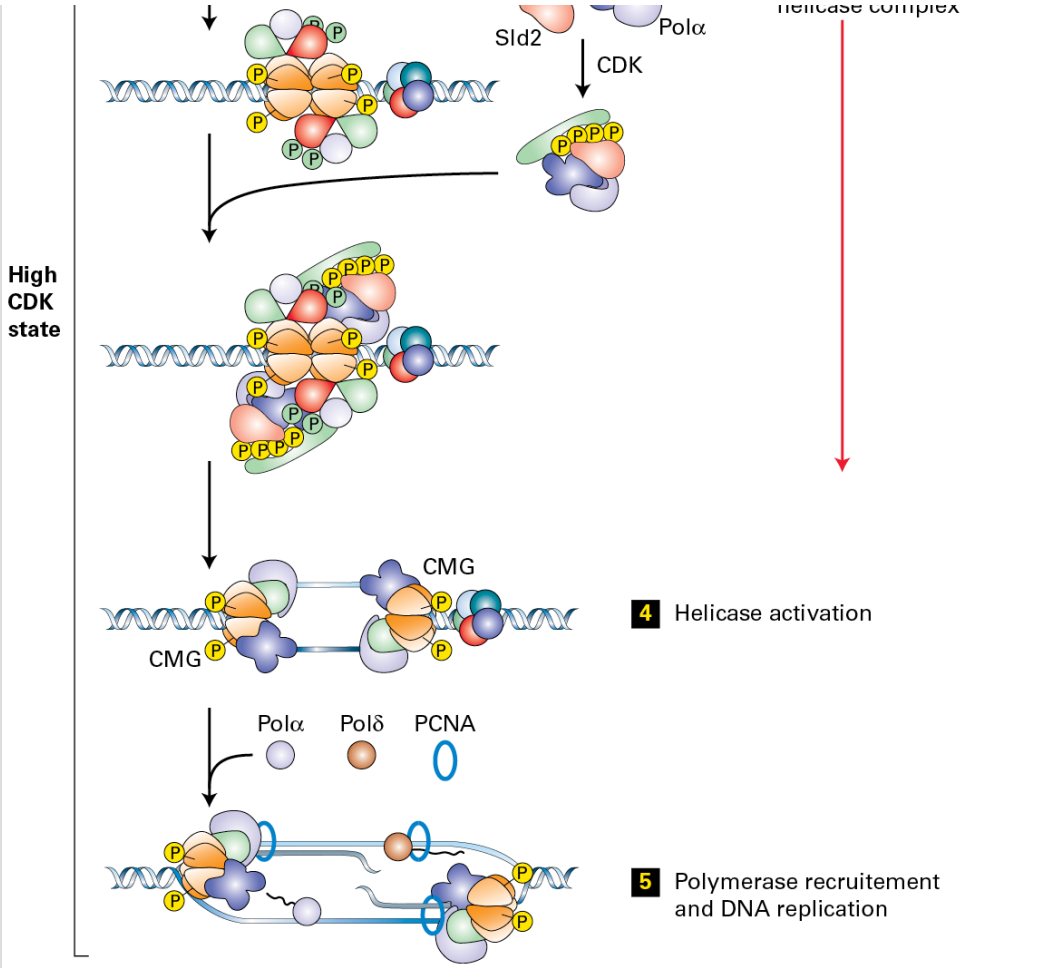


FIGURE 19-19 The molecular mechanisms governing the initiation of DNA replication. Step 1: During exit from mitosis and early G_1, when CDK activity is low, ORC recognizes and binds to replication origins. In step 2, the MCM loading factors Cdc6, and Cdt1 load the inactive replicative helicase, the MCM complex, onto DNA at replication origins, forming the pre-RC complex. Step 3: Activation of S phase CDKs and DDK marks the onset of S phase. They phosphorylate the MCM helicase, Sld2, and Sld3 (depicted as yellow phosphorylation events), to facilitate the loading of MCM helicase activators — the Cdc45 and GINS complexes — onto sites of replication initiation. Step 4: Loading of these activators leads MCM helicases to unwind DNA. S phase CDKs also prevent reloading of MCM helicases by phosphorylating Cdc6 and Cdt1 promoting their release from the replication origins and their nuclear export or degradation by ubiquitin ligases. S phase CDKs also phosphorylate MCM helicases (depicted as orange phosphorylation events), which leads to their export from the nucleus when the helicases disengage from the DNA when

replication is complete. Step **5**: The additional DNA polymerases Pol α (the primase) and Pol δ, along with the processivity factor PCNA, are recruited to origins, which leads to the initiation of DNA replication on both the leading and lagging strands (see Figure 5-30).

[Data from Y. Li and H. Araki, 2013, *Genes Cells* **18**:266–277.]

Even though DNA replication will not occur until S phase, the first 2 steps actually occur in G_1 phase, when the activity of G_1 and S phase CDKs is low. As shown in Figure 19-19, in step **1**, a complex of six proteins known as the *origin-recognition complex (ORC)* binds to all of the DNA replication origins. In budding yeast, an ORC binds to an 11-bp conserved core sequence found in every replication origin. The situation is more complicated in multicellular organisms because DNA replication origins lack any recognizable consensus sequence. Instead, it is thought that additional chromatin-associated factors target ORC to the DNA at replication origins. Once bound to DNA, ORC is thought to function as a landing pad for the assembly of additional factors necessary for DNA replication to commence. During step **2**, two replication initiation factors, Cdc6 and Cdt1, bind to ORC at the origins and then proceed to load another complex of six proteins, known as the *MCM helicase* (for *m*ini *c*hromosome *m*aintenance), onto the DNA to form a pre-RC. The MCM complexes appear to be loaded as double hexamers facing in opposite directions, allowing DNA unwinding to progress in both directions from the origins. Once loaded onto the origin DNA in G_1 phase, the helicases remain inactive until the cell has entered S phase. This restriction on activity is conferred by CDKs, which regulate two

opposing phosphorylation states. The MCM helicases can only be loaded onto the DNA in a state of low CDK activity that occurs when CDKs are inactivated during exit from mitosis and during early G_1. In other words, MCM helicases are loaded onto DNA to form pre-RCs only when they are nonphosphorylated. In contrast to DNA loading, activation of the MCM helicases and the recruitment of DNA polymerases to the unwound origin DNA can only occur when the MCM helicase proteins are phosphorylated. This happens only in S phase, when they are phosphorylated by S phase CDKs and a second protein kinase called *DDK* (for *D*bf4-*d*ependent *k*inase).

Just like CDKs, which require cyclin-binding for activity, the monomeric DDK kinase subunit (also known as *Cdc7*) is inactive in its monomeric form. Only when DDK binds to the Dbf4 subunit does it acquire kinase catalytic activity. During G_1, the Dbf4 protein is continuously targeted for degradation by the APC/C. Only during S phase, when the APC/C is inactivated by CDK phosphorylation, does Dbf4 accumulate and activate DDK. (Recall that S phase CDKs become active only when G_1/S phase CDK levels reach their peak, allowing the CKIs of S phase CDKs to be destroyed.) Putting all of this phosphorylation and ubiquitination-mediated control together, the key concept that emerges is that the increase in CDK and DDK activity during S phase enforces the strict temporal separation of pre-RC formation (i.e., loading of the MCM helicases), which only happens during G_1, from firing of the origins to start DNA replication, which only happens during S phase. Formation of the

pre-RCs is sometimes referred to as *replication licensing*, meaning that the MCM-loaded origins are now licensed to fire later in S phase. By only licensing origins during G_1 (Figure 19-19, steps **1** and **2**), and only firing origins during S phase (Figure 19-19, steps **3**–**5**), the cell ensures that each origin only triggers DNA replication once with each passage through the cell cycle, thereby preventing inappropriate genomic re-replication.

So how do S phase CDKs and DDK collaborate at the mechanistic level to initiate DNA replication and prevent re-replication? The process is rather complicated, and the exact order of events differs somewhat between yeast and metazoans, but the basic concept, which is the same in all cells, is to recruit helicase cofactors (step **3**) that will activate the MCM helicases (step **4**) at the same time that the replicative polymerase machinery (the replisome) is loaded onto the DNA (step **5**). This coordination prevents excessive unwinding of DNA, which would result in the accumulation of large amounts of single-stranded DNA. The activation process involves binding of inactive MCM helicase to two other protein complexes, one containing Cdc45 and the other containing GINS, in order to form the active Cdc45-MCM-GINS (CMG) helicase complex. Formation of the CMG complex requires phosphorylation events catalyzed by both DDK and S phase CDKs, which are activated in late G_1. In yeast it happens like this: DDK phosphorylates the N-termini of several subunits of the MCM helicase, altering their conformation, and leading to the recruitment of one of the two

critical MCM helicase cofactors, a protein known as *Cdc45*. Cdc45 is required for MCM helicase activity, and binds to the MCM helicase as part of a complex with two other proteins Sld3 and Sld7 (Figure 19-19, step **3**). Sld3 recruits another protein complex containing Pol ε (the replicative polymerase that synthesizes the leading strand), forming a fully active CMG helicase. The origins unwind (Figure 19-19, step **4**), Pol α (the primase), Pol δ (the replicative polymerase that synthesizes the lagging strand), and PCNA (the DNA sliding clamp that holds the polymerases anchored to the DNA to increase processivity) then bind, and replication ensues (Figure 19-19, step **5**).

S phase CDKs are not only essential for initiating DNA replication but also are responsible for ensuring that each origin fires only once during S phase. Refiring of origins during S phase is prevented by phosphorylation of several components of the MCM helicase-loading machinery and the MCM helicase complex itself. To distinguish these phosphorylation events from the ones required for the initiation of DNA replication, they are depicted in orange in Figure 19-19. Concomitant with activation of the MCM helicases, Cdc6 and Cdt1 dissociate from the sites of DNA replication initiation. Once they have done so, their phosphorylation leads to export of Cdt1 from the nucleus, and degradation of Cdc6 by the SCF ubiquitin-protein ligase in yeast. Phosphorylation of MCM helicases leads to the export of these proteins from the nucleus after they dissociate from the DNA on completion of DNA replication. Thus

only after CDK activity is lowered by APC/C^{Cdh1} during exit from mitosis can MCM helicases be reloaded onto DNA. As a result, helicase loading is restricted to late stages of mitosis and early G_1 (see Figure 19-19, steps **1** and **2**).

The general mechanisms governing the initiation of DNA replication in metazoan cells parallel those in *S. cerevisiae*, although some differences are found in vertebrates. The helicases are loaded in G_1, when CDK activity is low. Phosphorylation of MCM helicase activators by G_1/S phase CDKs and S phase CDKs activates the helicases and promotes polymerase loading. As in yeast, phosphorylation of the MCM helicase-loading factors Cdc6 and Cdt1 prevents reloading of MCM helicases until the cell passes through mitosis, thereby ensuring that replication from each origin occurs only once during each cell cycle. In addition, vertebrates contain a small protein, geminin, that also contributes to the inhibition of re-initiation at origins until cells complete a full cell cycle. Geminin is expressed in late G_1; it binds to and inhibits Cdt1 (see Figure 19-19, step **2**), thereby preventing reloading of the MCM helicase that is released from origins after DNA replication is initiated during S phase. Geminin contains a destruction box at its N-terminus that is recognized by APC/C^{Cdh1}, causing it to be ubiquitinylated in late anaphase and degraded by proteasomes. Its degradation frees the MCM helicase-loading factors, which are also de-phosphorylated as CDK activity declines, to bind to ORC on replication origins and load MCM helicases during the following G_1 phase. During S phase, the

combined ubiquitin-mediated degradation of Cdt1 and the inactivation of any residual Cdt1 left over by geminin provides redundant regulation to ensure that there is no accidental chromosome re-replication.

Duplicated DNA Strands Become Linked During Replication

It is essential that the sister chromatids generated during S phase remain together until they separate during the metaphase-to-anaphase transition of mitosis. As chromosomes are duplicated to form sister chromatids, they become tethered to each other by protein links. The protein ring complexes that establish these linkages between sister chromatids are called **cohesins**. They are composed of four subunits: Smc1, Smc3, Scc1, and Scc3 (Figure 19-20a). Smc1 and Smc3 are members of the structural maintenance of chromosomes (SMC) protein family, which is characterized by globular domains at the N- and C-termini connected by a long α helix. The helix is broken in the middle by a globular hinge domain, so that the two halves of the α helix in each protein fold back on themselves to form a coiled-coil structure. The globular head domains bind and hydrolyze ATP. The ATPase domains interact with the proteins Scc1 and Scc3, which together form a ring structure. These cohesin rings embrace one or both copies of the replicated

DNA. When cohesins are experimentally inactivated, sister chromatids do not associate properly with each other.

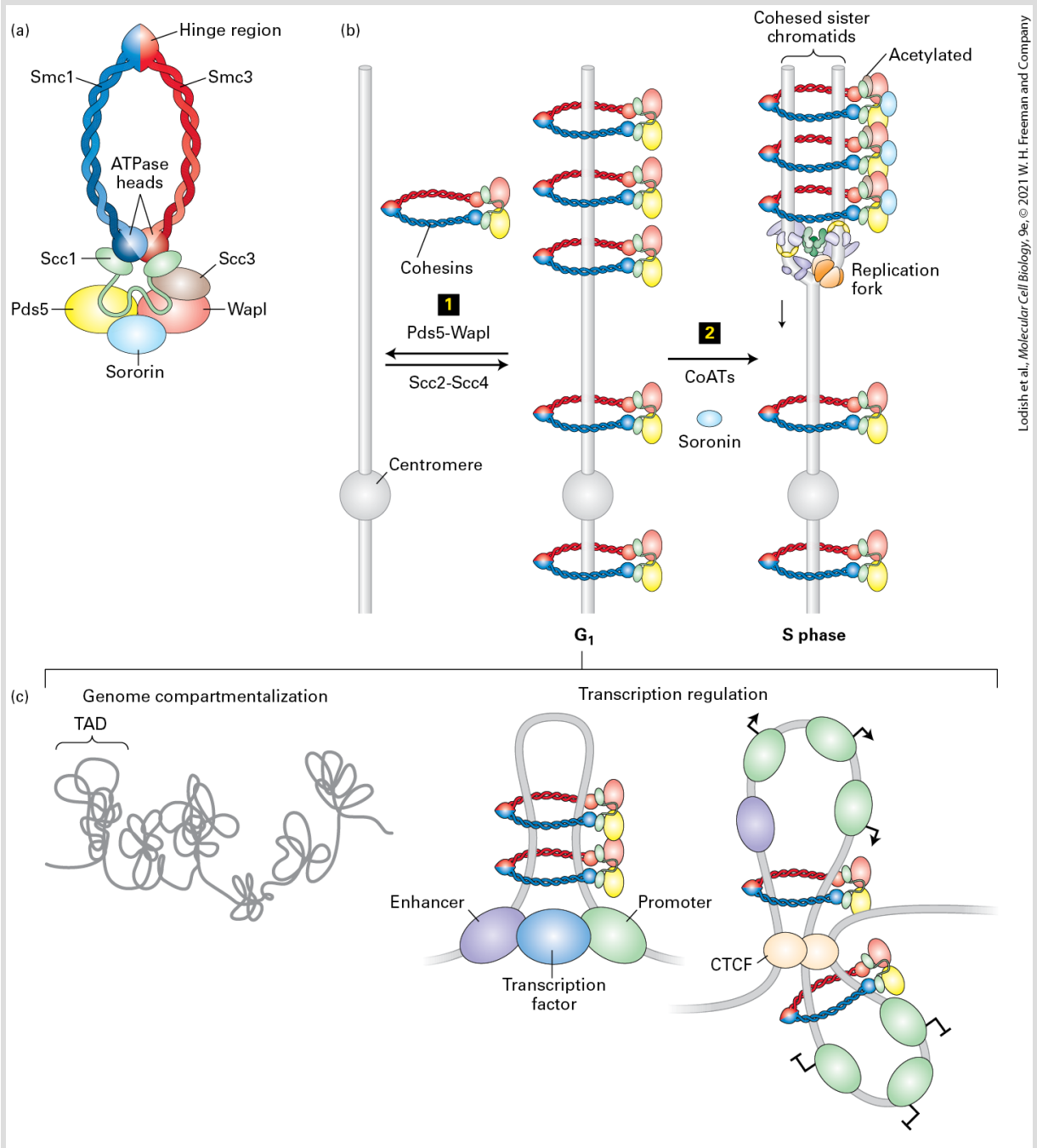


FIGURE 19-20 Model for establishment of cohesin linkage of sister chromatids. Cohesin complexes form rings that link sister chromatids by embracing the two sister DNA molecules. (a) Schematic structure of the yeast cohesin complex. The structure of vertebrate

cohesion is essentially the same, though some of the proteins have different names. (b) Mechanism whereby cohesins are loaded onto DNA and acquire their cohesive properties. Step **1**: Cohesins are loaded onto chromosomes during G_1 by the cohesin-loading complex Scc2-Scc4, but the cohesins do not possess cohesive (gluelike) properties. In this state, cohesins are dynamic and can slide along the DNA or dissociate from the DNA with the help of the Pds5-Wapl complex, which associates with cohesins. Step **2**: Concomitant with DNA replication, closely behind the replication fork cohesins are converted into cohesive molecules, able to glue sister chromatids together through acetylation of Smc3 by cohesin acetyltransferases (CoATs). Acetylation, indicated by brown circles on SMC3, is accompanied in vertebrates by the binding of sororin to cohesin, which helps stabilize cohesins on chromosomes. In budding yeast, both sister chromatids are embraced within a single cohesin ring, as shown. (c) Cohesion-independent functions of cohesins during G_1 include partitioning of the genomic DNA into discrete compartments of the nucleus called *topologically associated domains (TADs)* and controlling gene expression by bringing enhancers close to promoters and transcription start sites through chromosomal DNA looping. In combination with CCCTC-binding factor (CTCF), cohesin-looping can activate some genes in a cluster, shown by arrows, while repressing others, indicated by the T symbol.

[Part (c) Information from A. Losada, 2014, *Nat. Rev. Cancer* **14**:389–393.]

Cohesin-mediated cohesion between sister chromatids is established by a two-step process and is tightly tied to DNA replication. In the first step, which occurs during G_1, cohesins associate with chromosomes, aided by cohesin loading factors Scc2 and Scc4 (Figure 19-20b, step **1**). In G_1, the chromosome-cohesin interaction is quite dynamic. In addition to Scc2-Scc4–mediated binding, cohesins are unloaded continuously by a cohesin-associated complex composed of the Pds5 and Wapl proteins. This continual loading and unloading of cohesins is likely to be

important in regulating interphase chromatin structure and gene expression (Figure 19-20c).

Cohesins acquire their cohesive properties during DNA replication. In the second step, the two duplicated DNA strands become entrapped within the cohesin rings as replication forks replicate the DNA (Figure 19-20b, step **2**). Converting DNA-bound G_1 cohesins into cohesive complexes requires acetylation of the Smc3 subunit by cohesin acetyl transferases (CoATs). This acetylation prevents unloading of cohesins by Pds5-Wapl, thereby stabilizing the cohesins on chromosomes. In vertebrates, stabilization of cohesins requires the cohesin-associated factor sororin (Figure 19-20a and b, step **2**). As we will see in Section 19.6, cohesins are essential for accurate attachment of the replicated sister chromatids to the mitotic spindle and for their segregation during mitosis.

Cohesins also regulate the topological localization of chromosomes within the interphase cell nucleus and help control gene expression (Figure 19-20c). In some instances, cohesins promote gene expression, whereas in others they restrain it. The mechanism cohesins use to accomplish gene expression control is, however, the same in both cases: cohesins promote chromatin loop formation, which brings enhancer or repressor elements close to the promoter and transcriptional start site. Defects of cohesin's gene expression regulation function are the cause of a group of diseases collectively called *cohesinopathies*. In these diseases, mutations in

cohesin subunits or cohesin-loading factors disrupt the expression of genes that are critical for development, causing limb and craniofacial abnormalities and intellectual disabilities. Cohesin's sister chromatid cohesion function, however, is intact in these diseases. In contrast, as we will see in Section 19.8, defects in cohesin's cohesion function during meiosis cause miscarriages and intellectual disabilities. ■

KEY CONCEPTS OF SECTION 19.4

The Transition from G_1 into S Phase and DNA Replication

- The molecular events promoting entry into the cell cycle are conserved across species. G_1 CDKs phosphorylate and inhibit a transcriptional repressor. This permits transcription of G_1/S phase cyclin genes and other genes important for S phase (see Figure 19-16).
- Extracellular signals such as nutritional state (in yeast) and the presence of mitogens and anti-mitogens (in vertebrates) regulate entry into the cell cycle.
- Various polypeptide growth factors called *mitogens* stimulate cultured mammalian cells to proliferate by inducing expression of early response genes. Many of these genes encode transcription factors that stimulate expression of genes encoding the G_1/S phase cyclins and E2F transcription factors.
- The G_1/S phase CDKs phosphorylate and inhibit Cdh1, the factor that directs the anaphase-promoting complex (APC/C) to ubiquitinylate S phase and M phase cyclins. This allows S phase cyclins to accumulate in late G_1 (see Figure 19-17).
- In yeast, S phase CDKs are initially inhibited by Sic1. A series of phosphorylation sites marks Sic1 for binding to a phosphoserine/threonine-binding domain on an SCF ubiquitin-protein ligase, leading to its ubiquitinylation and proteosomal degradation. This releases activated S phase CDKs that trigger an abrupt step-like onset of the S phase (see Figure 19-18). A similar phospho-dependent degradation of Cip/Kip CKI proteins functions in a similar manner to control S phase entry in animal cells.
- DNA replication is initiated from helicase loading sites known as *replication origins*.
- Loading and activation of MCM helicases occur in mutually exclusive cell cycle states: MCM helicase loading can occur only when CDK activity is low (during early G_1); MCM

helicases are activated when CDK activity is high (during S phase).

- S phase CDKs and DDK trigger the initiation of DNA replication by recruiting MCM helicase activators to origins (see Figure 19-19).
- Initiation of DNA replication occurs at each origin only once during the cell cycle because S phase CDKs activate the helicases and at the same time prevent additional helicases from loading onto DNA.
- Cohesins establish linkages between the replicated DNA molecules (see Figure 19-20), which are essential for their accurate segregation later in the cell cycle. This linking mechanism is coupled to DNA replication.

19.5 The G_2/M Transition and the Irreversible Engine of Mitosis

Once S phase has been completed and the entire genome has been duplicated, the pairs of duplicated DNA chromosomes — the sister chromatids — are segregated to the future daughter cells. This process requires not only the formation of the apparatus that facilitates this segregation — the mitotic spindle — but essentially a complete remodeling of the cell. Chromosomes condense and attach to the mitotic spindle, the nuclear envelope is disassembled, and almost all organelles are rebuilt or modified. These events are triggered by mitotic CDKs and a handful of other key mitotic kinases, particularly the Aurora kinases and the Polo-like kinases. This section first discusses how the mitotic CDKs are precipitously activated during G_2, following the completion of DNA replication, through a series of positive feedback loops involving the CDKs themselves, together with Aurora kinase A and Polo-like kinase 1 (Plk1). We then describe how these kinases bring about the dramatic changes in the cell necessary to facilitate sister chromatid segregation during anaphase, focusing on the events as they occur in metazoans.

Precipitous Activation of Mitotic CDKs by Positive Feedback Loops Initiates

Mitosis

Mitotic cyclin-CDKs initiate mitosis. Whereas levels of the catalytic CDK1 subunit (the mitotic CDK) are constant throughout the cell cycle, mitotic cyclins gradually accumulate during S phase. Most eukaryotes contain multiple mitotic cyclins, which for historical reasons are subdivided into the cyclin A and cyclin B families. As they assemble, mitotic cyclin-CDK complexes are maintained in an inactive state through inhibitory phosphorylation of the CDK1 subunit. Recall from Section 19.3 that two highly conserved tyrosine and threonine residues in mammalian CDKs are subject to regulated phosphorylation. In CDK1, phosphorylation of threonine 14 and tyrosine 15 maintains mitotic cyclin-CDK complexes in an inactivate state (see Figure 19-12). The phosphorylation state of T14 and Y15 is controlled by a dual-specificity protein kinase known as **Wee1** and a pair of dual-specificity phosphatases, **Cdc25B** and **Cdc25C**). Dual-specificity kinases and phosphatases can phosphorylate and de-phosphorylate serines, threonines, and tyrosines. Their regulation of mitotic CDK1 underlies the abrupt activation of CDK1 kinase activity at the G_2/M phase transition and explains the observation that although mitotic cyclins gradually accumulate during S phase and G_2, mitotic CDKs are not active until cells enter mitosis.

The critical roles of Wee1 and Cdc25 in controlling the entry of G_2 cells into mitosis emerged first from studies in the fission yeast *Schizosaccharomyces pombe*. Recall from Section 19.2 that fission

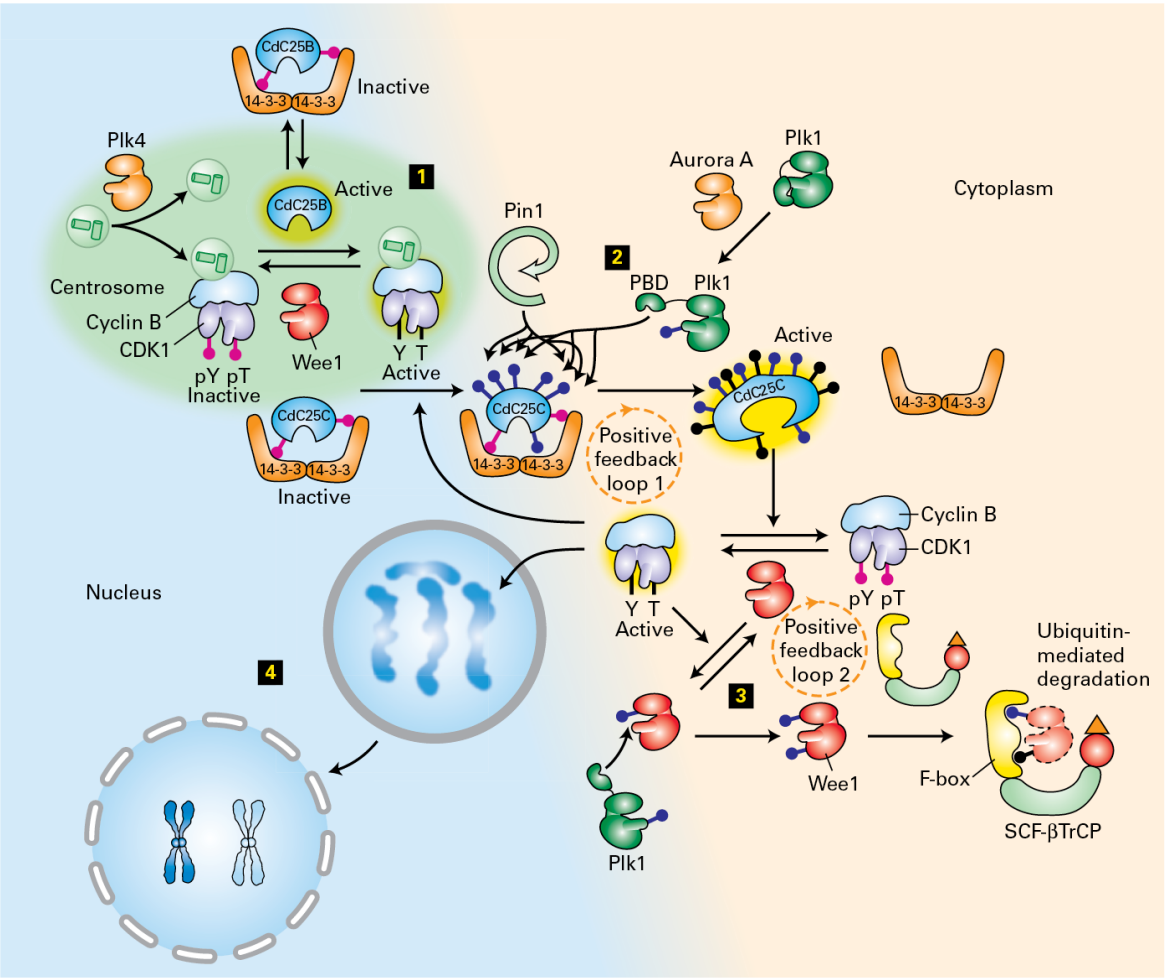
yeast coordinate cell size and cell division during G_2 (see Figure 19-4). This is in contrast to budding yeast and vertebrate cells, which coordinate cell size and cell division during G_1 and therefore spend most of the cell cycle in this phase. This means that the length of fission yeast at the time they divide is a direct reflection of the duration of G_2. In *S. pombe* the dual-specificity protein kinase Wee1 phosphorylates CDKs on the inhibitory tyrosine 15 (see Figure 19-13). (Threonine 14 is not phosphorylated in *S. pombe* CDK1.) Yeast cells with a defective $wee1^+$ gene activate mitotic CDKs prematurely and hence experience premature entry into mitosis. These wee1 mutants are therefore smaller than normal cells. In fact, the Scottish term for small is "wee," hence the name for this gene that, when mutated, resulted in small cells. Consistent with the importance of Wee1 in phosphorylating Tyr-15 in CDK1, fission yeast cells carrying a CDK1 mutation in which the tyrosine-15 residue is replaced by phenylalanine (which is structurally similar to tyrosine but cannot be phosphorylated) show the same premature mitotic CDK activation and entry into mitosis. In contrast, fission yeast cells carrying mutations in the *cdc25* gene arrest in G_2, giving rise to very long cells. This indicates that this phosphatase, which opposes Wee1 activity, is essential for entry into mitosis.

Vertebrates contain the same Wee1 protein kinase as *S. pombe* but have multiple Cdc25 phosphatases. Two family members, Cdc25B and Cdc25C, remove the inhibitory phosphorylation on CDK1 during G_2. Cdc25B is thought to act as a *starter* phosphatase that initiates

CDK1 activation, while Cdc25C is the *workhorse* phosphatase that activates the vast majority of cyclin-CDK1 complexes.

Abrupt activation of mitotic CDKs at the G_2/M transition is the consequence of rapid inactivation of Wee1 together with activation of Cdc25. Central to this rapid transition are two positive feedback loops built from CDKs, Polo-like kinase 1 (Plk1), and phosphoserine/threonine-binding domains, in which mitotic CDKs and Plk1 simultaneously activate Cdc25 and inactivate Wee1 (Figure 19-21). The process begins with Cdc25B, the starter phosphatase, activating an initial small pool of cyclin B-CDK1 at the centrosome (Figure 19-21, step **1**), which is thought to facilitate centrosome maturation. This initial pool of active cyclin B-CDK1 that is generated then phosphorylates Cdc25C, the workhorse phosphatase, in the cytoplasm. This recruits Pin1 and Plk1, which recognize Cdc25C through their phospho-binding domains. These proteins further modify and phosphorylate Cdc25C to stimulate its phosphatase activity (Figure 19-21, step **2**). This pool of active Cdc25C that is created, in turn activates the large pool of cytosolic cyclin B-CDK1, which activates more cytosolic Cdc25C as part of positive feedback loop 1. At the same time that Cdc25C is being activated, mitotic CDKs and Plk1 phosphorylate Wee1, which targets it for destruction through a phospho-binding SCF protein-ubiquitin ligase. This forms positive feedback loop 2 (Figure 19-21, step **3**). This combination of two positive feedback loops — one that activates a CDK1 activator (Cdc25C), and one that inhibits a CDK1

inhibitor (Wee1) — results in the abrupt and irreversible transition from G_2 into mitosis.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-21 Positive feedback loops involving CDKs, Plk1, and phospho-binding domains result in abrupt mitotic CDK activation and the transition from G_2 into M phase. Although mitotic cyclins are synthesized during late S phase and G_2 and bind to CDK1, the cyclin-CDK complex is not active because threonine 14 and tyrosine 15 of the CDK1 subunit are phosphorylated by the protein kinase Wee1 (shown as red kinases), while the activating protein phosphatase, Cdc25, is phosphorylated and bound and inhibited by the phosphoserine/threonine-binding protein 14-3-3. The entire process of mitotic cyclin-CDK activation and entry into mitosis involves two positive feedback loops and coordination between events at the centrosome, in the cytoplasm, and in the nucleus. Step 1: At the centrosome, an initial pool of cyclin B-CDK1 is activated by Cdc25B, which removes the

inhibitory phosphorylations on threonine 14 and 15. This initial pool of active cyclin B-CDK1 then phosphorylates 14-3-3-bound Cdc25C on multiple serine-proline or threonine-proline sequences in the cytoplasm (shown as blue circles). Step **2**: The resulting phosphoserine/phosphothreonine-proline motifs created by cyclin B-CDK target Cdc25 for recognition by specific phospho-binding domains within the proline isomerase Pin1 and the protein kinase Plk1, which then bind to and further activate Cdc25C. Pin1 changes the conformation of Cdc25C while Plk1 catalyzes additional activating phosphorylations (shown as black circles). This releases Cdc25C from 14-3-3, allowing it to further de-phosphorylate and activate the larger cytosolic pool of cyclin B-CDK1. The resulting cyclin B-CDK1 drives further release and activation of more 14-3-3-bound Cdc25C, forming positive feedback loop 1. Step **3**: At the same time, the active cyclin B-CDK1 in the cytosol phosphorylates its inhibitory kinase Wee1 on serine-proline motifs, targeting it for recognition by the phospho-binding domain in Plk1. Plk1 then phosphorylates Wee1 on a motif that is recognized by the phospho-binding domain in the SCF-βTrCP ubiquitin protein ligase, targeting Wee1 for ubiquitination and destruction by the proteasome. This loss of Wee1 inhibition further enhances cyclin B-CDK1 activation, forming positive feedback loop 2. Step **4**: Active cytoplasmic cyclin B-CDK1 translocates into the nucleus, where it phosphorylates nuclear lamins to cause nuclear envelope breakdown and, together with Plk1 and the Aurora kinases, induces chromosome condensation and formation of the mitotic spindle.

Once the precipitous activation of mitotic CDKs is initiated, these protein kinases set in motion all the events necessary to ready the cell for chromosome segregation. The initial activation of mitotic CDKs at the centrosome, followed by activation in the cytoplasm, results in changes in the subcellular localization of these kinases. They enter the nucleus, where they bring about chromosome condensation and nuclear envelope breakdown (Figure 19-21, step **4**). In what follows, we discuss how the mitotic CDKs accomplish the coordinated execution of mitosis.

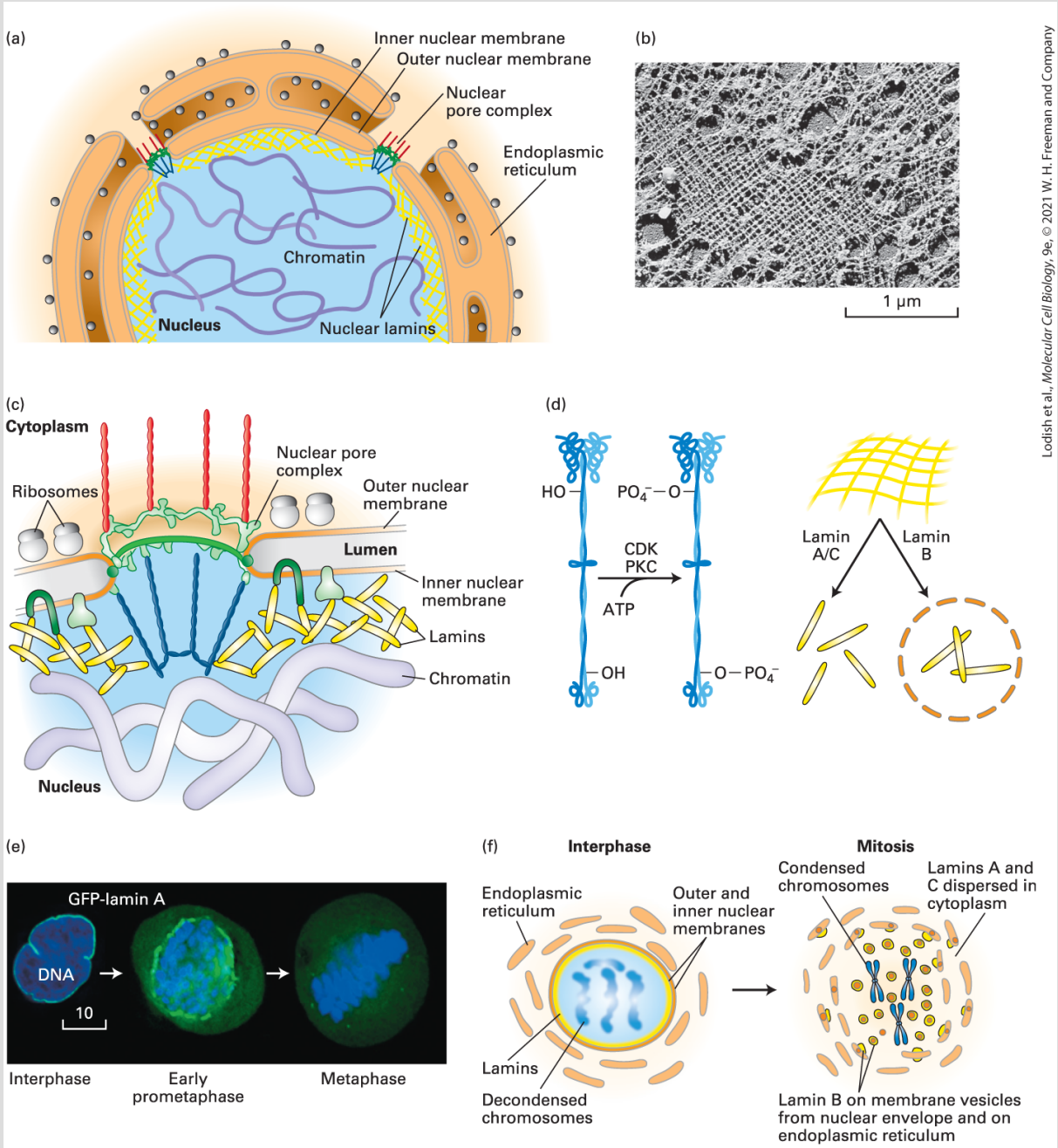
Just as S phase CDKs work together with DDK to promote MCM helicase activation during DNA replication initiation (see Section 19.4), mitotic CDKs collaborate in a similar manner with other protein kinases to bring about the mitotic events. Plk1 is critical for formation of the mitotic spindle as well as for chromosome segregation. The **Aurora kinase** family members Aurora A and Aurora B play key roles in mitotic spindle formation and in ensuring that chromosomes attach correctly to the mitotic spindle so that they are segregated accurately during mitosis. Their contributions to the various mitotic events will also be discussed.

Mitotic CDKs Promote Nuclear Envelope Breakdown

During interphase, chromosomes are surrounded by the nuclear envelope, while the centrosomes that nucleate the mitotic spindle are located in the cytoplasm. For nuclear chromosomes to interact with the cytoplasmic microtubules nucleated by the centrosomes, the nuclear envelope must be dismantled.

The nuclear envelope is a double-membrane extension of the endoplasmic reticulum containing many nuclear pore complexes (Figure 19-22a; see also Figures 1-12, 1-15, 13-33, and 18-52). The lipid bilayer of the inner nuclear membrane is associated with the nuclear lamina, a meshwork of lamin filaments adjacent to the

inside face of the nuclear envelope (Figure 19-22b) that participate in maintenance of nuclear structure and chromosome positioning within the nucleus (Figure 19-22c; see also Figure 1-15). The three nuclear lamins (A, B, and C) present in vertebrate cells belong to a class of cytoskeletal proteins, the intermediate filaments, that are critical in supporting cellular membranes. Once mitotic CDKs are activated at the end of G_2, they phosphorylate specific serine residues in all three nuclear lamins. This phosphorylation, together with phosphorylation of lamin B by protein kinase C, causes depolymerization of the lamin intermediate filaments (Figure 19-22d). Depolymerization of the nuclear lamins leads to disintegration of the nuclear lamina and contributes to disassembly of the nuclear envelope (Figure 19-22e).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-22 Dissolution of the nuclear envelope during mitosis. (a) The nuclear lamins form a meshwork between the inner nuclear membrane and the underlying chromatin (see Figure 18-53). (b) Electron micrograph of the nuclear lamina from a *Xenopus* oocyte. (c) The nuclear lamin meshwork helps tether chromatin to nuclear pore complexes and to specific proteins in the nuclear membrane. (d) Phosphorylation of lamins adjacent to the head domains by cyclin B-CDK1 and protein kinase C, results in lamin disassembly and dissolution of the nuclear membrane into vesicles or retraction into the endoplasmic

reticulum. Phosphorylation of lamins A and C results in their solubilization and dispersal in the cytoplasm, while phosphorylation of lamin B dissociates it into membrane-bound subunits. (e) Human cells in which lamin A has been fused with GFP. The DNA has been stained blue with DAPI. The nuclear lamins are beginning to dissolve in early prometaphase. By metaphase, all of the lamin A fibers have been solubilized in the cytoplasm. (f) An overview of nuclear envelope breakdown during mitosis including disassembly of nuclear lamins and dispersal of the nuclear membrane.

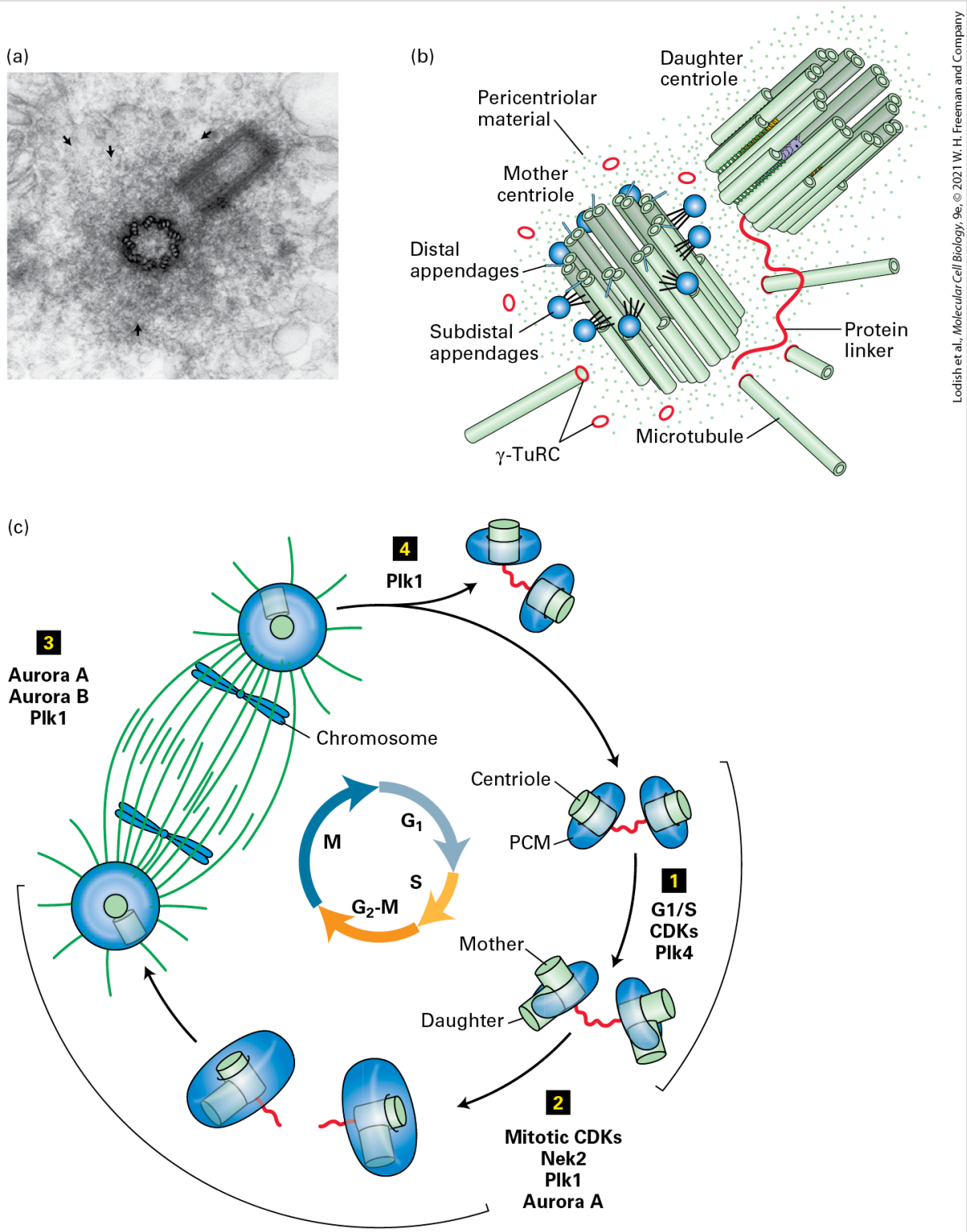
[Part (b) Republished with permission from *Nature*, from Electron micrograph of the nuclear lamina from a *Xenopus* oocyte U. Aebi et al., 323, 1986; permission conveyed through Copyright Clearance Center, Inc. Parts (c) and (d) Data from T. D. Pollard et al., 2016, *Cell Biology*, 3d ed., Elsevier, Fig. 44.6. Part (e) Republished with permission from Elsevier, from T. D. Pollard et al., 2016, *Cell Biology*, 3d ed.; permission conveyed through Copyright Clearance Center, Inc.]

Mitotic CDKs also affect other nuclear envelope components. The CDKs phosphorylate specific **nucleoporins**, which causes nuclear pore complexes to dissociate during prophase. Phosphorylation of integral membrane proteins of the inner nuclear membrane is thought to decrease their affinity for chromatin and further contributes to the disassembly of the nuclear envelope. Weakening of the associations between the inner nuclear membrane proteins and the nuclear lamina and chromatin allows sheets of inner nuclear membrane to retract into the endoplasmic reticulum, which is continuous with the outer nuclear membrane (Figure 19-22f).

Centrosomes Duplicate During S phase and Separate During Mitosis

A key function of mitotic CDKs, together with other protein kinases and ubiquitin-protein ligases, is to induce the formation of the **mitotic spindle**, also known as the *mitotic apparatus*. The function of the mitotic spindle is to segregate chromosomes so that the sister chromatids separate from each other and are moved to opposite poles of the cell (see Figure 18-39). In what follows, we describe how the mitotic spindle forms, how chromosomes attach to it, and how cells correct faulty attachments.

As we saw in Chapter 18, the mitotic spindle is made of microtubules that attach to chromosomes via specialized protein structures known as **kinetochores**, which form at the central centromeric region of chromosomes. In most organisms, the mitotic spindle is built between two **centrosomes** (called **spindle pole bodies** in yeast), that lie at opposite ends (poles) of the spindle (Figure 19-23; also see Figure 18-39). Centrosomes contain a specialized type of tubulin, γ tubulin, which forms specific ringlike complexes (γ TuRCs) that, together with a variety of associated proteins, have the ability to nucleate microtubules and form a bipolar spindle. [Notable exceptions to these centrosome-based spindle assembly mechanisms are seen in higher plants and metazoan oocytes. In these cells, the (−) ends of microtubules are cross-linked, and the microtubules self-assemble into a spindle.] In either case, the ends of the mitotic spindle are always referred to as *spindle poles*, and in animal cells these typically arise from centrosomes.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-23 The centrosome cycle is coordinated with the cell cycle to form a bipolar spindle that can correctly segregate chromosomes in mitosis. (a) Thin section electron

micrograph of an interphase centrosome showing two centrioles at right angles to each other surrounded by pericentriolar material. (b) Diagram of the centrosome showing each centriole composed of nine triplet outer microtubules embedded in pericentriolar material that contains γ-TuRC nucleating structures. The original mother centriole from the prior cell division is marked with additional appendages shown as blue spheres, and the centrioles are connected by a flexible protein linker. (c) Stages of the centrosome cycle. Step **1**: G_1/S CDKs and Plk4 induce each centriole (green) to produce a daughter procentriole, which elongates during S phase. Step **2**: The kinase Nek2 drives cleavage of the protein linker connecting the two original centrioles, while CDK1, Plk1, and Aurora A induce accumulation of large amounts of PCM (blue). Step **3**: Each centriole-procentriole pair together with its associated PCM separates from the other, forming two individual, mature centrosomes located on opposite sides of the cell nucleus. The mitotic spindle that forms between them segregates the chromosomes during the next steps in mitosis. Step **4**: At the end of mitosis, Plk1, together with a protease called *separase*, promotes centriolar disengagement, separating the former centriole-procentriole pair from the orthogonal arrangement they had during mitosis. This Plk1-mediated separation so to speak licenses the centrosome, so that procentriolar formation and centrosome duplication can only happen once each cell cycle. This is conceptually similar to how CDK activity licenses replication origins (see Figure 19-19) so that they can only fire once each cell cycle.

[Part (a) Reprinted with permission from Nature Publishing Group, from G. Sluder, 2005, "Two-Way Traffic: Centrosomes and the Cell Cycle," *Nat. Rev. Mol. Cell Biol*. **6**:743–748; permission conveyed through Copyright Clearance Center, Inc. Part (c) Data from P. T. Conduit, A. Wainman, and J. W. Raff, 2015, *Nat. Rev. Mol. Cell Biol.* **16**:611–624.]

During G_1, cells contain a single centrosome, which functions as the major microtubule nucleating center of the cell. In order for this single centrosome to create a mitotic spindle, it has to duplicate into two centrosomes and acquire additional proteins and modifications through a process called *centrosome maturation*. The two centrosomes then have to move to opposite sides of the cell and nucleate a mitotic spindle. Once mitosis is completed, the

centrosome must re-form its G_1-like structure. This sequence of events is sometimes called the *centrosome cycle* (Figure 19-23c), which is tightly coordinated with the cell cycle overall.

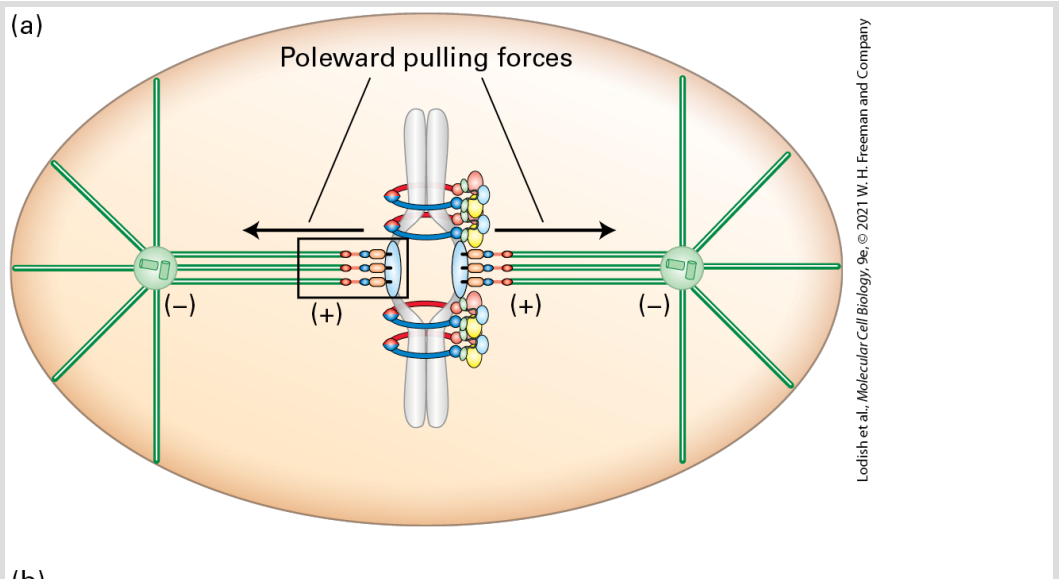
Mitotic spindle formation actually begins at the G_1–S phase transition with the duplication of the centrosome. How duplication occurs is not completely understood, but at the heart of this process is duplication of the pair of **centrioles**, short, microtubule-based structures arranged orthogonally to each other and surrounded by proteinaceous pericentriolar material (PCM). As discussed in Chapter 18, G_1 cells contain a single pair of centrioles, which are connected together by a flexible protein linker (Figure 19-23a and b). As a cell enters S phase, each of the two centrioles begins to duplicate by growing a new daughter procentriole (Figure 19-23c). This process is initiated by the G_1/S cyclin-CDKs and is largely controlled by a member of the Polo-like kinase family, Plk4, which phosphorylates additional centriole-associated proteins. They recruit γ-tubulin complexes that help form the new procentrioles (Figure 19-23c, step **1**). The procentrioles enlarge during S phase so that by G_2 each centriole has made an exact copy of itself. The two centriole-procentriole pairs remain in close proximity to each other. At the G_2/M transition, the protein linker connecting the two original centrioles is dissolved (Figure 19-23c, step **2**). This process, called **centrosome disjunction**, creates two separate centrosomes. Each centrosome expands its complement of pericentriolar material. The centrosomes recruit more γ TuRCs,

allowing the centrosome to efficiently organize and nucleate the large number of microtubules that are required for mitotic spindle formation and function (Figure 19-23c, step **3**). Once the cell enters prophase, the mature centrosomes separate from one another and move to opposite sides of the cell nucleus, where they create the mitotic spindle. This occurs in close proximity to the still-intact nuclear envelope through the actions of a kinesin motor protein (see Chapter 18), which pushes apart overlapping microtubules emanating from each centrosome. The specifics of microtubule formation, mitotic spindle assembly, and microtubule motors were discussed in Chapter 18. Here we briefly consider how chromosomes attach to the mitotic spindle and how mistakes in the process are corrected.

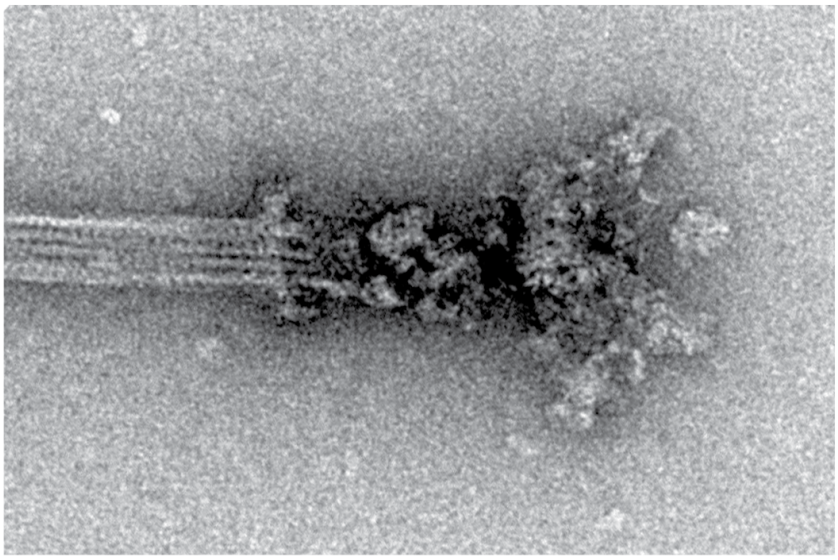
Mitotic CDKs, Polo-like Kinases, and Aurora Kinases Drive Assembly of a Mitotic Spindle That Attaches to the Kinetochores of Condensed Chromosomes

For chromosomes to be accurately segregated during mitosis, they must attach to the mitotic spindle in such a way that one kinetochore of each sister chromatid pair attaches to microtubules emanating from opposite spindle poles. Once this happens, the

sister chromatids are said to be **bi-oriented** (Figure 19-24a). This type of attachment, in which the kinetochores of each sister chromatid are attached to microtubules coming from opposing centrosomes, is called an **amphitelic attachment**. How is this accomplished? Once centrosomes have moved apart from each other during prophase, and the nuclear envelope has broken down during prometaphase (see Figure 19-3), microtubules, in a search-and-capture mechanism, begin to interact with the kinetochores of sister chromatid pairs. Initially, chromosomes glide along the length of microtubules, propelled by motor proteins. When a chromosome reaches the (+) end of a microtubule, the kinetochores attach to microtubules in an end-on attachment, the final configuration in which chromosomes are linked to the mitotic spindle (Figure 19-24b and c; also see Figure 18-42). Kinetochores of sister chromatids then bind microtubules emanating from the opposite spindle poles, allowing the chromosomes to separate by poleward pulling forces coming from motor proteins at the kinetochore.



(b)



(c)

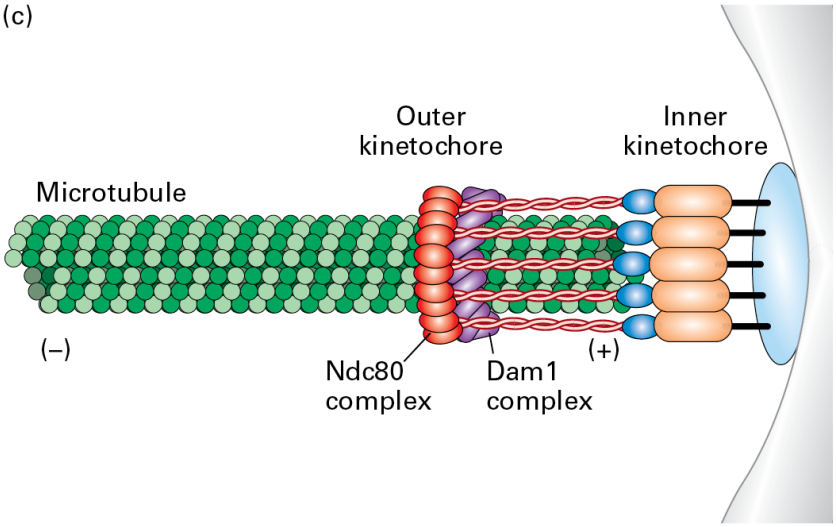


FIGURE 19-24 Chromosome attachment to the mitotic spindle. (a) Chromosomes attach to the mitotic spindle and accumulate at the spindle center. They then attach, via their kinetochores, to the ends of microtubules (called *end-on attachments*), and these attachments are stabilized by additional microtubules. The final chromosome attachment, in which the chromosome is stably bi-oriented on the mitotic spindle, is shown. "(–)" indicates the minus end of the microtubule, "(+)" the plus end. The boxed region at the kinetochore is explained in parts b and c. (b) Electron micrograph of a microtubule bound end-on to a purified yeast kinetochore. (c) Schematic cartoon illustrating the key features shown in part (b) and the boxed area in part (a). The ring structure embracing the microtubule most likely represents

the outer kinetochore Dam1 complex (called the *Ska complex* in metazoans) and part of the Ndc80 complex, also an outer kinetochore component. The globular structure at the end of the complex most likely reflects the inner kinetochore and protein complexes that link the inner kinetochore to the outer kinetochore.

[Part (b) Reprinted with permission from Nature Publishing Group, from S. Gonen et al., 2012, *Nat. Struct. Mol. Biol*. **19**:925–929, Fig. 2d; permission conveyed through Copyright Clearance Center, Inc. Part (c) Information from S. Gonen et al., 2012, *Nat. Struc. Mol. Biol.* **19**:925–929, Fig. 1a.]

Before the chromosomes can be separated by poleward pulling forces, each and every chromosome in the cell has to be properly attached to the mitotic spindle in a bi-oriented amphitelic attachment (Figure 19-25a). How does the cell know that this has occurred? Microscopic analysis of chromosome attachment has shown that, initially, many chromosomes attach to microtubules in faulty ways. A kinetochore can attach to microtubules emanating from both poles at the same time, a situation called **merotelic attachment** (Figure 19-25b). Alternatively, both kinetochores of a sister chromatid pair can attach to microtubules from the same pole (**syntelic attachment**; Figure 19-25c), or only one kinetochore can attach to microtubules (**monotelic attachment**; Figure 19-25d). Clearly none of these attachments would result in accurate chromosome segregation, since the sister chromatids would not be pulled apart to opposite spindle poles. Thus mechanisms must be in place that detect and correct such faulty kinetochore-microtubule attachments.

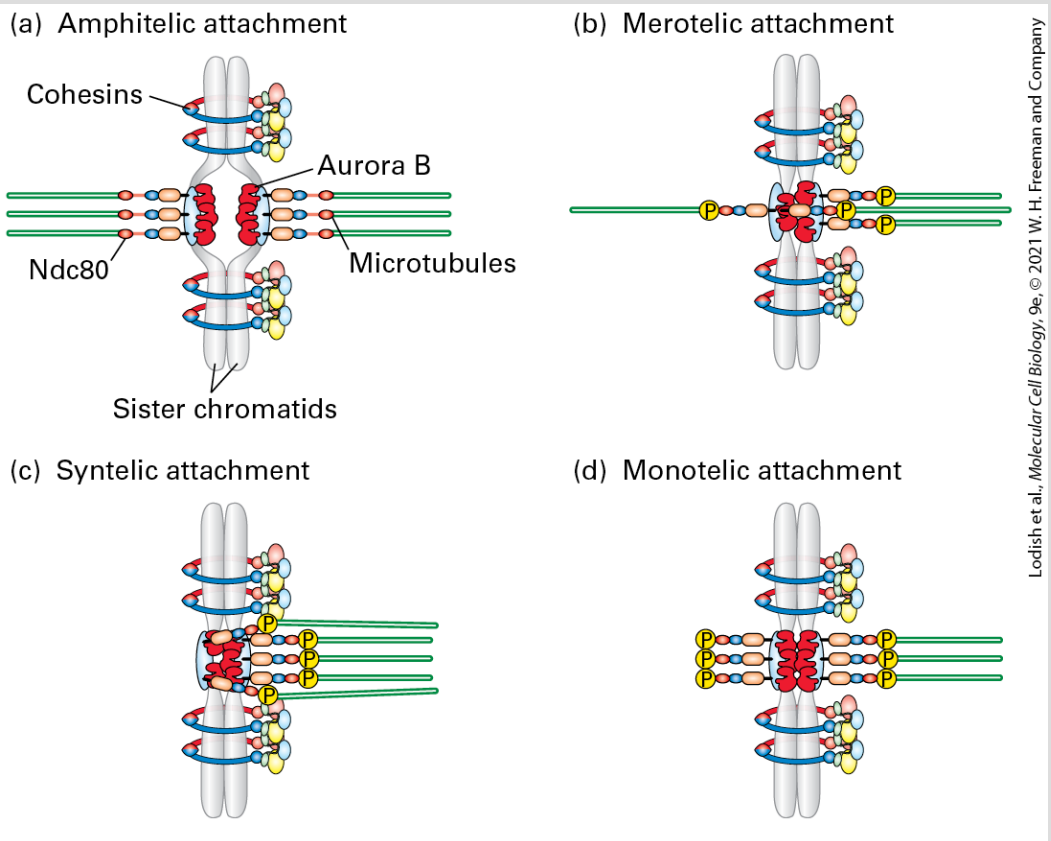


FIGURE 19-25 Stable and unstable chromosome attachments. When sister kinetochores attach to microtubules emanating from opposite spindle poles, they are stably attached. This configuration is called *amphitelic attachment* (a). Microtubules (green) pull kinetochores; cohesins resist this pulling force. The resulting tension leads to the outer kinetochore component Ndc80 (orange) being pulled away from the protein kinase Aurora B (colored red), which localizes to the inner kinetochore when chromosomes are near the spindle midzone. As a result, Aurora B can no longer phosphorylate Ndc80, and kinetochore-microtubule attachments are stable. When one of the two sister chromatid kinetochores simultaneously attaches to microtubules emanating from two opposite spindle poles [merotelic attachment, (b)], or both sister kinetochores attach to microtubules emanating from the same spindle pole [syntelic attachment, (c)], or only one of the two sister kinetochores attaches to microtubules [monotelic attachment, (d)], Ndc80 is not pulled away from Aurora B. As a result, Aurora B phosphorylates Ndc80 (yellow circles labeled 'P'), and Ndc80 can no longer form stable attachments to microtubules.

The sensing mechanism used by cells to detect incorrect chromosome attachments is based on tension. When sister chromatids are correctly attached to microtubules, their kinetochores are under tension (Figure 19-25a). Microtubules attached to the kinetochores pull at them, while the cohesin molecules that hold the sister chromatids together withstand these forces, creating tension at the kinetochores. Merotelic, syntelic, or monotelic attachment leads to insufficient tension at kinetochores, allowing the cell to distinguish these faulty forms of attachment from the correct amphitelic one.

How does the cell measure whether or not kinetochores are under tension? The Aurora protein kinase family members *Aurora A* and *Aurora B* and their associated regulatory factors sense kinetochores that are not under tension and sever these microtubule attachments, giving cells a second chance to get the attachment right. The molecular basis for this sensing mechanism is reasonably well understood. Recall that outer kinetochore components, especially the Ndc80 complex, bind to microtubules (see Figure 19-24c) in a manner that is stabilized by the Ska complex (or the Dam1 complex in yeast). Aurora kinases phosphorylate the N-terminal tail of Ndc80. When phosphorylated, Ndc80 loses its ability to form stable interactions with microtubules. Aurora B localizes to the inner kinetochore and is thought to be the dominant error-correcting kinase for chromosomes along most of the spindle. When kinetochores are not under tension, Ndc80 is in close proximity to

Aurora B, and the protein kinase can phosphorylate the protein, destabilizing any kinetochore-microtubule attachments (Figure 19-25b–d). When microtubules are attached correctly to kinetochores, microtubule forces pull Ndc80 away from Aurora B, and the kinase can no longer phosphorylate Ndc80 (Figure 19-25a). Protein phosphatase 1 (PP1) localizes to the outer kinetochore and continuously de-phosphorylates Ndc80. Thus, when kinetochores are under tension and pulled away from Aurora B, Ndc80 is quickly de-phosphorylated by PP1 and microtubule–kinetochore attachments are stabilized. A similar mechanism is thought to operate on kinetochore-microtubule interactions close to the spindle poles, where Aurora A substitutes for Aurora B.

Microtubules continuously pull on chromosomes. Once all the chromosomes have attached to microtubules in an amphitelic manner, only the cohesins that hold them back in the middle of the spindle (Figure 19-25a) prevent chromosomes from segregating to the poles. As we will see in Section 19.6, it is the severing of these cohesins that initiates anaphase chromosome segregation and drives mitotic exit. Again, just as in S phase entry, mitotic entry, chromosome replication, and the centrosome cycle, protein phosphorylation and de-phosphorylation by protein kinases and phosphatases control all of the events leading to chromosome separation.

Chromosome Condensation Facilitates Chromosome Segregation

Chromosome segregation not only requires building the apparatus that segregates chromosomes, but it also requires that the DNA be compacted into travel-friendly structures. Any attempt to segregate the long and intertwined DNA-protein complexes present in interphase cells would lead to DNA breakage and loss of genetic material. To avoid this fate, cells compact their chromosomes during prophase into the dense structures we have become acquainted with through light and electron microscopy (Figure 19-26a and see Figure 19-2).

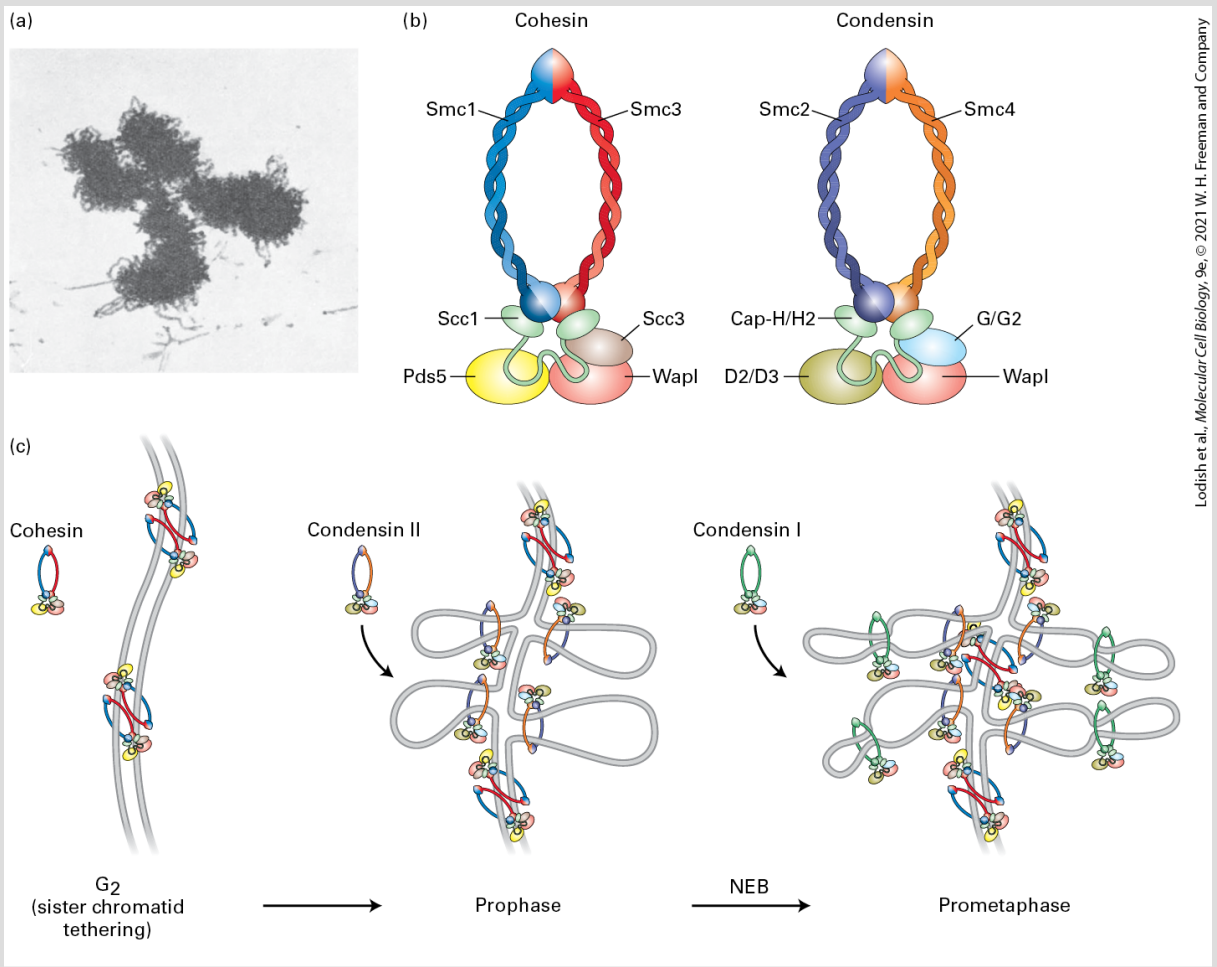


FIGURE 19-26 The Condensin complex packages chromosomes for efficient segregation on the mitotic spindle. (a) Electron micrograph of a metaphase chromosome. By metaphase, the chromosomes are fully condensed, and the two individual sister chromatids are visible. (b) The condensin complexes have a similar architecture to the cohesin complex discussed previously. (c) A model of cell cycle regulated chromosome compaction involving progressive chromosome looping. During prophase, prior to nuclear envelope breakdown, condensin II binds DNA and extrudes large loops. After nuclear envelope breakdown in prometaphase, condensin I forms additional loops nested within the loops formed by condensin II.

[Part (a) Republished with permission from Nature Publishing Group, from E. J. DuPRAW and P. M. M. Rae, 1966, “Polytene Chromosome Structure in Relation to the ‘Folded Fibre’ Concept”. *Nature* **212**:598–600; permission conveyed through Copyright Clearance Center, Inc.]

Chromosome condensation results in a dramatic reduction in chromosome length, up to 10,000-fold in vertebrates. During chromosome compaction, the intertwined sister chromatids are also untangled. This process, called **sister chromatid resolution**, is mediated in part by topoisomerase II, which cuts one chromatid in a pair of sister chromatid molecules, passes the other one through, and closes the cut molecule. As part of this process, during prophase most of the cohesin proteins that tether sister chromatids together along the chromosome arms are shed through the actions of Plk1 and Aurora B.

Central to the process of chromosome condensation is another ring-shaped protein complex known as **condensin**. Most eukaryotes have two closely related condensin complexes, called *condensin I* and *II*. These protein complexes are closely related to the cohesins that link sister chromatids together (Figure 19-26b). Condensins were first identified based on their ability to promote chromosome condensation in frog egg extracts. Like cohesins, condensins are composed of two large coiled-coil SMC protein subunits that associate through their ATPase domains with non-SMC subunits. When condensin function is lost in cells, chromosomes do not condense and sister chromatid tangles are not resolved.

Recent studies have shown that chromosome compaction occurs via the generation of consecutive loops that lead to the formation of a fiber at the bases of the loops (Figure 19-26c). These fibers are

compressed, leading to further chromosome compaction. Condensins likely package chromosomes by creating these loops by forming intrachromosomal linkages, through the process explained in Figure 19-26c. Condensin association with chromosomes is facilitated by mitotic CDKs and Aurora B, which phosphorylate histone H2A (see Figure 7-26b), allowing condensins to bind chromatin.

KEY CONCEPTS OF SECTION 19.5

The G_2/M Transition and the Irreversible Engine of Mitosis

- Mitotic CDKs induce entry into mitosis in all eukaryotes.
- Mitotic CDKs are kept inactive by inhibitory phosphorylation of the CDK subunit until the completion of DNA replication (see Figure 19-21).
- Mitotic CDKs promote their own activation through two positive feedback loops that inactivate Wee1 kinase and activate Cdc25 phosphatase (see Figure 19-21).
- Mitotic CDKs induce nuclear envelope breakdown in most eukaryotes by phosphorylating lamins (see Figure 19-22).
- Centrosome duplication occurs during S phase. Mitotic CDKs induce the separation of the duplicated centrosomes, which initiates mitotic spindle formation (see Figure 19-23).
- Sister chromatids attach to the mitotic spindle via their kinetochores in a bi-oriented manner, with one sister kinetochore attaching to microtubules emanating from one spindle pole and the other one to microtubules nucleated by the other spindle pole (see Figure 19-24).
- Cells sense bi-orientation of sister chromatids through a tension-based mechanism. When kinetochores are not under tension, the protein kinases Aurora A and Aurora B phosphorylate the microtubule-binding subunits of the kinetochore, which limits their ability to form stable long-lasting interactions (see Figure 19-24).
- Chromosomes must be compacted for segregation to occur.
- Condensins, protein complexes that are related to cohesins, facilitate chromosome condensation and are activated by mitotic CDKs (see Figure 19-26).

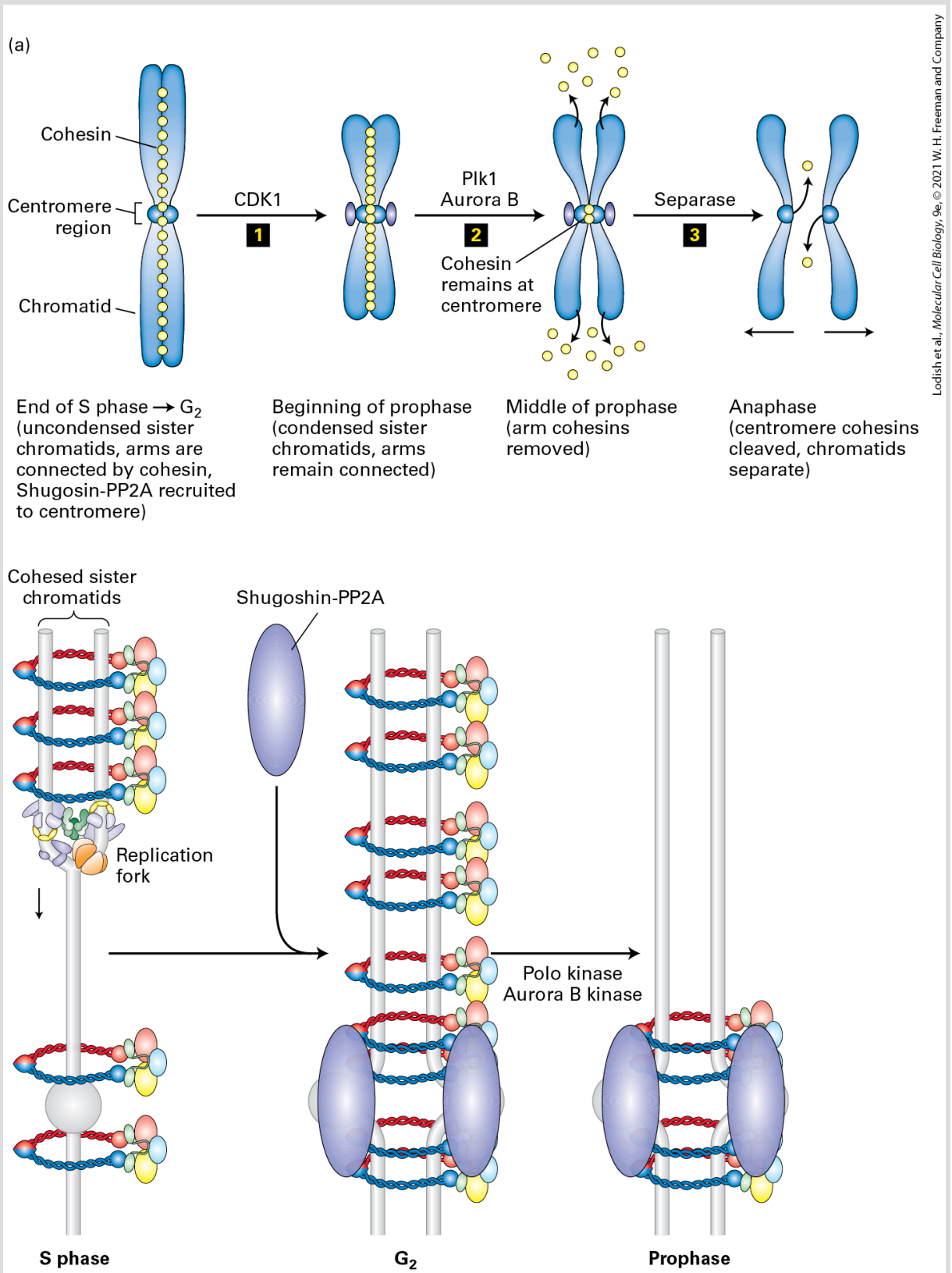
19.6 The Mitotic Spindle, Chromosome Segregation, and Exit from Mitosis

Once all chromosomes have condensed and have correctly attached to the mitotic spindle, chromosome segregation commences. In this section, we discuss how cleavage of cohesins by a protease known as *separase* triggers anaphase chromosome movement and how this cleavage is regulated by phosphorylation and protein ubiquitination. We then see how the same machinery that initiates cohesin cleavage at the metaphase-anaphase transition also initiates mitotic CDK inactivation. Next we describe how phosphatases activated at the end of mitosis also participate in mitotic CDK inactivation, bringing about the disassembly of mitotic structures and the resetting of the cell to the G_1 state. We end this section with a discussion of cytokinesis, the process that produces two daughter cells.

Separase-Mediated Cleavage of Cohesins Initiates Chromosome Segregation

As mentioned in the previous section, chromosome condensation and dissociation of cohesins along the chromosome arms leads to further compaction of chromosomes. Most of the cohesins along the

chromosome arms are removed from chromosomes during prophase (Figure 19-27a, steps **1** and **2**). This process is mediated by phosphorylation of cohesins by Plk1 and Aurora B kinase. In most organisms, by metaphase, cohesins are maintained only around centromeres. These centromere-associated cohesins are specifically protected from phosphorylation-dependent removal by protein phosphatase 2A (PP2A). This phosphatase is recruited to centromeric regions in human cells from G_2 through metaphase by a member of a family of PP2A targeting factors known as the *Shugoshin* family of proteins (Figure 19-27b). The protected pool of cohesins provides the resistance to the pulling force exerted by microtubules that are necessary to establish tension at bi-oriented kinetochores. As we will see in Section 19.8, this protection mechanism also plays an essential role in establishing the meiotic chromosome segregation pattern.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-27 Cleavage of cohesin at centromeres by separase triggers anaphase onset. (a) A multistep process removes the cohesin that links sister chromatids. Step **1**: During G_2, when sister chromatids are linked along their entire length by cohesins, the Shugoshin proteins begin recruiting the protein phosphatase 2A (PP2A) to centromeric regions. This recruitment process continues through the beginning of prophase when CDK1 triggers chromosome condensation. Step **2**: Later in prophase and early metaphase, cohesin along the chromosome arms is shed by the action of a protein complex in combination with phosphorylation of cohesins by Polo kinase and Aurora B kinase. However, cohesin remains tightly associated with centromeric regions of the chromosome because PP2A localized there removes any phosphates added by the Plk1 kinase. Finally, in step **3** separase cleaves the residual centromeric cohesin, triggering the onset of anaphase, releasing the sister chromatids, which move rapidly toward opposite spindle poles. (b) Diagrammatic representation of Shugoshin-PP2A recruitment and retention at centromeric chromatin.

Each sister chromatid of a metaphase chromosome is attached to microtubules via its kinetochore (see Figure 19-24). Once all of the chromosomes have made stable amphitelic attachments to the spindle pole microtubules and congressed to the spindle midzone at metaphase, the chromosomes will be in a state of tension, with forces pulling the two kinetochores toward opposite spindle poles. Sister chromatids do not separate from each other yet, however, because they are held together at their centromeres by what remains of the cohesion attachments. In all organisms analyzed to date, the loss of this residual cohesin from the centromeric regions of chromosomes is what triggers anaphase chromosome movement (Figure 19-27a, step **3**). The mechanism that brings about this loss of cohesins from chromosomes is conserved as well. A protease known as **separase** cleaves the cohesin subunit Scc1 (Figure 19-28), breaking the protein circles linking sister chromatids. Once this link

is broken, anaphase begins as poleward force exerted on the kinetochores moves the split sister chromatids toward opposite spindle poles.

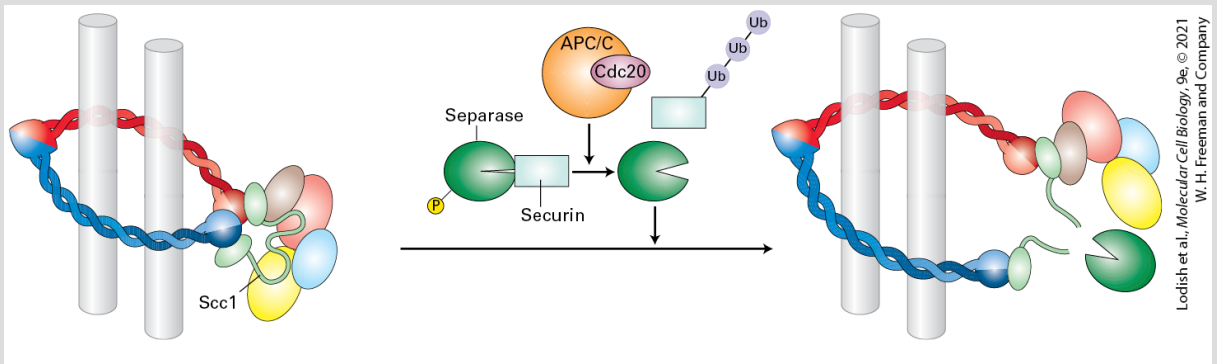


FIGURE 19-28 Regulation of separase through phosphorylation and ubiquitin-mediated degradation of an inhibitor establishes the timing of cohesin cleavage. Separase is inhibited before anaphase by the binding of the protein securin. Mitotic CDKs also inhibit separase by phosphorylating it. When all the kinetochores have properly attached to spindle microtubules and the spindle apparatus is properly assembled and oriented, the Cdc20 specificity factor associated with APC/C directs it to ubiquitinylate securin and mitotic cyclins. Following securin degradation and a decrease in mitotic CDK activity, separase cleaves the Scc1 subunit, breaking the cohesin circles and allowing sister chromatids to be pulled apart by the spindle apparatus that is pulling them toward opposite spindle poles.

Cohesin cleavage was discovered in budding yeast. Insight into the identity of the protein that was responsible for the cleavage of cohesin came from analysis of previously identified yeast mutants that failed to segregate chromosomes during anaphase. A mutant form of the gene encoding Esp1 — what we now call separase — failed to produce the cleavage fragment. Subsequent analyses revealed that separase is a protease and that cleavage of cohesin by separase is essential for chromosome segregation. Given the

irreversible nature of Scc1 cleavage, it is absolutely essential that separase activity be tightly controlled. Next we discuss its regulation.

APC/C Activates Separase Through Securin Ubiquitinylation

Prior to anaphase, the protein **securin** binds to and inhibits separase (see Figure 19-28). Once all kinetochores have attached to spindle microtubules in the correct bi-oriented manner, the APC/C ubiquitin-protein ligase, directed by specificity factor Cdc20 (known as APC/C^{Cdc20}) (see Figure 19-14), ubiquitinylates securin. The polyubiquitinylated securin is rapidly degraded by proteasomes, thus releasing separase.

$\mathrm{APC/C^{Cdc20}}$ is phosphorylated and activated in prophase by mitotic CDK phosphorylation. We will see in Section 19.7, however, that phosphorylated $\mathrm{APC/C^{Cdc20}}$ is inhibited by a checkpoint pathway — the spindle assembly checkpoint pathway — that ensures anaphase is not initiated until all of the chromosomes have achieved proper attachment to the mitotic apparatus (i.e., are bi-oriented). Cdc20 is inhibited until every kinetochore has attached to microtubules and proper tension has been applied to the kinetochores of all sister chromatids. In vertebrate cells, separase itself is also negatively regulated by CDK phosphorylation during prophase and metaphase.

Separase becomes active and triggers chromosome segregation when APC/C^{Cdc20}-mediated protein degrades because mitotic CDK activity begins to decline at the metaphase-anaphase transition.

Once cohesins are cleaved, anaphase chromosome movement ensues. As discussed in Chapter 18, chromosome segregation is mediated by microtubule depolymerization and motor proteins as the spindle poles move away from each other. Decline in mitotic CDK activity is important for these anaphase chromosome movements. If mitotic CDK inactivation is inhibited, anaphase does occur, but it is abnormal. De-phosphorylation of a number of microtubule-associated proteins that affect microtubule dynamics appears to be important for this process. In budding yeast, this de-phosphorylation is brought about by the protein phosphatase Cdc14, which, as we will see, plays an essential role in the final cell cycle stage, exit from mitosis.

Mitotic CDK Inactivation and Protein De-phosphorylation Triggers Exit from Mitosis

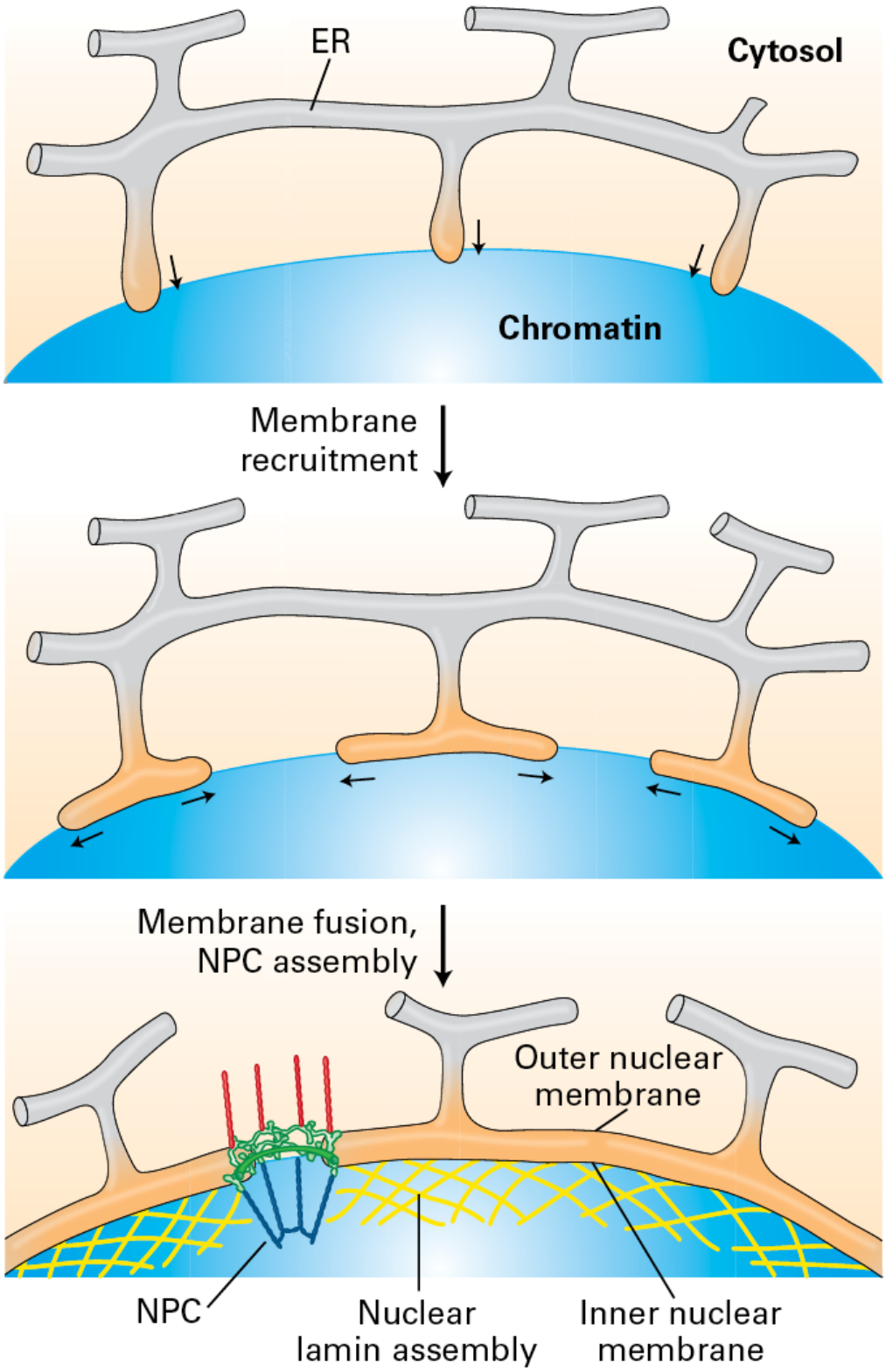
Anaphase spindle elongation and the events associated with exit from mitosis — mitotic spindle disassembly, chromosome decondensation, and nuclear envelope re-formation — require the de-phosphorylation of CDK substrates. In other words, the

phosphorylation events that triggered all of the different mitotic events need to be undone for the cell to revert to the G_1 state.

De-phosphorylation of mitotic CDK substrates involves two steps. First, the mitotic CDKs themselves have to be inactivated. In most organisms, mitotic CDK inactivation is triggered by APC/C^{Cdc20}-mediated degradation of mitotic cyclins. However, in budding yeast, only about 50 percent of mitotic cyclins are degraded by APC/C^{Cdc20}. An additional pathway called the *mitotic exit network* plays a particularly important role in shutting off mitotic cyclins in yeast by controlling the activation of the **Cdc14 phosphatase** through a GTPase pathway called the mitotic exit network. Activation of Cdc14 at anaphase de-phosphorylates Cdh1, allowing formation APC/C^{Cdh1}, which then degrades the remaining mitotic cyclins. In addition, Cdc14 promotes the re-accumulation of the CDK inhibitor Sic1, which inhibits any residual mitotic CDK activity, similarly to how it inhibits G_1/S phase cyclins in previous stages of the cell cycle (Section 19-4). This process leads to exit from mitosis. Phosphatase activity is also essential for exit from mitosis in vertebrate cells, although they do not appear to possess a mitotic exit network like that found in budding yeast.

Ultimately, reversal of mitotic CDK phosphorylation changes the activities of many proteins back to their interphase states. De-phosphorylation of condensins and other chromatin-associated proteins leads to decondensation of mitotic chromosomes in

telophase. De-phosphorylated inner nuclear membrane proteins are thought to bind to chromatin once again. As a result, multiple projections of regions of the ER membrane containing these proteins are thought to associate with the surfaces of the decondensing chromosomes and then fuse with one another, directed by an unknown mechanism to form a continuous double membrane around each chromosome (Figure 19-29). De-phosphorylation of nuclear pore subcomplexes allows them to reassemble into complete NPCs traversing the inner and outer membranes soon after fusion of the ER projections. Ran·GTP, required for driving most nuclear import and export (see Chapter 13), stimulates both fusion of the ER projections to form daughter nuclear envelopes and assembly of NPCs (Figure 19-29). The Ran·GTP concentration is highest in the vicinity of the decondensing chromosomes because the Ran-guanine nucleotide–exchange factor (Ran-GEF) is bound to chromatin. Consequently, membrane fusion is stimulated at the surfaces of decondensing chromosomes. Sheets of nuclear membrane with inserted NPCs then fuse with one another to form one nuclear membrane around all of the chromosomes.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-29 Model for reassembly of the nuclear envelope during telophase. Extensions of the endoplasmic reticulum (ER) associate with each decondensing chromosome and then fuse with one another, forming a double membrane around the chromosome. De-phosphorylated nuclear pore subcomplexes reassemble into nuclear pores, forming individual mini-nuclei called *karyomeres*. The enclosed chromosome further decondenses, and subsequent fusion of the nuclear envelopes of all the karyomeres at each spindle pole forms a single nucleus containing a full set of chromosomes. NPC: nuclear pore complex. See B. Burke and J. Ellenberg, 2002, *Nat. Rev. Mol. Cell Biol.* **3**:487–497.

Cytokinesis Creates Two Daughter Cells

When chromosome segregation is completed, the cytoplasm and organelles are distributed between the two future daughter cells. This process is called **cytokinesis**. With the exception of higher plants, the division of the cell is brought about by a contractile ring made of actin and the actin motor myosin (see Figure 17-34). During cytokinesis, the ring contracts in a manner similar to muscle contraction, pulling the membrane inward and eventually closing the connection between the two daughter cells (Figure 19-30).

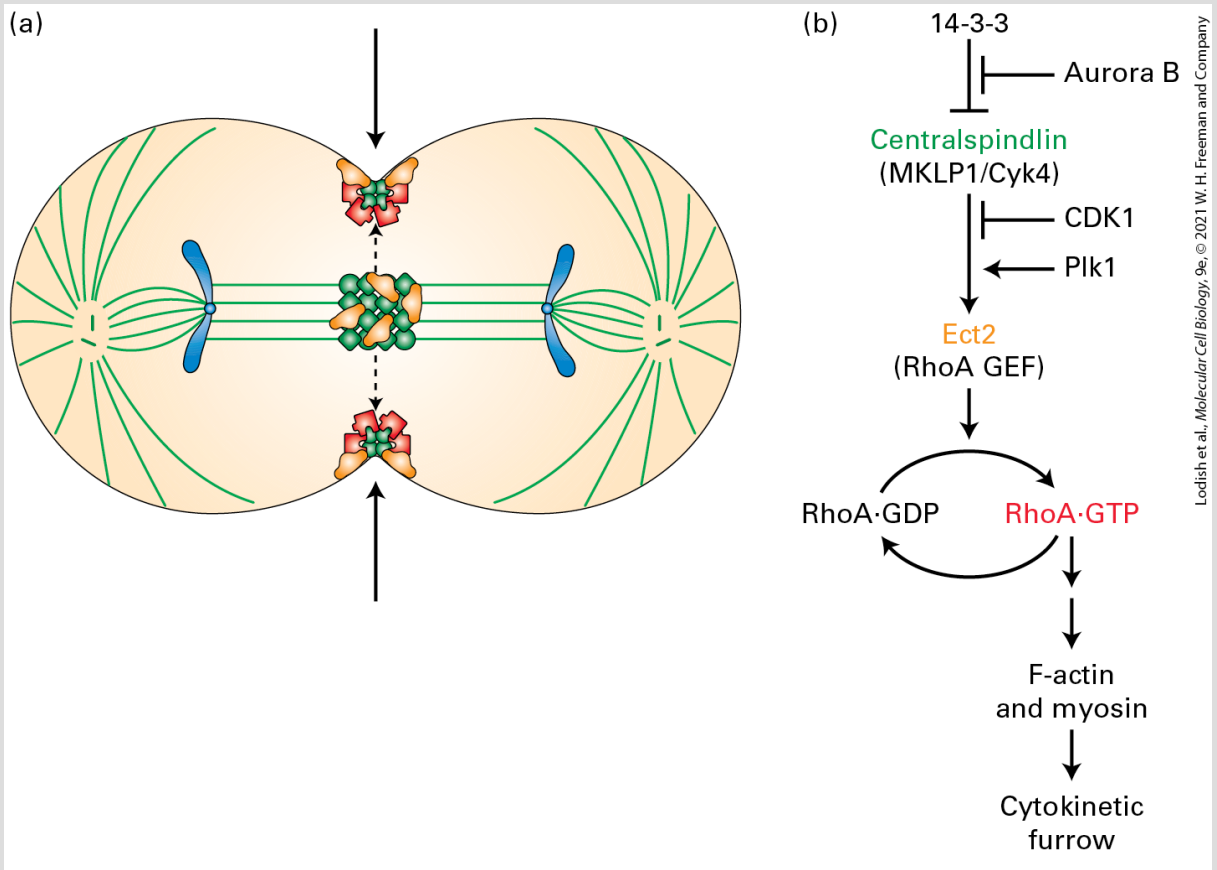


FIGURE 19-30 Control of RhoA regulates cytokinesis. (a) The Centralspindlin complex (green) recruits and activates the RhoGEF Ect2 (orange) at the central spindle during anaphase. These colocalize with and activate RhoA (red) on the cell cortex opposite the central spindle, generating an actomyosin-based contractile ring that creates a cleavage furrow and divides the cell in half. (b) Molecular pathway of RhoA activation and control of cleavage furrow formation by phosphorylation.

[Information from A. Basant and M. Glotzer, 2018. *Curr. Biol*. **28**:R570-R580.]

Cytokinesis must be coordinated with other cell cycle events in space and time. For cell division to produce two daughter cells each containing half of the parent cell's cytoplasmic content and exactly half of the genetic content, the division plane must be properly spatially positioned. Cytokinesis must also be temporally

coordinated with completion of mitosis. In what follows, we explore both these aspects of cytokinesis regulation.

In animal cells, the contractile ring forms during anaphase and must be placed exactly in the middle of the anaphase spindle to ensure that each daughter cell receives half of the genetic material. Contractile ring assembly requires the small GTPase RhoA (see Section 17.7). Regions of active RhoA form on the cortical surface of the cell opposite the central part of the spindle at sites where the contractile ring will form (Figure 19-30a). Active GTP-bound RhoA both stimulates formin-mediated F-actin assembly and promotes the action of myosin II (see Figure 17-42). Recall that in yeast, small GTPases play a critical role in forming the mitotic exit network, which is absent in animal cells. Instead, small GTPases in animal cells are required for cytokinesis through a process that is tightly controlled by mitotic protein kinases (Figure 19-30b).

RhoA itself is activated by a conserved Rho guanine nucleotide–exchange factor (GEF) called *Ect2*, which is itself tightly regulated by a multiprotein complex called *Centralspindlin*. Centralspindlin is a heterotetramer composed of two copies of the kinesin-6 motor protein MKLP1 and a dimer of a Rho GTPase-activating protein (GAP). As one can guess from its name, most of the Centralspindlin localizes to the midzone of the anaphase spindle, where the MKLP1 components form oligomers that bundle spindle microtubules and organize the central spindle in preparation for cell abscission at the

end of cytokinesis. A variable amount of Centralspindlin, however, localizes to the cell cortex adjacent to the central spindle, together with the RhoA activator Ect2 and RhoA (Figure 19-30a). MKLP1 is inhibited by the phosphoserine/threonine-binding protein 14-3-3 (see Figure 19-30b). This ensures that the spindle midzone microtubules do not compact until the end of anaphase. Similarly, the RhoA activator Ect2 is auto-inhibited during prior stages of mitosis, but it contains two phospho-binding BRCT domains that are important for its activation. During anaphase, when cyclin B–CDK1 activity is extinguished but Aurora and Polo kinases are still active, Aurora B phosphorylates MKLP1 and displaces the 14-3-3 inhibitor, allowing MKLP1 to acquire catalytic motor activity and compact the spindle midzone. At the same time, Plk1 phosphorylates the Rho-GAP component, recruiting the BRCT domains of Ect2 to form an active Centralspindlin-Ect2 complex. This catalyzes the GTP-dependent activation of RhoA, leading to cleavage furrow formation and division of the cytoplasm to form two completely separate cells.

This concludes our discussion of the molecular events of cell division. As we have seen, cyclin-dependent kinases, phosphoserine/threonine-binding domains, and ubiquitin-mediated protein degradation are at the center of its control (Figure 19-31). In Section 19.7, we discuss the mechanisms that ensure that a cell cycle stage is not initiated until the previous one has been completed and that each cell cycle step occurs accurately.

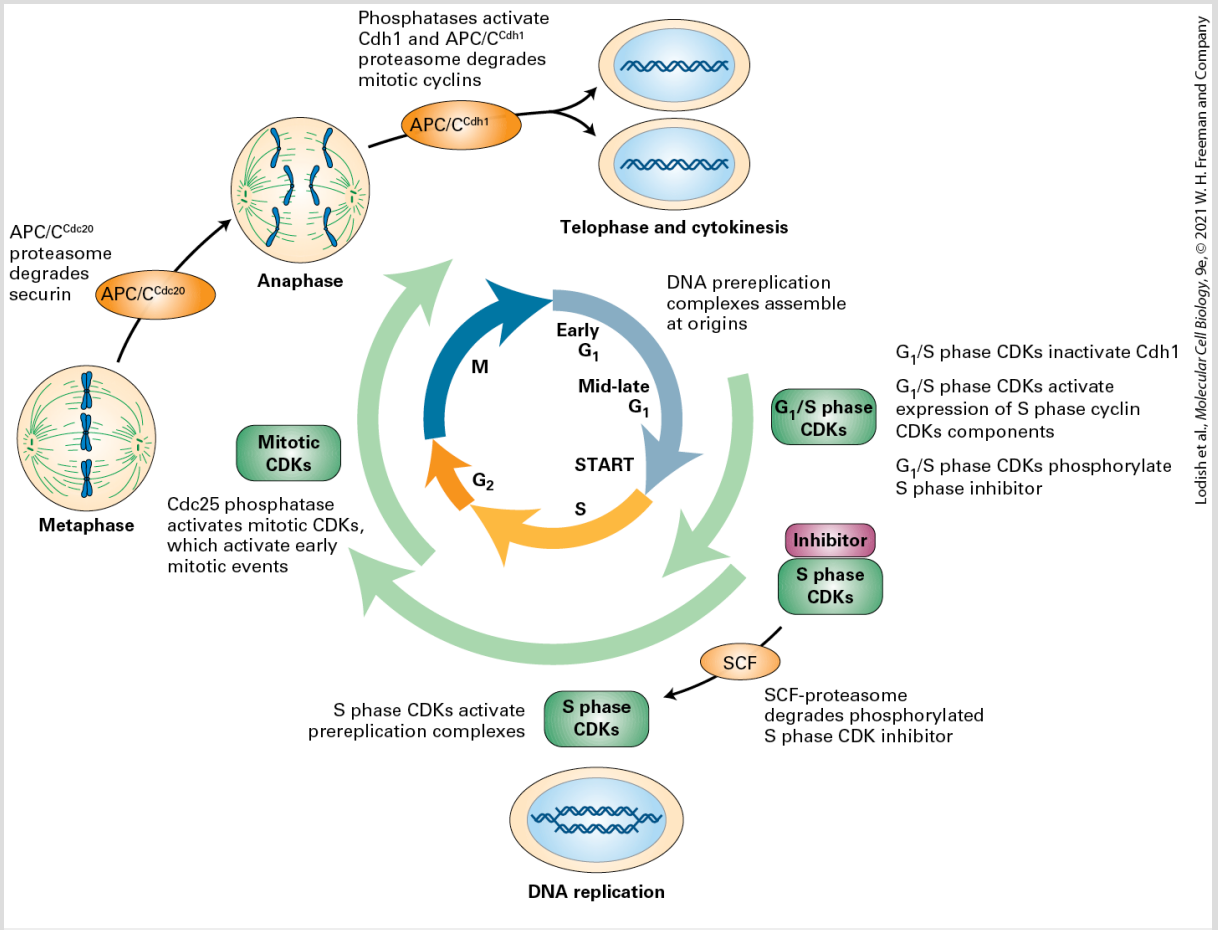


FIGURE 19-31 Fundamental processes in the eukaryotic cell cycle. A synopsis of major events that regulate progression through the cell cycle. See text for details.

KEY CONCEPTS OF SECTION 19.6

The Mitotic Spindle, Chromosome Segregation, and Exit from Mitosis

- Cleavage of cohesin by separase induces chromosome segregation during anaphase.
- At the onset of anaphase, APC/C is directed by Cdc20 to ubiquitinylate securin, which is subsequently degraded by proteasomes. The degradation of securin, along with phosphorylation by Plk1, activates separase (see Figure 19-28).
- Exit from mitosis is triggered by mitotic CDK inactivation mainly brought about by mitotic cyclin degradation.

- Exit from mitosis requires the activity of protein phosphatases such as Cdc14 to remove mitotic phosphates from many different proteins, permitting mitotic spindle disassembly, decondensation of chromosomes, and reassembly of the nuclear envelope.
- Cytokinesis finalizes cell fission and must be coordinated with the site of nuclear division. In animal cells, this coordination involves the small GTPase RhoA and is regulated by Plk1, Aurora B, and phospho-binding domains (see Figure 19-30).

19.7 Surveillance Mechanisms in Cell Cycle Regulation

Surveillance mechanisms known as **checkpoint pathways** operate to ensure that the next cell cycle event is not initiated until the previous one has been completed. Checkpoint pathways consist of **sensors** that monitor a particular cellular event, a **signaling cascade** that initiates the response, and an **effector** that halts cell cycle progression and activates repair pathways when necessary until the defect is corrected. If the defect cannot be repaired, the checkpoint pathway induces apoptosis.

Cell cycle events monitored by checkpoint pathways include growth, DNA replication, DNA damage, kinetochore attachment to the mitotic spindle, and positioning of the spindle within the cell. These pathways are responsible for the extraordinarily high fidelity of cell division, ensuring that each daughter cell receives the correct number of accurately replicated chromosomes. Checkpoint pathways function through protein phosphorylation, targeting proteins to bind to phosphoserine/threonine-binding domains that control the activities of the cyclin-CDKs. This control involves a variety of mechanisms: blocking synthesis and increasing degradation of cyclins, phosphorylating CDKs at inhibitory sites, sequestering CDK activators away from the CDKs, stabilizing CDK inhibitors (CKIs) that inactivate cyclin-CDK complexes, and inactivating the APC/C ubiquitin-protein ligase.

The DNA Damage Response System Halts Cell Cycle Progression and Recruits DNA Repair Machinery When DNA Is Compromised

The complete and accurate duplication of the genetic material is essential for cell division. If cells enter mitosis when DNA is incompletely replicated or otherwise damaged, genetic changes occur. In many instances, those changes lead to cell death, but as we will see in Chapter 25, they can also lead to genetic alterations that result in loss of control over cell growth and division and, eventually, cancer. This risk is underscored by the finding that many proteins involved in sensing DNA damage and its repair are frequently found mutated in human cancers.

The enzymes that replicate DNA are accurate, but not enough to ensure complete accuracy during DNA synthesis. Furthermore, environmental insults such as x-rays and UV light can cause DNA damage, and this damage must be repaired before a cell enters mitosis. Cells possess a *DNA damage response system* that senses many different types of DNA damage and responds by activating repair pathways and halting cell cycle progression until the damage has been repaired. Cell cycle arrest can occur in G_1, S phase, or G_2, depending on whether DNA damage occurred before cell cycle entry or during DNA replication. In multicellular organisms, when the

DNA damage is severe, cells forego repair and instead initiate either apoptosis ("programmed cell death"), a mechanism that is discussed in detail in Chapter 22, or they senesce, becoming large, flattened cells that have permanently stopped dividing.

DNA damage exists in many different forms and degrees of severity. A break of the DNA helix, known as a *double-strand break*, is perhaps the most severe form of damage because such a lesion would almost certainly lead to DNA loss if mitosis ensued in its presence. This is because the severed piece of DNA that was not attached to the centromere of the chromosome, where the kinetochore assembles, would not be reliably separated into the two daughter cells. More subtle defects include single-strand breaks, structural changes in nucleotides, and DNA mismatches. For our discussion here, it is important to note that cells have sensors for all these different types of damage. These sensors scan the genome and, when they detect a lesion, assemble signaling and repair factors at the site of the lesion.

Central to the detection of these different types of lesions is a pair of homologous protein kinases called **ATM** (for **ataxia telangiectasia mutated**) and **ATR** (for **ataxia telangiectasia and Rad3-related protein**). In addition to ATM and ATR, a related protein kinase called **DNA-PK** (for **DNA-dependent protein kinase**) is also critical for DNA damage signaling and repair. All three protein kinases are recruited to sites of DNA damage through specific protein complexes that form on the chromatin at the site of the damage. They then

initiate the sequential recruitment of adapter proteins and DNA repair proteins near the site of the DNA lesion. In addition, they activate another set of downstream effector protein kinases called *Chk1, Chk2*, and *MK2* that establish and maintain cell cycle checkpoints that stop progression through the cell cycle (Figure 19-32).

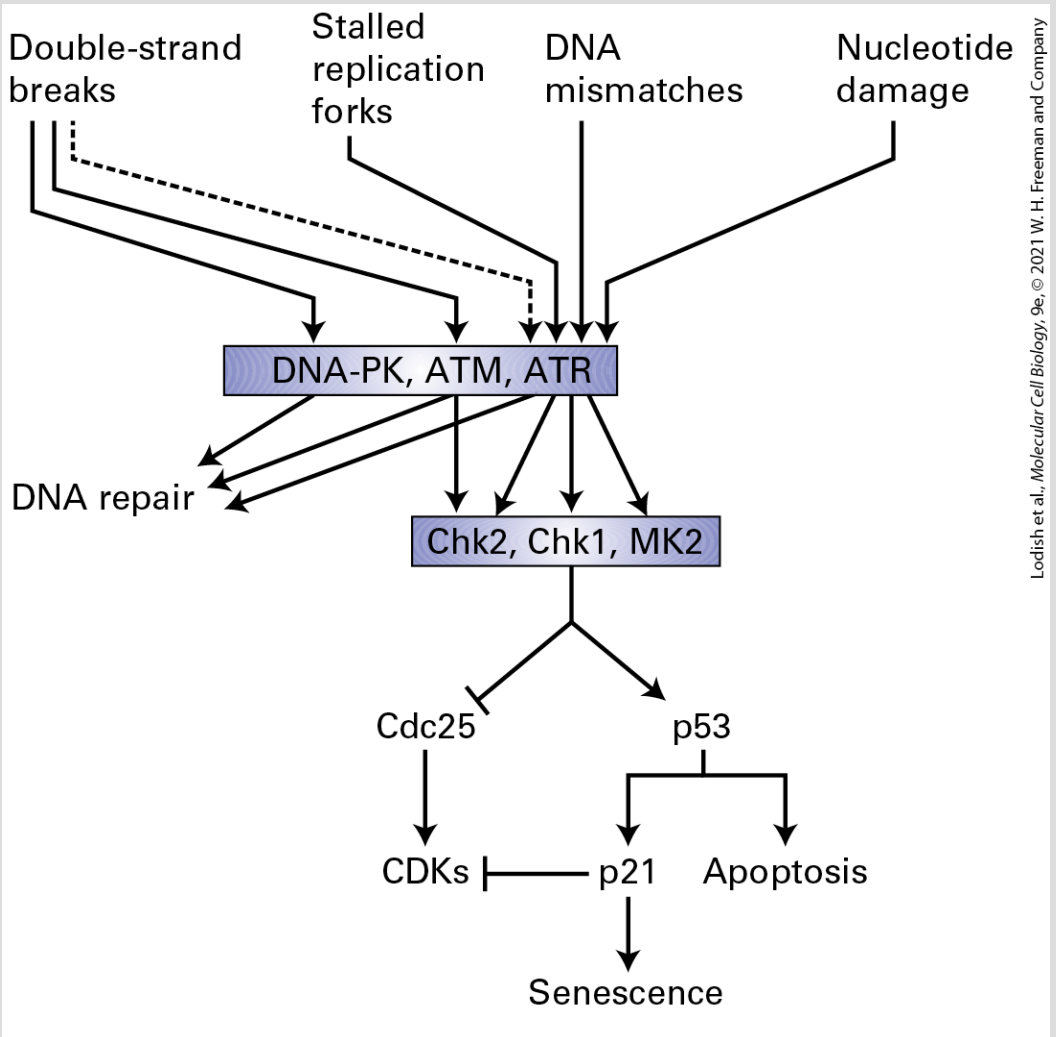


FIGURE 19-32 An overview of the DNA damage response system. The protein kinases DNA-PK, ATM, and ATR are activated at sites of damaged DNA. ATR responds to a variety of DNA damage — most likely to the single-stranded DNA that exists either as a result of the damage itself or as a result of repair. ATM and

DNA-PK are specifically activated by double-strand breaks. Because double-strand breaks are converted into single-stranded DNA as a part of the repair process, they also, albeit indirectly (as depicted by a dashed line), activate ATR. These kinases phosphorylate additional proteins at the DNA lesion resulting in the recruitment and activation of DNA repair machinery. ATM and ATR, once activated by DNA damage, activate a triad of related protein kinases, Chk1, Chk2, and MK2. These kinases cause cell cycle arrest by phosphorylating and thus inhibiting Cdc25 and also by stabilizing the transcription factor p53, which induces the transcription of the CDK inhibitor CKI p21. If the DNA damage is severe, p21 can cause the cells to senesce or p53 can induce apoptosis.

ATR, ATM, and DNA-PK recognize different types of DNA damage. ATM and DNA-PK respond primarily to double-strand breaks. ATR recognizes diverse types of DNA damage, such as stalled replication forks, DNA mismatches, damaged nucleotides, and single-strand breaks. ATR recognizes these diverse types of damage because all of them contain some amount of *single-stranded DNA*, either as part of the damage itself or because repair enzymes create single-stranded DNA as part of the repair process. ATR binding to single-stranded DNA is mediated through an ATR-interacting protein called *ATRIP* and results in increased ATR protein kinase activity. ATR then phosphorylates adapter proteins whose function is to recruit DNA repair proteins and help activate the Chk1 and MK2 kinases that stop progression through the cell cycle until the damage is repaired.

Both DNA-PK and ATM recognize double-strand DNA breaks. Two subunits of DNA-PK, called *Ku70/80*, form a ringlike structure that encircles the free DNA end. In contrast, ATM is directly recruited to

the DNA ends by a protein complex known as the *MRN complex* (Figure 19-33, step **1**). Binding of DNA-PK and ATM to the broken ends helps to hold them together. ATM also activates Chk2 and MK2 to stop progression through the cell cycle until repair is complete. In addition, DNA-PK, ATM, and ATR phosphorylate a variant histone called *H2AX* within the chromatin near the damage site (Figure 19-33, step **2**). Phosphorylated H2AX acts as a landing pad for recruiting additional protein complexes that control DNA repair in a manner that is conceptually similar to other aspects of cell cycle progression and control described in previous sections. That is, by using additional protein phosphorylations, recruiting additional proteins through phospho-binding domains, and marking proteins at the damage site, including H2AX using ubiquitin protein ligases (Figure 19-33, step **3**), as well as other protein modifications like methylation. The end result is the recruitment of two proteins that determine what specific type of DNA repair occurs at the break. One type of repair is *homologous recombination*, as described in Chapter 5. This involves the creation of single-stranded overhangs that in turn recruit and activate ATR and its effectors, further enhancing the DNA damage response. Repair by this process requires the presence of a homologous chromosome and so is restricted to late S and G_2 phases of the cell cycle. This type of repair requires a protein called *BRCA1* (Figure 19-33, step **4**). In contrast, recruitment of another protein called *53BP1* to the DNA break (Figure 19-33, step **5**) suppresses homologous recombination and promotes an alternative repair pathway known as *nonhomologous end joining*. Here the two

broken ends of the DNA are simply glued back together, sometimes with deletion of some of the DNA bases at the break joining site; hence this type of repair can be mutagenic. This end joining process requires the catalytic activity of DNA-PK. Since no homologous DNA sequence is required, nonhomologous end joining is the primary mechanism of repair for cells in G_1 and is also the dominant mechanism of repair in S and G_2, even though homologous DNA from chromosome replication is available. Other than cell cycle stage, we do not know what controls which repair pathway is used at a particular break.

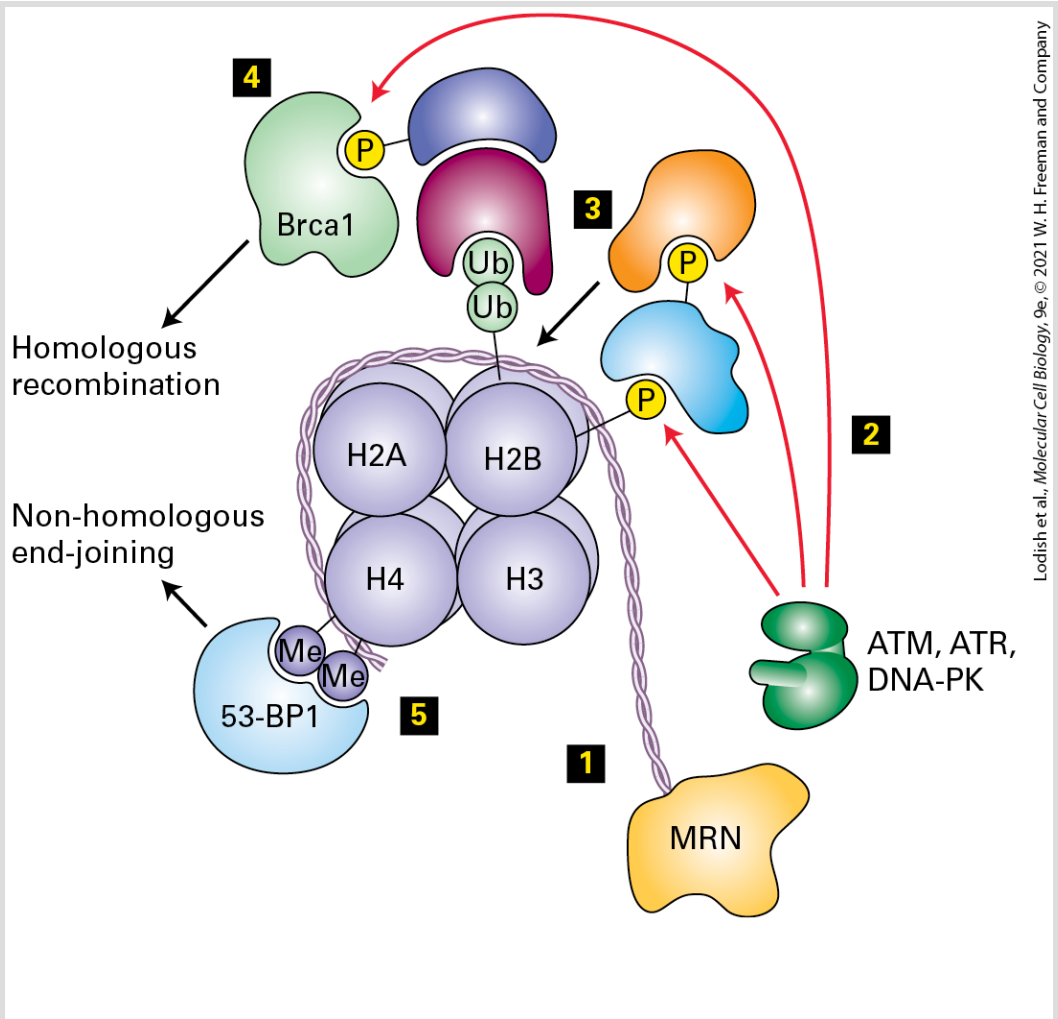


FIGURE 19-33 Activation of cell cycle checkpoints and recruitment of repair machinery at DNA double-strand breaks. The activation and recruitment process depends on a concerted series of protein phosphorylations and ubiquitin modifications, together with recruitment of additional proteins through phosphoserine/threonine-binding domains and ubiquitin-binding modules. Step **1**: The ends of a double-strand break are recognized by the MRN complex, which activates ATM. The MRN complex also has nuclease activity that helps process the broken end for repair. Alternatively, the free double-stranded DNA end can bind to Ku70/80 (not shown) to activate DNA-PK. Step **2**: ATM, ATR, or DNA-PK phosphorylates the variant histone H2AX within nucleosomes located close to the break. This recruits additional proteins that recognize phospho-H2AX through their phospho-binding domains and are themselves then phosphorylated. Step **3**: One of the recruited phospho-binding proteins ubiquitinates H2AX. This allows H2AX to now bind to a ubiquitin-binding protein complex. Step **4**: Phosphorylation of the ubiquitin-binding complex recruits BRCA1 through its phospho-binding BRCT domains, which promotes homologous recombination. Step **5**: Alternatively, methylated histones on histone H4 near the break recruit 53BP1 through its methyl lysine-binding domain. Together with the catalytic activity of DNA-PK (not shown), this suppresses DNA end resection at the break, blocking homologous recombination and promoting nonhomologous end joining.

The identification of BRCA1 as a molecule involved in DNA repair emerged from genetic studies of breast-cancer prone families in whom the development of breast tumors showed a Mendelian-like pattern of inheritance. The genomic locus associated with this trait was localized to a single gene on chromosome 17 by Mary-Claire King and colleagues and designated *BR*east *C*ancer *A*ssociated-*1* (BRCA1). The gene was subsequently cloned by several academic and industrial research teams and shown to be mutated in the patients with breast cancer. Subsequent work by many labs

implicated the BRCA1 protein as a key component of the homologous DNA repair pathway. Cancer-associated mutations in BRCA1 disrupt its DNA repair function. ■

Downstream from ATR and ATM, the kinases Chk1, Chk2, and MK2 halt the cell cycle (see Figure 19-32). These protein kinases inhibit Cdc25 family members by phosphorylating them on sites that are distinct from the CDK-activating phosphorylation sites. These damage-induced phosphorylation sites target Cdc25 proteins to bind to various phosphoserine/threonine-binding proteins (Figure 19-34). When the DNA damage occurs during G_1, phosphorylation of Cdc25A targets it for binding to an F-box–containing ubiquitin-protein ligase (see Figure 19-14), targeting Cdc25A for proteasomal degradation. Loss of Cdc25A inhibits G_1/S phase CDKs and S phase CDKs (see Figure 19-34). As a result, these kinases cannot initiate DNA replication. When the DNA damage occurs during S phase or in G_2, phosphorylation of Cdc25B and Cdc25C by Chk1/2 and MK2 targets them for binding to 14-3-3 proteins. This reduces the catalytic activity of Cdc25B and Cdc25C and sequesters them, preventing them from binding to and activating mitotic CDKs, which results in arrest in G_2 (see Figures 19-34 and 19-21).

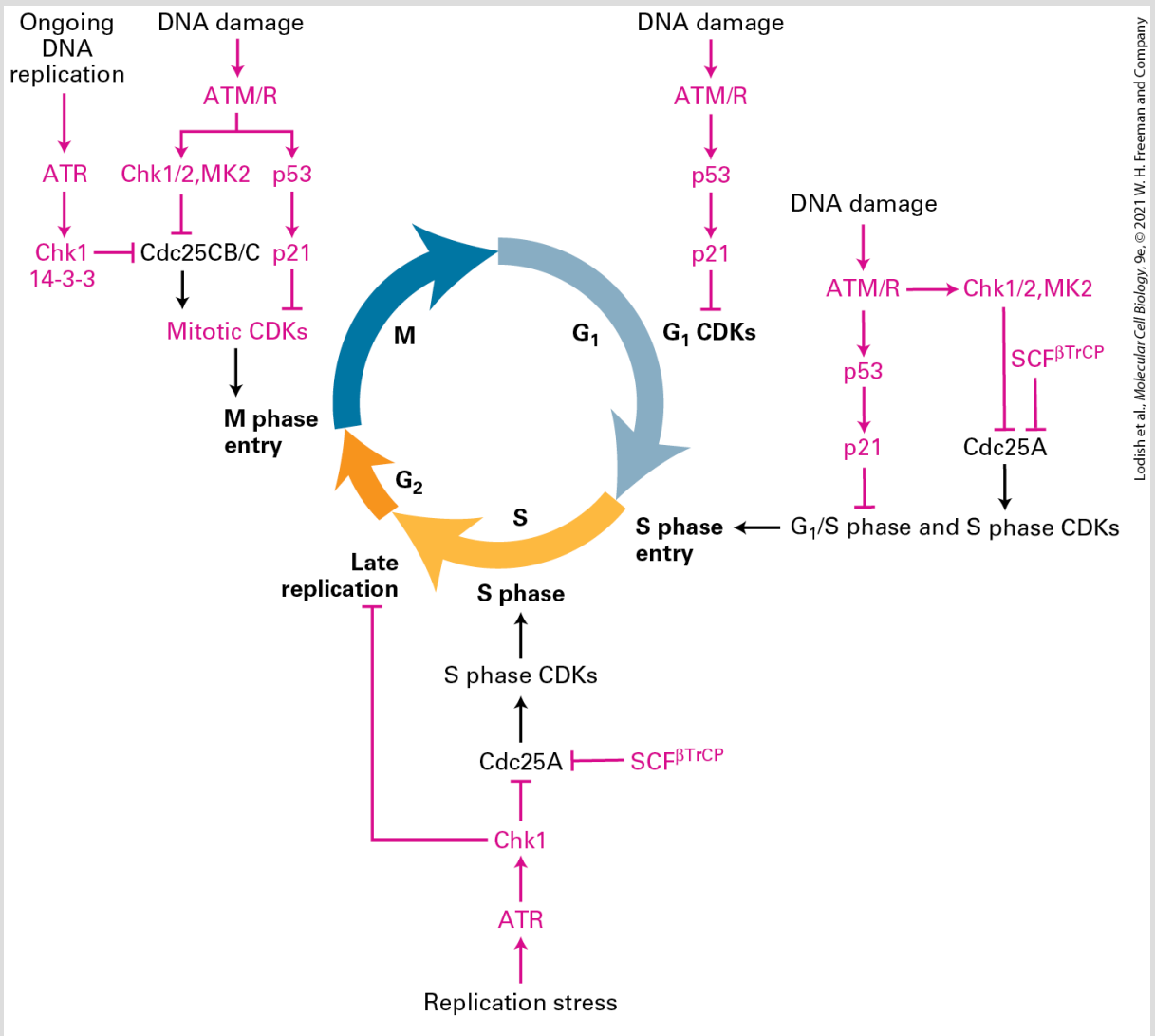


FIGURE 19-34 Overview of DNA damage checkpoint controls in the cell cycle. In response to DNA damage, the ATM or ATR protein kinases (ATM/R) inhibit Cdc25 family members via the Chk1/2 and MK2 protein kinases. They also activate p53, which induces production of the CKI p21. During G_1, the p53-p21 pathway inhibits G_1 CDKs. In S and G_2 replication stress (slow DNA replication fork movement or DNA replication fork collapse), activates the ATR-Chk1 protein kinase cascade, while DNA double-strand breaks activate the ATM-Chk2 protein kinase cascade. Chk1 and Chk2, together with MK2, phosphorylates and inactivates Cdc25B and C, blocking activation of mitotic CDKs and entry into mitosis. During G_1 and S, the DNA damage checkpoint pathways also phosphorylate Cdc25A, targeting it for $SCF^{\beta TrCP}$ -mediated degradation, thereby inhibiting G_1/S phase CDKs and S phase CDKs and blocking entry into or passage through S phase. Red symbols indicate pathways that inhibit progression through the cell cycle.

Even in the absence of DNA damage, active DNA replication also inhibits entry into mitosis. ATR activates Chk1 and MK2 to inhibit Cdc25B/C until all replication forks complete DNA replication and disassemble. This mechanism makes the initiation of mitosis dependent on the completion of chromosome replication. Finally, cells also sense DNA replication stress that results in stalling or slowing of replication forks. Such stress triggers activation of the ATR-Chk1 checkpoint pathway and results in down-regulation of S phase CDK activity and prevents the firing of late-replicating origins. In addition, ATR also stabilizes the stalled fork so that resumption of DNA replication can ensue once the replication stress is resolved.

Chk1-, Chk2-, and MK2-mediated inhibition of Cdc25 phosphatases is not the only means of inhibiting cell cycle progression when there is DNA damage or incomplete replication. DNA damage also leads to the activation of p53, a transcription factor that induces expression of the gene encoding the CDK inhibitor p21. p21 binds to and inhibits all metazoan cyclin-CDK complexes. As a result, cells are arrested in G_1 and G_2 (see Figure 19-34).

The transcription factor **p53** is known as a tumor suppressor because its normal function is to limit cell proliferation in the face of DNA damage. This protein is extremely unstable and generally does not accumulate to high enough levels to stimulate transcription under normal conditions. The instability of p53 results from its ubiquitinylation by a ubiquitin-protein ligase called *Mdm2* and its

subsequent proteasomal degradation. Rapid degradation of p53 is inhibited by ATM and ATR, which phosphorylate p53 at a site that interferes with Mdm2 binding. This and other modifications of p53 in response to DNA damage greatly increase its ability to activate transcription of specific genes that help the cell cope with DNA damage, such as p21.

Under some circumstances, such as when DNA damage is extensive, p53 also activates expression of genes that lead to apoptosis, the process of programmed cell death (see Figure 19-32). In metazoans, the p53 response leading to induce apoptosis (see Chapter 22) evolved presumably to prevent the accumulation of multiple mutations that might convert a normal cell into a cancer cell. The dual role of p53 in both cell cycle arrest and induction of apoptosis may account for the observation that nearly all cancer cells have mutations in the *p53* gene or in the pathways that stabilize p53 in response to DNA damage (see Chapter 25). The consequences of mutations in *p53, ATM*, and *Chk2* provide dramatic examples of the significance of cell cycle checkpoint pathways to the health of a multicellular organism.

The Spindle Assembly Checkpoint Pathway Prevents Chromosome Segregation Until Chromosomes Are

Accurately Attached to the Mitotic Spindle

The *spindle assembly checkpoint pathway* is a global surveillance mechanism that ensures that all of the chromosomes are properly attached to spindle microtubules through amphitelic attachments before entry into anaphase occurs. If even a single kinetochore is unattached or not under tension, then anaphase is inhibited, since any attempt to pull the sister chromatids apart in the presence of such a defect would almost certainly lead to chromosome loss. As we saw in Section 19.5, the spindle assembly checkpoint pathway monitors for kinetochores that are not attached to spindle microtubules. Microtubule-kinetochore interactions that are under insufficient tension are destabilized by the protein kinase Aurora B, which phosphorylates the microtubule-binding component of the kinetochore, the Ndc80 complex. This leads to the loss of stable kinetochore-microtubule interactions that are then recognized by the spindle assembly checkpoint pathway. In this manner, the kinase activity of Aurora B at the level of individual chromosomes and the spindle assembly checkpoint pathway that surveys attachments of all of the chromosomes collaborate during every cell cycle to accurately ensure that every single pair of sister chromatids is attached to the mitotic spindle in the correct, bi-oriented manner.

Components of the spindle assembly checkpoint pathway bind to unoccupied microtubule binding sites at kinetochores and create an

anaphase inhibitory signal that ultimately inhibits APC/C^{Cdc20}. Recall that APC/C^{Cdc20} is the ubiquitin ligase that triggers cyclin B degradation, and the ubiquitinylation and subsequent degradation of securin, triggering the activation of separase to degrade cohesins, and entry into anaphase (see Figure 19-28). When a kinetochore is not attached to microtubules, the outer kinetochore component Knl1 is phosphorylated by the spindle assembly checkpoint pathway protein kinase Mps1 (Figure 19-35). This phosphorylation, in turn, recruits other checkpoint pathway components to the unattached kinetochore. Critical for the shutting down of APC/C^{Cdc20} is the recruitment and activation of the Mad1-Mad2 complex by unattached kinetochores. Importantly, the activated Mad1-Mad2 complex can convert inactive ("open") Mad2 molecules in the cytoplasm to an active ("closed") conformation that is capable of binding to and inhibiting APC/C^{Cdc20}. Mad2 bound to APC/C^{Cdc20} recruits additional checkpoint factors to the complex to form the mitotic checkpoint complex (MCC). The MCC then prevents APC/C^{Cdc20} from recognizing and ubiquitinylating substrates. This spindle assembly checkpoint pathway enables a single unattached kinetochore to inhibit all cellular Cdc20, and thus anaphase, until the kinetochore becomes properly associated with spindle microtubules.

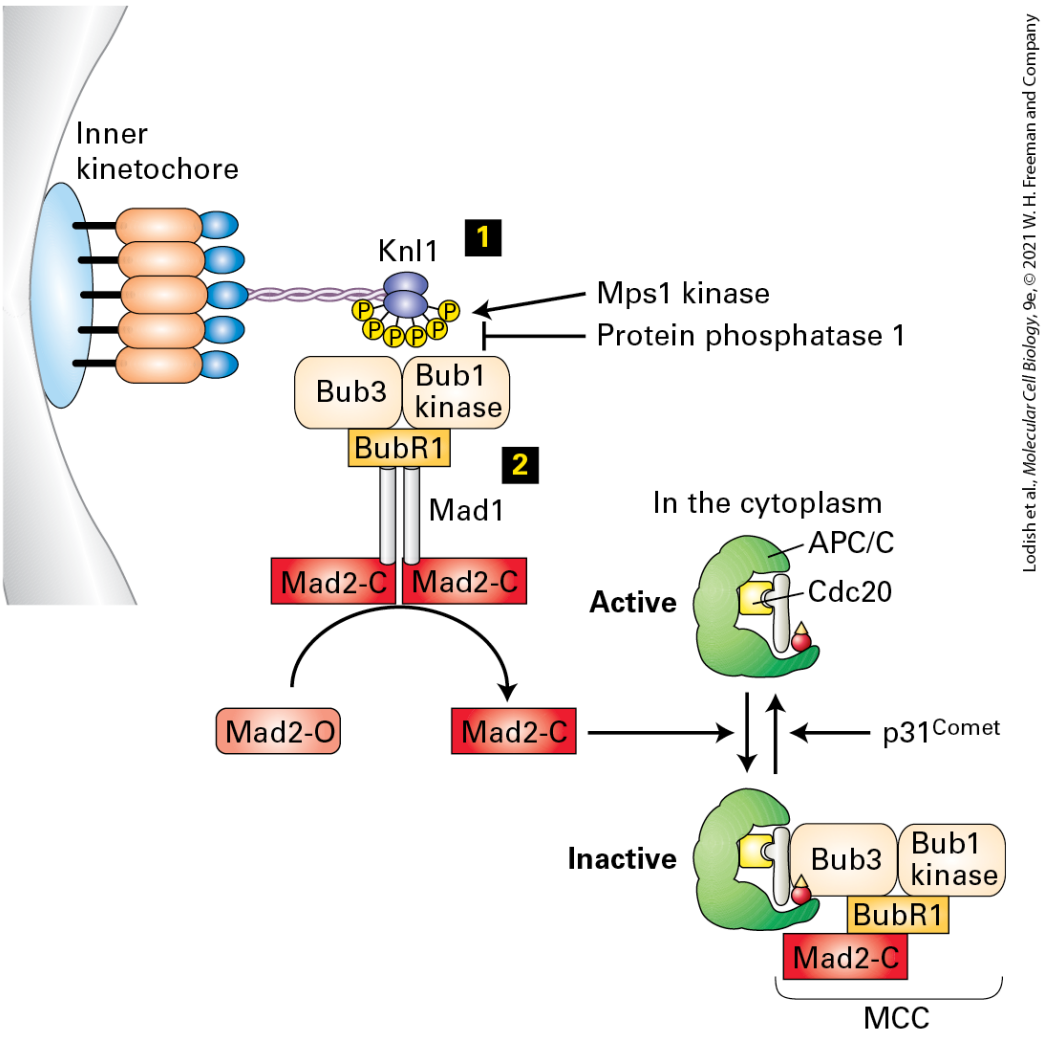


FIGURE 19-35 The spindle assembly checkpoint pathway. The spindle assembly checkpoint pathway is active until every single kinetochore has attached properly to spindle microtubules. Step **1**: When a kinetochore is unattached, the outer kinetochore component Knl1 is phosphorylated by the checkpoint kinase Mps1. Phosphorylated Knl1 then binds the checkpoint protein kinase Bub1-Bub3 and the checkpoint protein BubR1. These three proteins, in turn, recruit the Mad1-Mad2 complex to the kinetochore. The Mad1-Mad2 complex bound to a kinetochore adopts the active closed form (shown as Mad2-C; step **2**) and has the ability to convert inactive open Mad2 (Mad2-O) in the cytoplasm into active closed Mad2 that is able to bind $\mathrm{APC/C^{Cdc20}}$ and inhibit it. Complete inhibition of $\mathrm{APC/C^{Cdc20}}$ requires the recruitment of the checkpoint factors Bub1-Bub3 and BubR1 into the complex. Together, these proteins form the mitotic checkpoint complex (MCC) that prevents $\mathrm{APC/C^{Cdc20}}$

from recognizing and ubiquitinylating its substrates. Silencing of the spindle assembly checkpoint pathway occurs once all kinetochores have attached to microtubules in a tension-generating manner. Protein phosphatase 1 then de-phosphorylates Knl1, thereby eliminating checkpoint protein binding sites at kinetochores. In addition, $\mathrm{p31}^{\mathrm{comet}}$ disassembles the MCC, allowing $\mathrm{APC/C}^{\mathrm{Cdc20}}$ to degrade mitotic cyclins and securin to trigger anaphase.

[Data from E. A. Foley and T. M. Kapoor, 2013, *Nat. Rev. Mol. Cell Biol.* **14**:25–37; and A. Mussachio, 2015, *Curr. Biol.* **25**:R1002–R1018.]

Once all chromosomes have attached to the mitotic spindle in the correct amphitelic manner, the spindle assembly checkpoint pathway must be silenced to allow $\mathrm{APC/C}^{\mathrm{Cdc20}}$ to degrade securin and initiate anaphase. Silencing of the spindle assembly checkpoint pathway occurs through multiple mechanisms. Protein phosphatase 1 de-phosphorylates Knl1, thereby eliminating checkpoint protein-binding sites at kinetochores. In addition, a protein known as $p31^{comet}$ disassembles the MCC, thereby activating $\mathrm{APC/C}^{\mathrm{Cdc20}}$ and initiating anaphase (Figure 19-35).

The spindle assembly checkpoint pathway is essential for viability in mice, highlighting the importance of this quality-control pathway during every cell division. If anaphase is initiated before both kinetochores of a replicated chromosome become attached to microtubules from opposite spindle poles, chromosomes are mis-segregated in a process called **nondisjunction**. The resulting condition, in which cells have either lost or gained whole chromosomes, is called **aneuploidy**. Aneuploidy has profound

effects on health and fitness. It leads to the misregulation of genes and can contribute to the development of cancer. When nondisjunction occurs during the meiotic division that generates a human egg or sperm, trisomy (gain of a chromosome) or monosomy (loss of a chromosome) can be the result. As we will see in Section 19.8, meiosis is especially prone to nondisjunction, which can lead to spontaneous abortions or Down syndrome. ■

KEY CONCEPTS OF SECTION 19.7

Surveillance Mechanisms in Cell Cycle Regulation

- Surveillance mechanisms known as *checkpoint pathways* establish dependencies among cell cycle events and ensure that the next cell cycle event does not occur prior to the completion of a preceding event.
- Checkpoint pathways consist of sensors that monitor a particular cellular event, a signaling pathway, and an effector that halts cell cycle progression and activates repair pathways when necessary. These pathways make extensive use of protein kinases, phospho-binding domains, and ubiquitin-protein ligases.
- Cells are able to detect and respond to a wide variety of DNA damage, and their response depends on what stage of the cell cycle cells are in (see Figure 19-34).
- In response to DNA damage, three related protein kinases, DNA-PK, ATM, and ATR, are recruited to the damage site, where they activate signaling pathways that lead to cell cycle arrest, repair, and under some circumstances, apoptosis. ATM and ATR establish cell cycle checkpoints through the actions of three effector checkpoint kinases Chk1, Chk2, and MK2 (see Figure 19-32).
- The spindle assembly checkpoint pathway, which prevents premature initiation of anaphase, utilizes Mad2 and other proteins to regulate $\mathrm{APC/C}^{\mathrm{Cdc20}}$, which targets securin and mitotic cyclins for ubiquitinylation (see Figure 19-35).

19.8 Meiosis: A Special Type of Cell Division

In nearly all diploid eukaryotes, **meiosis** is the form of nuclear division that generates haploid germ cells — eggs or sperm. An egg or sperm from one individual can then fuse with a germ cell from another individual to generate a diploid zygote that develops into a new individual. Meiosis is a fundamental aspect of the evolution of eukaryotes because it results in the reassortment of the chromosome sets received from an individual's two parents. Chromosome reassortment and homologous recombination between parent DNA molecules during meiosis guarantee that each haploid germ cell will receive a unique combination of alleles that is distinct from each parent and from every other haploid germ cell formed. Meiosis, like mitosis, is preceded by a G_1, S, and G_2 phase.

The mechanisms of meiosis are analogous to those of mitosis. However, several features of meiosis generate genetically diverse haploid cells (see Figure 6-3). In the mitotic cell cycle, each S phase is followed by chromosome segregation and cell division. In contrast, during meiotic cell division, one round of DNA replication is followed by *two consecutive chromosome segregation phases*. This leads to the formation of haploid daughter cells. During the two divisions, maternal and paternal chromosomes are shuffled so that the daughter cells are different in genetic makeup from the parent cell. In this section, we discuss the similarities between mitosis and

meiosis and the meiosis-specific mechanisms that transform the canonical mitotic cell cycle machinery so that it leads to the formation of haploid daughter cells.

Extracellular and Intracellular Cues Regulate Germ Cell Formation

Identifying signals triggering entry into meiosis in metazoans is an active area of research. We know that extracellular signals induce a transcriptional program, which produces meiosis-specific cell cycle factors that bring about the meiotic cell divisions. This modification of the cell cycle goes hand in hand with a program that induces characteristic features of gametes, such as the development of a flagellum in sperm or the production of a stress-resistant cell wall in fungi. One of the extracellular signals inducing entry into meiosis in mammals is retinoic acid. Binding of retinoic acid to the retinoic acid receptor initiates activation of many different developmental processes (see Figure 9-43). The cellular targets of this hormone and how it functions to specify the meiotic fate remain to be discovered.

The molecular mechanisms underlying the decision to enter meiosis are well understood in *S. cerevisiae*. The decision to enter the meiotic divisions is made in G_1. Depletion of nitrogen and carbon sources induces diploid cells to undergo meiosis instead of mitosis, yielding haploid spores (see Figure 1-24). During the meiotic

divisions, budding is repressed. Pre-meiotic S phase and the two meiotic divisions thus occur within the confines of the parent cell. Spore walls are produced around the four meiotic products. Recall that budding and the initiation of DNA replication are induced by G_1/S phase CDKs. Their expression needs to be inhibited to prevent budding. Nutritional starvation represses expression of G_1/S phase cyclins, thereby inhibiting budding. However, DNA replication also relies on G_1/S phase CDKs. How can pre-meiotic DNA replication occur in the absence of G_1/S phase CDKs? The sporulation-specific protein kinase *Ime2* takes over the role of G_1/S phase CDKs in promoting DNA replication. Ime2 promotes (1) phosphorylation of the APC/C specificity factor Cdh1, inactivating it so that S phase and M phase cyclins can accumulate (see Figure 19-17); (2) phosphorylation of transcription factors to induce expression of genes whose products are required for S phase, including DNA polymerases and S phase cyclins (see Figure 19-18); and (3) phosphorylation of the S phase CDK inhibitor Sic1, leading to release of active S phase CDKs and the onset of pre-meiotic DNA replication.

Several Features Distinguish Meiosis from Mitosis

The meiotic divisions differ in several ways from mitotic divisions. These differences are summarized in Figure 19-36. The two

consecutive cycles of cell division are termed *meiosis I* and *meiosis II* (Figure 19-37). Meiosis II resembles mitosis in that sister chromatids are segregated. However, meiosis I is very different. During this division, homologous chromosomes — the chromosome inherited from your mother and the same chromosome inherited from your father — are segregated. This unusual chromosome segregation requires three meiosis-specific modifications to the chromosome segregation machinery. Next we discuss these modifications and explain why they are needed.

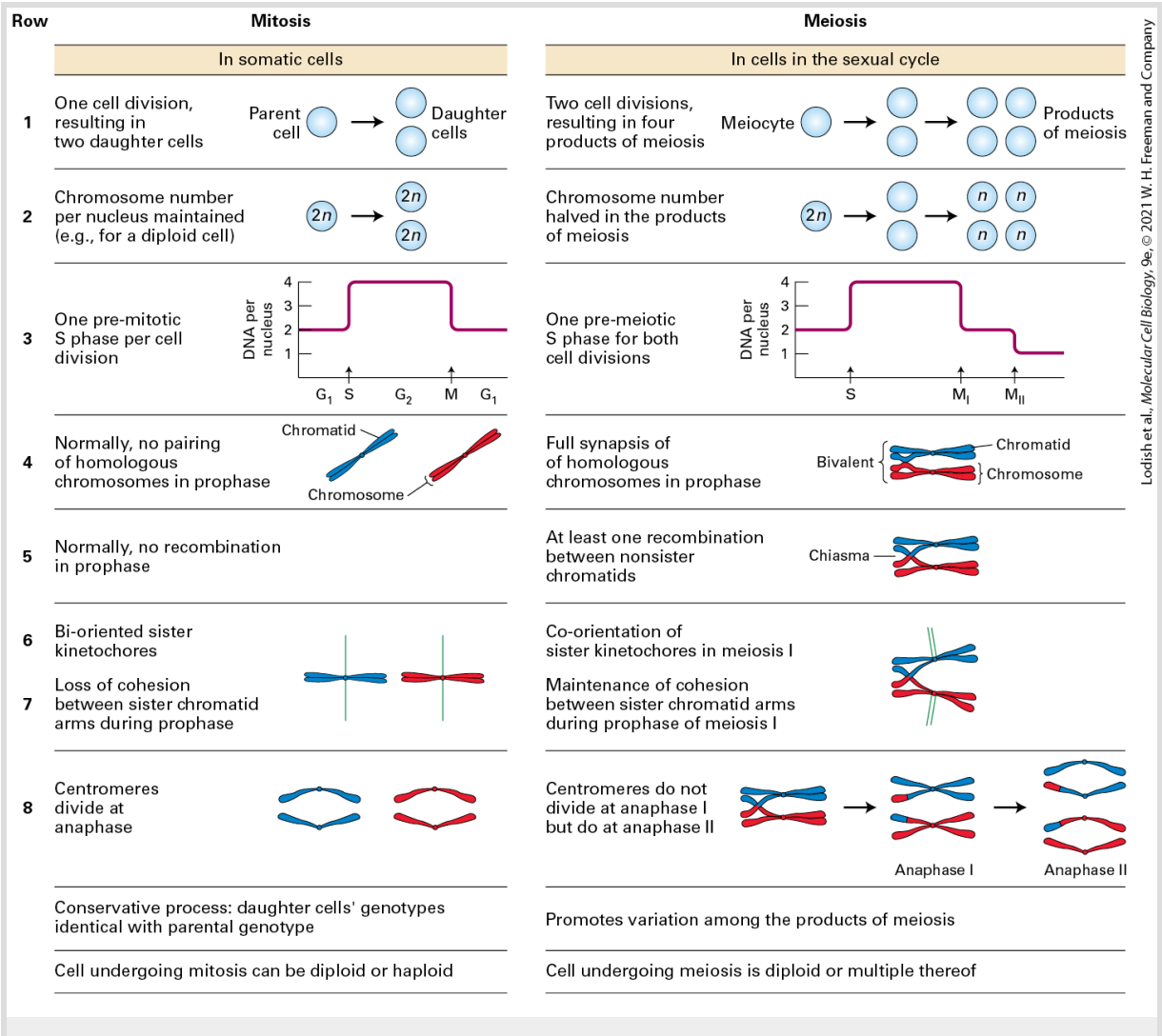
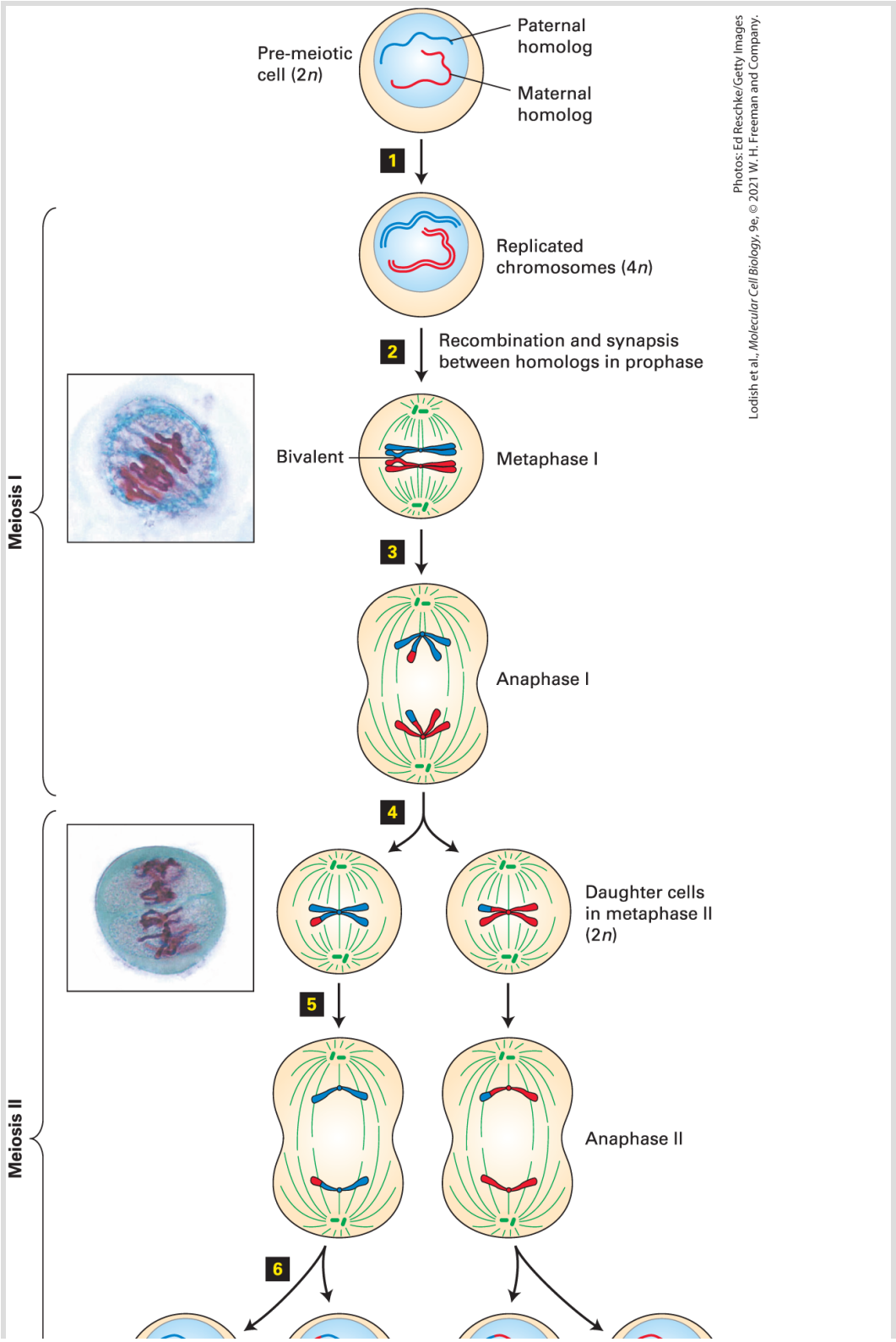


FIGURE 19-36 Comparison of the main features of mitosis and meiosis. See text for details.



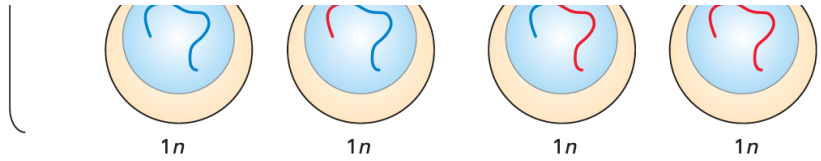


FIGURE 19-37 Meiosis. Pre-meiotic cells have two copies of each chromosome (2*n*), one derived from the paternal parent and one from the maternal parent. For simplicity, the paternal and maternal homologs of only one chromosome are diagrammed. Step **1**: All chromosomes are replicated during S phase before the first meiotic division, giving a 4*n* chromosomal complement. Cohesin complexes (not shown) link the sister chromatids composing each replicated chromosome along their full lengths. Step **2**: As chromosomes condense during the first meiotic prophase, the replicated homologs pair and undergo homologous recombination, leading to at least one crossover event. At metaphase I, shown here, both chromatids of one chromosome associate with microtubules emanating from one spindle pole, but each member of a homologous chromosome pair associates with microtubules emanating from opposite poles. Step **3**: During anaphase of meiosis I, the homologous chromosomes, each consisting of two chromatids, are pulled to opposite spindle poles, and the chiasmata are dissolved as the recombination between homologous chromosomes is completed. Step **4**: Cytokinesis yields two daughter cells (now 2*n*), which enter meiosis II without undergoing DNA replication. At metaphase of meiosis II, shown here, the sister chromatids associate with spindle microtubules from opposite spindle poles, as they do in mitosis. Steps **5** and **6**: Segregation of sister chromatids to opposite spindle poles during anaphase of meiosis II, followed by cytokinesis, generates haploid gametes (1*n*) containing one copy of each chromosome. Micrographs on the left show meiotic metaphase I and metaphase II in developing gametes from *Lilium* (lily) ovules. Chromosomes are aligned at the metaphase plate.

The tension-based sensing mechanism responsible for accurately attaching chromosomes to the spindle during mitosis is also responsible for segregating chromosomes during meiosis I. Thus homologous chromosomes must be linked so that this tension-based

mechanism can function. Homologous recombination between homologous chromosomes creates these linkages (see Figure 19-36). The molecular mechanisms of homologous recombination are discussed in detail in Chapter 4. Here we restrict our discussion to the importance of homologous recombination to successful meiotic divisions.

In G_2 and prophase of meiosis I, the two replicated chromatids of each chromosome are linked together by cohesin complexes along the full length of the chromosome arms, just as they are following DNA replication in a mitotic cell cycle. In prophase of meiosis I, homologous chromosomes (i.e., the maternal and paternal chromosome 1, the maternal and paternal chromosome 2, etc.) pair with each other and undergo homologous recombination. The pairing of homologous chromosomes with each other, a unique feature of meiosis, is called *synapsis*, and is mediated by a proteinaceous complex known as the **synaptonemal complex (SC)** that initially forms at specific points on the chromosomes and spreads all along their lengths. This pairing precisely aligns the genes of homologous chromatids adjacent to one another in order to support homologous recombination. Significantly, at least one recombination event must occur between a maternal and a paternal chromosome. *Crossing over* of chromatids produced by recombination occurs at large multiprotein assemblies, called *recombination nodules*, and can be observed microscopically in the first meiotic prophase and metaphase as structures called *chiasmata*

(singular, *chiasma*). Homologous chromosomes are linked through chiasmata by a meiotic-specific cohesin called *Rec8* and are called *bivalents* (see Figure 19-36). During the later part of prophase, the synaptonemal complex is disassembled as the chromosomes condense until the homologous chromosomes are only attached at the chiasmata and adjacent cohesin molecules, which now provide the resistance to the pulling force exerted by microtubules on the metaphase I spindle (see Figure 19-37). At the meiotic metaphase to anaphase transition, Rec8 is cleaved by separase, much like how the mitotic cohesin Scc1 is cleaved by separase during mitosis, allowing the homologous chromosome pairs to separate.

Recombination between homologous chromosomes that occurs in prophase of meiosis I has two functional consequences: first, it contributes to genetic diversity among individuals of a species by ensuring new combinations of alleles in different individuals; second, it connects homologous chromosomes during meiosis I metaphase. The homologs, now connected through at least one chiasma, must align on the spindle in metaphase of meiosis I so that maternal and paternal chromosomes segregate away from each other during anaphase of meiosis I.

Recombination and a Meiosis-Specific Cohesin Subunit Are Necessary for the

Specialized Chromosome Segregation in Meiosis I

As we have seen, in metaphase of meiosis I, both sister chromatids in one (replicated) chromosome associate with microtubules emanating from the *same* spindle pole, rather than from opposite poles as they do in mitosis (Figure 19-38). Two physical links between homologous chromosomes resist the pulling force of the spindle until anaphase: (1) the chiasmata that result from crossing over between chromatids, and (2) cohesins distal to the crossover point (see Figure 19-39b, *top*). Evidence for the linking function of recombination during meiosis comes from the observation that when recombination is blocked by mutations in proteins essential for the process, chromosomes segregate randomly during meiosis I; that is, homologous chromosomes do not necessarily segregate to opposite spindle poles.

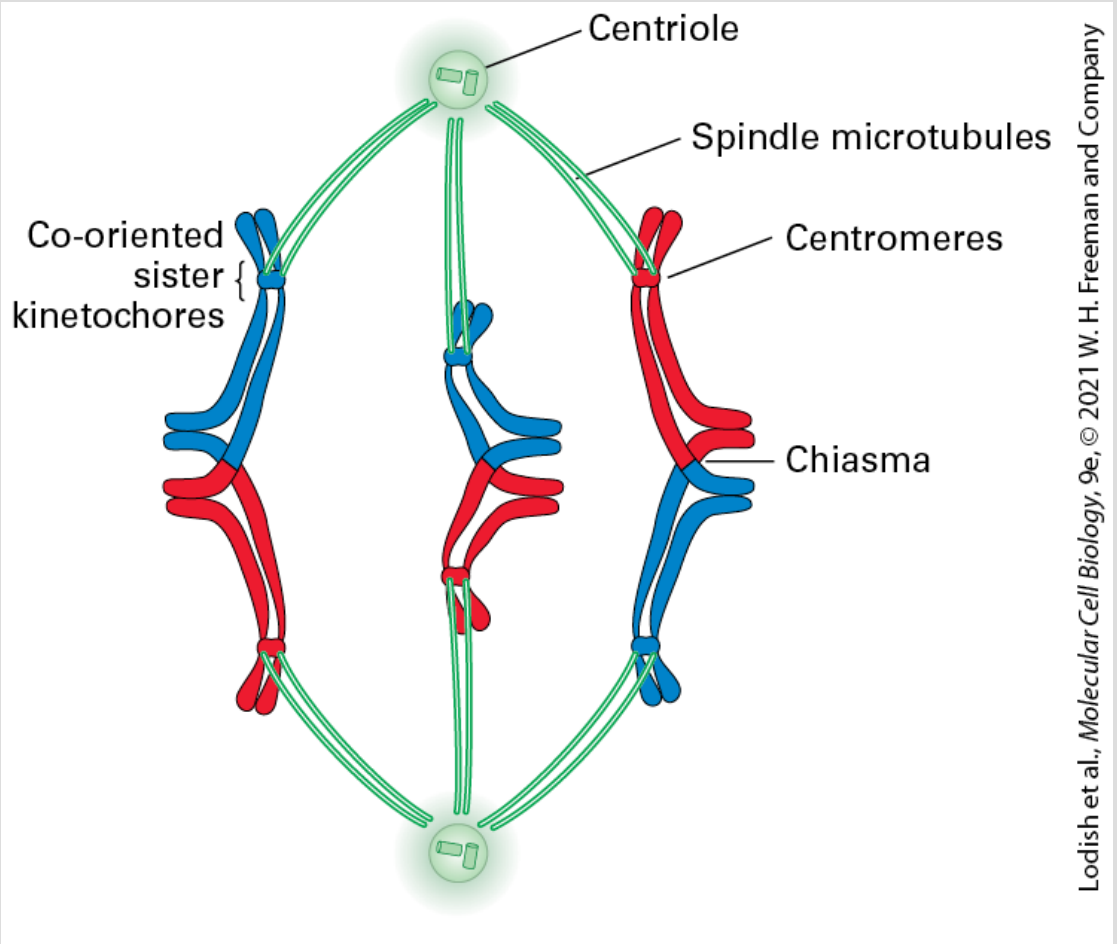
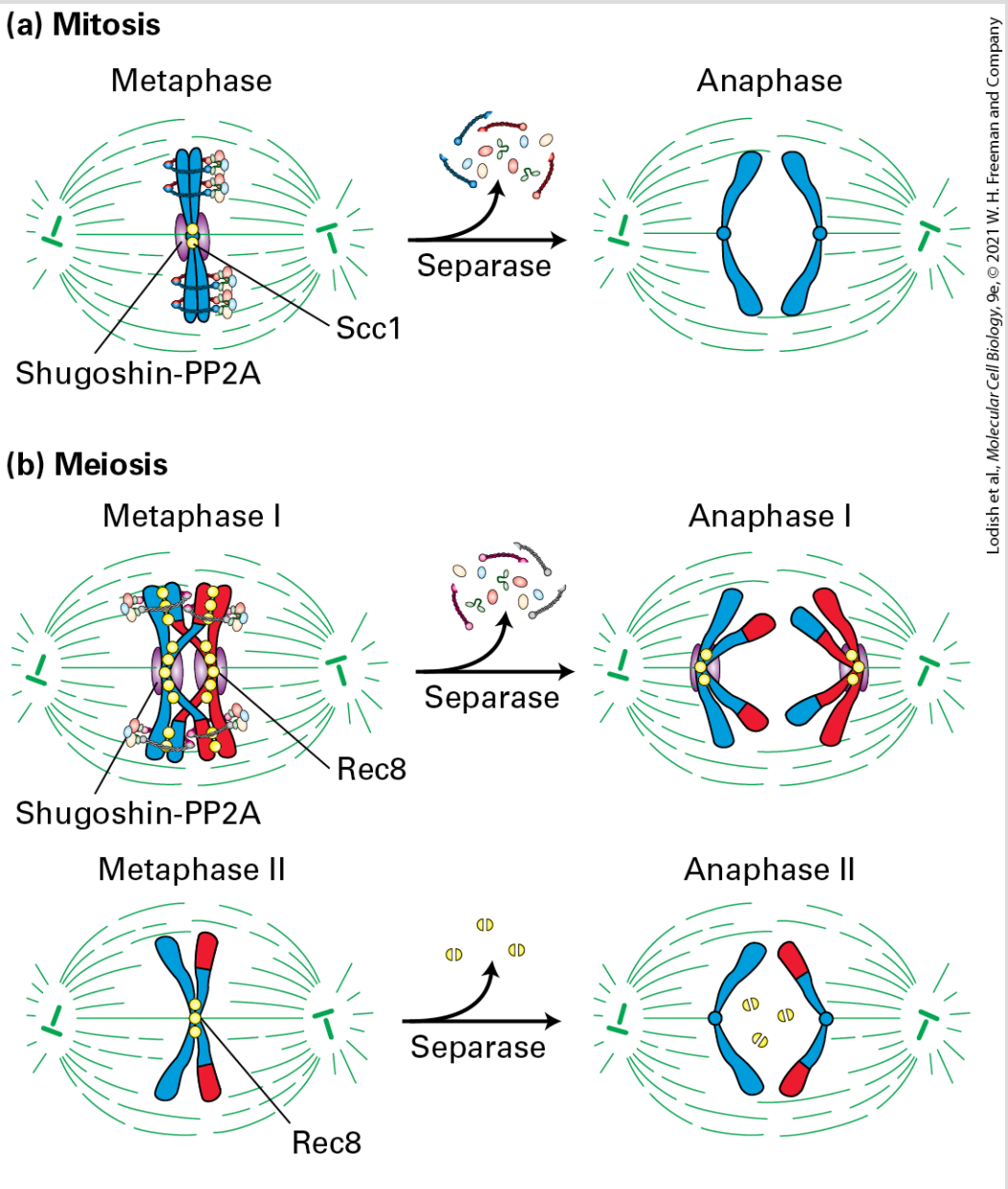


FIGURE 19-38 Chiasmata and cohesins distal to them link homologous chromosomes in meiosis I metaphase. Connections between chromosomes during meiosis I are most easily visualized in organisms with acrocentric centromeres, such as the grasshopper. The kinetochores at the centromeres of sister chromatids attach to spindle microtubules emanating from the same spindle pole, with the kinetochores of the maternal (red) and paternal (blue) chromosomes attaching to spindle microtubules from opposite spindle poles. The maternal and paternal chromosomes are attached to each other by chiasmata, which are formed by recombination between them and by the cohesion between sister chromatid arms that persists until metaphase I. Note that elimination of cohesion between sister chromatid arms by separase cleavage of the meiotic cohesin Rec8 is all that is required for the homologous chromosomes to separate at anaphase. See L. V. Paliulis and R. B. Nicklas, 2000, *J. Cell Biol.* **150**:1223–1232; and S. B. C. Buonomo et al., 2000, *Cell* **103**:387–398.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 19-39 Cohesin function during mitosis and meiosis. (a) During mitosis, sister chromatids generated by DNA replication in S phase are initially linked by cohesin complexes along the length of the chromatids. During chromosome condensation, cohesin complexes (yellow) become restricted to the region of the centromere at metaphase. Shugoshin (purple) recruits PP2A to centromeres, where they antagonize Polo kinase and Aurora B, preventing the dissociation of cohesins from centromeric regions. Dissociation of Shugoshin from centromeres and activation of separase lead to removal of cohesins at the centromere. Sister chromatids now separate, marking the onset of anaphase.

(b) In prophase of meiosis I, maternal and paternal chromatids establish linkages between each other by homologous recombination. By metaphase I, the chromatids of each replicated chromosome are cross-linked by cohesin complexes along their full length. Rec8, a meiosis-specific homolog of Scc1, is cleaved along chromosome arms but not around the centromere, allowing homologous chromosome pairs to segregate to daughter cells. Centromeric Rec8 is protected from cleavage by PP2A, recruited to centromeric regions by the PP2A regulator Shugoshin (shown in purple). By metaphase II, the Shugoshin-PP2A complex dissociates from chromosomes. Cohesins can now be cleaved during meiosis II, allowing sister chromatids to segregate. See F. Uhlmann, 2001, *Curr. Opin. Cell Biol.* **13**:754–761.

At the onset of meiotic anaphase I, cohesins between chromosome arms are cleaved by separase. This cleavage is required for homologous chromosomes to segregate. If cohesins were not lost from chromosome arms, the recombined sister chromatids would be torn apart during anaphase I. The maintenance of centromeric cohesion during meiosis I is necessary for the proper segregation of sister chromatids during meiosis II.

Studies in many organisms have shown that a specialized cohesin subunit, *Rec8*, is necessary for the stepwise loss of cohesins from chromosomes during meiosis. Expressed only during meiosis, Rec8 is homologous to Scc1, the cohesin subunit that closes the cohesin ring in the cohesin complex of mitotic cells (see Figure 19-20). Immunolocalization experiments revealed that during early anaphase of meiosis I, Rec8 is lost from chromosome arms but is retained at centromeres. During early anaphase of meiosis II, centromeric Rec8 is cleaved by separase, so the sister chromatids

can segregate, as they do in mitosis (see Figure 19-39b, *bottom*). Consequently, understanding the regulation of Rec8-cohesin complex cleavage is central to understanding chromosome segregation in meiosis I.

Cohesin removal differs for meiosis I because when Rec8 replaces Scc1 in the cohesin complex, the complex does not dissociate in prophase when it is phosphorylated. The meiotic cohesin complex can be removed from chromatin only by the action of separase. To be cleaved by separase, Rec8, unlike Scc1, requires phosphorylation by several protein kinases. During meiosis I, the centromere-specific isoform of PP2A targeted to centromeric chromatin by Shugoshin prevents this phosphorylation. The PP2A targeting factor and PP2A then dissociate from chromosomes by metaphase II, allowing separase to cleave Rec8.

Meiosis I is much more error-prone than mitosis. It is estimated that 10 percent of all conceptions in humans are aneuploid. These aneuploidies are largely caused by chromosome nondisjunction in meiosis I. When recombination fails to take place between homologous chromosomes, or when the chiasmata are too close to the end of the chromosomes, either no tension or too little tension is exerted on kinetochore–microtubule attachments, and they are destabilized, leading to nondisjunction. As a result, eggs or sperm receive too few or too many chromosomes. All monosomies (lack of a chromosome) and most trisomies (gain of a chromosome) result in

embryonic lethality or death shortly after birth. Only trisomy 21, also known as *Down syndrome,* results in viable offspring, but it causes developmental abnormalities and intellectual disabilities.

Nondisjunction in meiosis I increases dramatically with maternal age. Whereas fewer than 0.1 percent of babies born to women under the age of 30 have Down syndrome, this rate increases to 3.5 percent at age 45. This increase in chromosome nondisjunction with maternal age is not specific to chromosome 21. The incidence of trisomy in clinically recognized pregnancies is less than 5 percent in women under 30 but is as high as 35 percent in women over the age of 42 (Figure 19-40). The reason for this dramatic increase in nondisjunction with maternal age lies in the biology of female meiosis in vertebrates. In all vertebrates, pre-meiotic DNA replication and recombination occur in the female embryo. The oocytes then become arrested in G_2 of meiosis I until the female reaches sexual maturity, which in humans is between the ages of 12 and 16. It is at this time that the first oocyte enters meiosis and progresses to the second meiotic metaphase, where it arrests and awaits fertilization (see Figure 19-5). The oocytes that enter the meiotic divisions in a 40-year-old woman have been arrested in G_2 for 40 years, during which time cohesins that hold the homologous chromosomes together during this very long G_2 phase can deteriorate, causing the homologous chromosomes to dissociate from each other. This can cause the homologous chromosomes to mis-segregate, leading to the formation of aneuploid eggs.

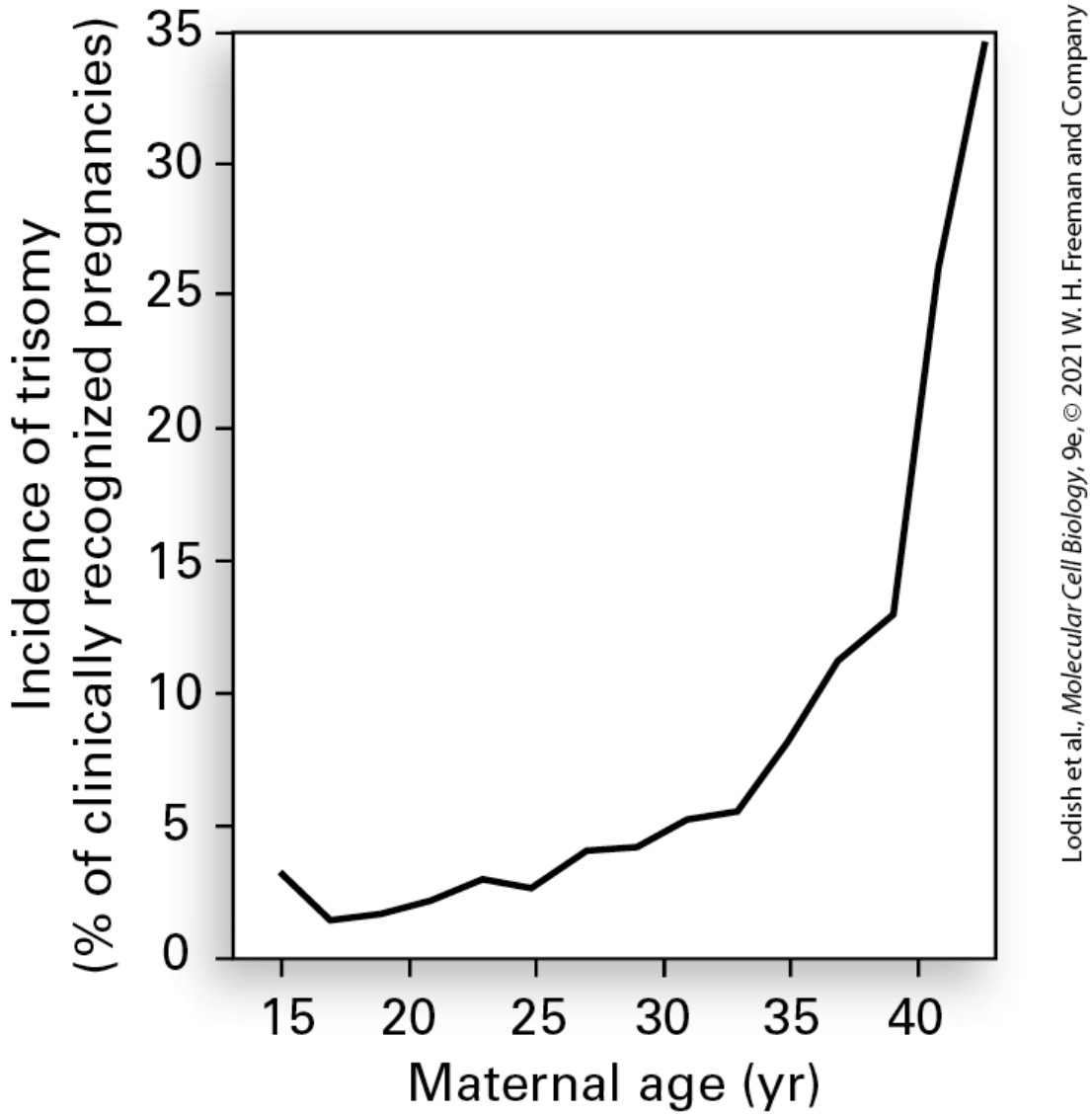


FIGURE 19-40 Fetal aneuploidy increases with maternal age. The percentage of trisomic embryos identified among clinically recognized pregnancies is shown as a function of maternal age.

[Data from T. Hassold and P. Hunt, 2001, *Nat. Rev. Genet.* **2**:280–291.]

Co-orienting Sister Kinetochores Is Critical for Meiosis I Chromosome

Segregation

At metaphase in mitosis and meiosis II, sister kinetochores attach to spindle microtubules emanating from opposite spindle poles; the kinetochores are said to be *bi-oriented*. This is essential for segregation of sister chromatids to different daughter cells. In contrast, at meiosis I metaphase, sister kinetochores attach to spindle microtubules emanating from the same spindle pole; these sister kinetochores are said to be **co-oriented** (see Figure 19-38). Obviously, attachment of sister kinetochores to the proper microtubules in meiosis I and II is critical for correct meiotic segregation of chromosomes.

Proteins required for meiosis I sister kinetochore co-orientation were first identified in *S. cerevisiae*. In this organism, a single microtubule attaches to each kinetochore. We now know that a protein complex known as the **monopolin complex** associates with sister kinetochores during meiosis I and fuses them into a single kinetochore unit, to which one microtubule attaches. In all other organisms, kinetochores attach to multiple microtubules. In vertebrates, Rec8-containing cohesins are essential for sister kinetochore co-orientation. These meiosis-specific cohesins impose a rigid kinetochore structure, restricting the movement of sister kinetochores and thereby favoring their attachment to microtubules from the same spindle pole.

Correct attachment of meiosis I chromosomes is mediated by a tension-based mechanism, as it is during mitosis and meiosis II. During meiotic metaphase I, kinetochore-associated microtubules are under tension (even though the co-oriented kinetochores of sister chromatids attach to microtubules coming from the same spindle pole) because chiasmata generated by recombination between homologous chromosomes and the cohesins distal to the chiasmata prevent them from being pulled to the poles (see Figure 19-38). Since kinetochore-microtubule attachments are unstable in the absence of tension (due to Aurora B–mediated phosphorylation), kinetochores that attach to microtubules from the wrong spindle release those microtubules, which enables them to bind microtubules again until attachments are made that generate tension. As in mitosis, once tension is generated, microtubule attachment to the kinetochores is stabilized.

KEY CONCEPTS OF SECTION 19.8

Meiosis: A Special Type of Cell Division

- Meiosis is a specialized type of cell division in which meiosis-specific gene products modulate the mitotic cell division program (see Figure 19-36).
- The meiotic division comprises one cycle of chromosome replication followed by two cycles of cell division to produce haploid germ cells from a diploid pre-meiotic cell. During meiosis I, homologous chromosomes are segregated; during meiosis II, sister chromatids separate.
- Specialized environmental conditions induce a developmental program that leads to the meiotic divisions.
- During prophase of meiosis I, homologous chromosomes pair with each other and undergo recombination. At least one recombination event must occur between the

chromatids of each pair of homologous chromosomes for proper chromosome segregation (see Figure 19-37).

- Chiasmata and cohesins distal to them are responsible for holding the homologous chromosomes together during prophase and metaphase of meiosis I (see Figure 19-38).
- At the onset of anaphase of meiosis I, cohesins on chromosome arms are phosphorylated and, as a result, cleaved by separase, but cohesins in the region of the centromere are protected from phosphorylation and cleavage. This protection is brought about by a meiosis-specific cohesin subunit and a protein phosphatase that associates with centromeres. As a result, the sister chromatids remain linked to each other during segregation in meiosis I (see Figure 19-39).
- Cleavage of centromeric cohesins during anaphase of meiosis II allows individual chromatids to segregate into germ cells (see Figure 19-39).
- Meiotic cohesins facilitate the attachment of sister kinetochores to microtubules emanating from the same pole during meiosis I.

End of Chapter

Key Terms

amphitelic attachment
anaphase-promoting complex (cyclosome; APC/C)
aneuploidy
ataxia telangiectasia and Rad3-related protein (ATR)
ataxia telangiectasia mutated (ATM)
Aurora kinase
bi-oriented
Cdc14 phosphatase
Cdc25B phosphatase
Cdc25C phosphatase
CDK-activating kinase (CAK)
centriole
centromere
centrosome (spindle pole body)
centrosome disjunction
checkpoint pathway
chromatid
cohesin
condensin
co-oriented
critical cell size

cytokinesis
DNA-PK (DNA-dependent protein kinase)
E2F transcription factor (E2F)
effector
G_1 cyclin-CDK
G_1/S phase cyclin-CDK
kinetochore
meiosis
merotelic attachment
mitogen
mitotic cyclin-CDK
mitotic spindle
monopolin complex
monotelic attachment
nondisjunction
nucleoporin
p53
S phase cyclin-CDKs
SCF
securin
sensor
separase
signaling cascade
sister chromatid resolution
synaptonemal complex (SC)
syntelic attachment
temperature-sensitive mutant
Wee1

Review the Concepts

1. What cellular mechanism(s) ensure that passage through the cell cycle is unidirectional and irreversible? What molecular machinery underlies these mechanism(s)?
2. What types of experimental strategies do researchers employ to study cell cycle progression? How do genetic and biochemical approaches to this topic differ?
3. Tim Hunt shared the 2001 Nobel Prize in Physiology or Medicine for his work in the discovery and characterization of cyclin proteins in eggs and embryos. Describe the experimental steps that led him to his discovery of cyclins.
4. What physiological differences between *S. pombe* and *S. cerevisiae* make them useful yet complementary tools for studying the molecular mechanisms involved in cell cycle regulation and control?
5. In *Xenopus*, one of the substrates of mitotic CDKs is the phosphatase Cdc25. When phosphorylated by mitotic CDKs, Cdc25 is activated. What is the substrate of Cdc25? How does this information help to explain the rapid rise in mitotic CDK activity as cells enter mitosis?
6. Explain how CDK activity is modulated by the following proteins: (a) cyclin, (b) CAK, (c) Wee1, (d) p21.
7. Explain the role of CDK inhibitors. If cyclin-CDK complexes are necessary to allow regulated progression through the eukaryotic cell cycle, what would be the physiological rationale for CDK inhibitors?

8. Cancer cells typically lose cell cycle entry control. Explain how the following mutations, which are found in some cancer cells, lead to a bypass of these controls: (a) overexpression of cyclin D, (b) loss of Rb function, (c) loss of p16 function, (d) hyperactive E2F.
9. The Rb protein has been called the "master brake" of the cell cycle. Describe how the Rb protein acts as a cell cycle brake. How is the brake released in mid to late G_1 to allow the cell to proceed to S phase?
10. A common feature of cell cycle regulation is that the events of one phase ensure progression to a subsequent phase. In *S. cerevisiae*, G_1 and G_1/S phase CDKs promote S phase entry. Name two ways in which they promote the activation of S phase.
11. For S phase to be completed in a timely manner, DNA replication must be initiated from multiple origins in eukaryotes. In *S. cerevisiae*, what role do S phase CDKs and DDKs play to ensure that the entire genome is replicated once and only once per cell cycle?
12. In 2001, the Nobel Prize in Physiology or Medicine was awarded to three cell cycle scientists. Paul Nurse was recognized for his studies with the fission yeast *S. pombe*, in particular for the discovery and characterization of the $wee1^+$ gene. What did the characterization of the $wee1^+$ gene tell us about cell cycle control?
13. What role do phosphoserine/threonine-binding proteins play in controlling entry from G_1 into S phase and entry from G_2 into M phase?

14. Describe how cells know whether sister kinetochores are properly attached to the mitotic spindle.
15. Describe the series of events by which APC/C promotes the separation of sister chromatids at anaphase.
16. Describe how cytokinesis is regulated so that cell cleavage only occurs after chromosome separation.
17. Leland Hartwell, the third recipient of the 2001 Nobel Prize in Physiology or Medicine, was acknowledged for his characterization of cell cycle checkpoint pathways in the budding yeast *S. cerevisiae*. What is a cell cycle checkpoint pathway? When during the cell cycle do checkpoint pathways function? How do cell cycle checkpoint pathways help to preserve the genome?
18. What role do tumor suppressors, including p53, play in mediating cell cycle arrest for cells with DNA damage?
19. Individuals with the hereditary disorder ataxia telangiectasia suffer from neurodegeneration, immunodeficiency, and an increased incidence of cancer. The genetic basis for ataxia telangiectasia is a loss-of-function mutation in the gene encoding ATM (*ATM; a*taxia *t*elangiectasia *m*utated). Besides p53, what other substrate is phosphorylated by ATM? How does the phosphorylation of this substrate lead to inactivation of CDKs to enforce cell cycle arrest?
20. Overall, meiosis and mitosis are analogous processes involving many of the same proteins. However, some proteins function uniquely in each of these cell-division

events. Explain the meiosis-specific function of the following: (a) Ime2, (b) Rec8, (c) monopolin.

21. Explain why the incidence of Down syndrome increases with maternal age.

CHAPTER 20

Integrating Cells into Tissues

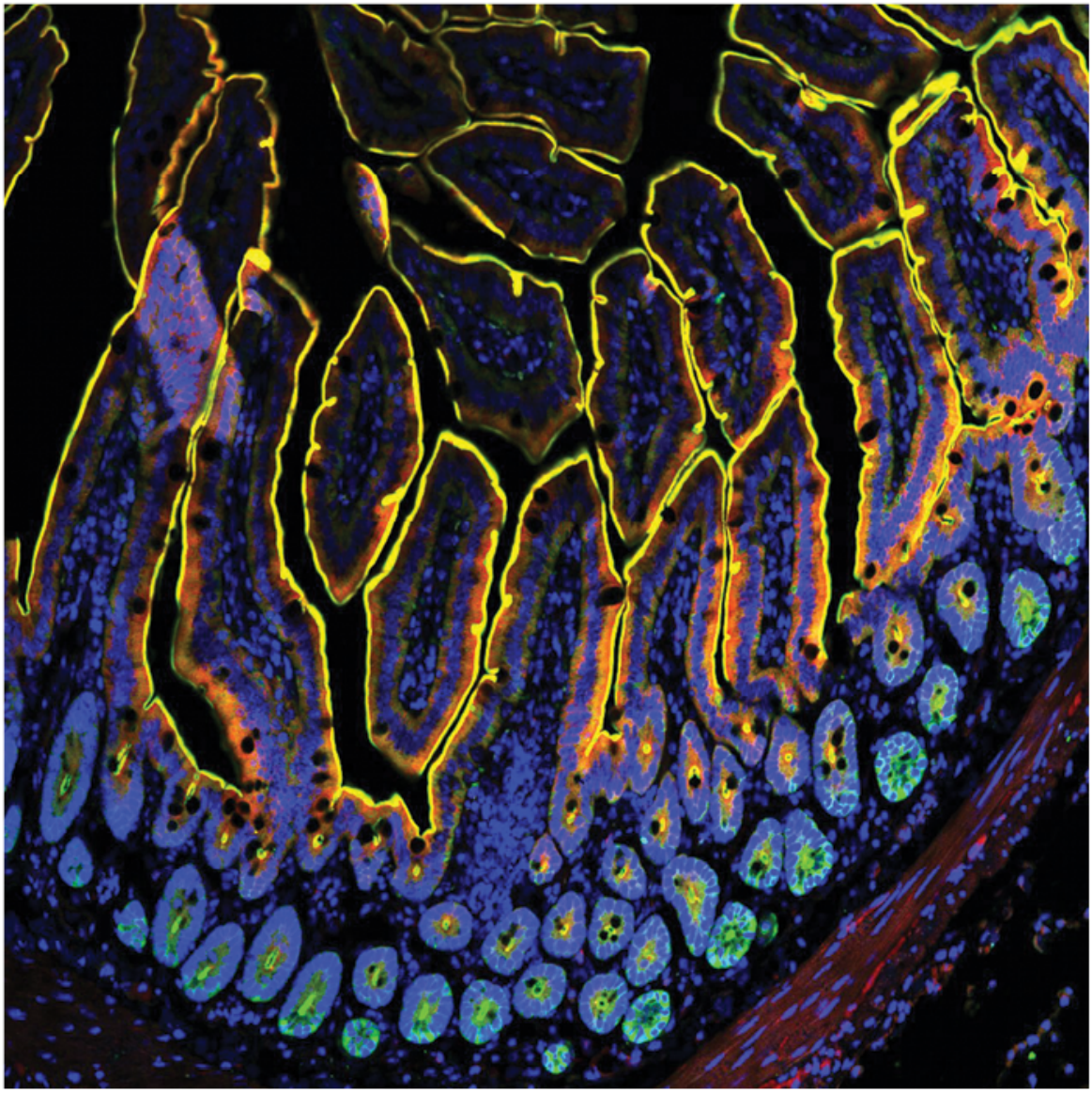


Image courtesy of B. L. Song.

Antibody staining of mouse small intestinal villi, fingerlike projections of cell layers (see Figures 20-13 through 20-15). NPC1L1, a protein located primarily on the plasma membrane involved with cholesterol metabolism, is green. Villin, which binds to actin bundles in

microvilli (very small membrane projections on the apical surfaces of absorptive cells that take up digested nutrients), is red. Nuclei (DNA) are blue. The image shows all colors merged together (colocalization of green and red appears yellow). The yellow lines show how the apical surfaces of the cells face the lumen of the small intestine (black) from which nutrients are absorbed.

OUTLINE

20.1 Cell-Cell and Cell–Extracellular Matrix Adhesion: An Overview
20.2 Cell-Cell and Cell–Extracellular Matrix Junctions and Their Adhesion Molecules
20.3 The Extracellular Matrix I: The Basal Lamina
20.4 The Extracellular Matrix II: Connective Tissue
20.5 Adhesive Interactions in Motile and Nonmotile Cells
20.6 Plant Tissues

In the development of complex multicellular organisms such as plants and animals, progenitor cells differentiate into distinct *types* that have characteristic compositions, structures, and functions. Cells of a given type often aggregate into a *tissue* to cooperatively perform a common function — muscle contracts, neural tissue conducts electric impulses, xylem tissue in plants transports water. Different tissues can be organized into an *organ*, again to perform one or more specific functions. For instance, the muscles, valves, and blood vessels of a heart work together to pump blood. The coordinated functioning of many types of cells and tissues permits the organism to move, metabolize, reproduce, and carry out other essential activities. Indeed, the complex and diverse morphologies of plants and animals are examples of the whole being greater than the sum of the individual parts.

Vertebrates have many hundreds of different cell types, including leukocytes (white blood cells) and erythrocytes (red blood cells), photoreceptors in the retina, fat-storing adipocytes, fibroblasts in connective tissue, and the hundreds of different subtypes of neurons in the human brain. A large-scale international effort, the Human Cell Atlas Project, is using single-cell RNA sequencing and other methods to generate a comprehensive catalogue of all human cell types and subtypes.

Multicellular animals — also called *metazoans* — usually comprise differentiated cells organized into distinct tissues, although there are some less complex exceptions (e.g., sponges). Even rather simple animals exhibit distinctive tissue organization. The adult form of the roundworm *Caenorhabditis elegans* contains a mere 959 cells, yet these cells fall into 12 different general cell types and many distinct subtypes. Despite their diverse forms and functions, animal cells typically can be classified as components of just five main classes of tissue: *epithelial tissue*, *connective tissue*, *muscular tissue*, *neural tissue*, and *blood*. Various cell types are arranged in precise patterns of staggering complexity to generate tissues and organs. The physiological costs of such complexity include increased requirements for information, material, energy, and time during the development of an individual organism. Although these costs of generating and maintaining complex tissues and organs are high, they confer the ability to thrive in varied and variable environments — a major evolutionary advantage.

One of the defining characteristics of animals with complex tissues and organs (most metazoans) is that the external and internal surfaces of most of their tissues and organs — and indeed, the exterior of the entire organism — are built from tightly packed sheet-like layers of cells known as **epithelia**. The formation of an epithelium and its subsequent remodeling into more complex collections of epithelial and nonepithelial tissues is a hallmark of the development of most metazoans. Sheets of tightly attached epithelial cells act as regulatable, selectively permeable barriers, which permit the generation of chemically and functionally distinct compartments in an organism, such as the stomach and bloodstream. As a result, distinct and sometimes opposite functions (e.g., digestion and synthesis) can efficiently proceed simultaneously within an organism. Such compartmentalization also permits more sophisticated regulation of diverse biological functions. In many ways, the roles of complex tissues and organs in an organism are analogous to those of organelles and membranes in individual cells.

The assembly of distinct tissues and their organization into organs are determined by molecular interactions at the cellular level (Figure 20-1). These interactions would not be possible without the temporally, spatially, and functionally regulated expression of a wide array of adhesion molecules. Cell-surface molecules engaged in adhesion are called **adhesion receptors**. Cells in tissues can adhere directly to one another (*cell-cell adhesion*) through specialized

membrane proteins called **cell-adhesion molecules (CAMs)** or **cell-cell adhesion receptors**, which often cluster on the plasma membrane, sometimes into specialized structures called **cell junctions**. In the fruit fly *Drosophila melanogaster*, at least 500 genes (~4 percent of the total) are estimated to be involved in cell adhesion, and in mammals there are over 1000 such genes. Cells in animal tissues also adhere indirectly (*cell-matrix adhesion*) through the binding of **cell-matrix adhesion receptors** in the plasma membrane to components of the surrounding **extracellular matrix (ECM)**, a complex interdigitating meshwork of proteins and polysaccharides secreted by cells into the spaces between them. Some classes of adhesion receptors (e.g., integrins) can mediate both cell-cell adhesion as well as cell-matrix adhesion.

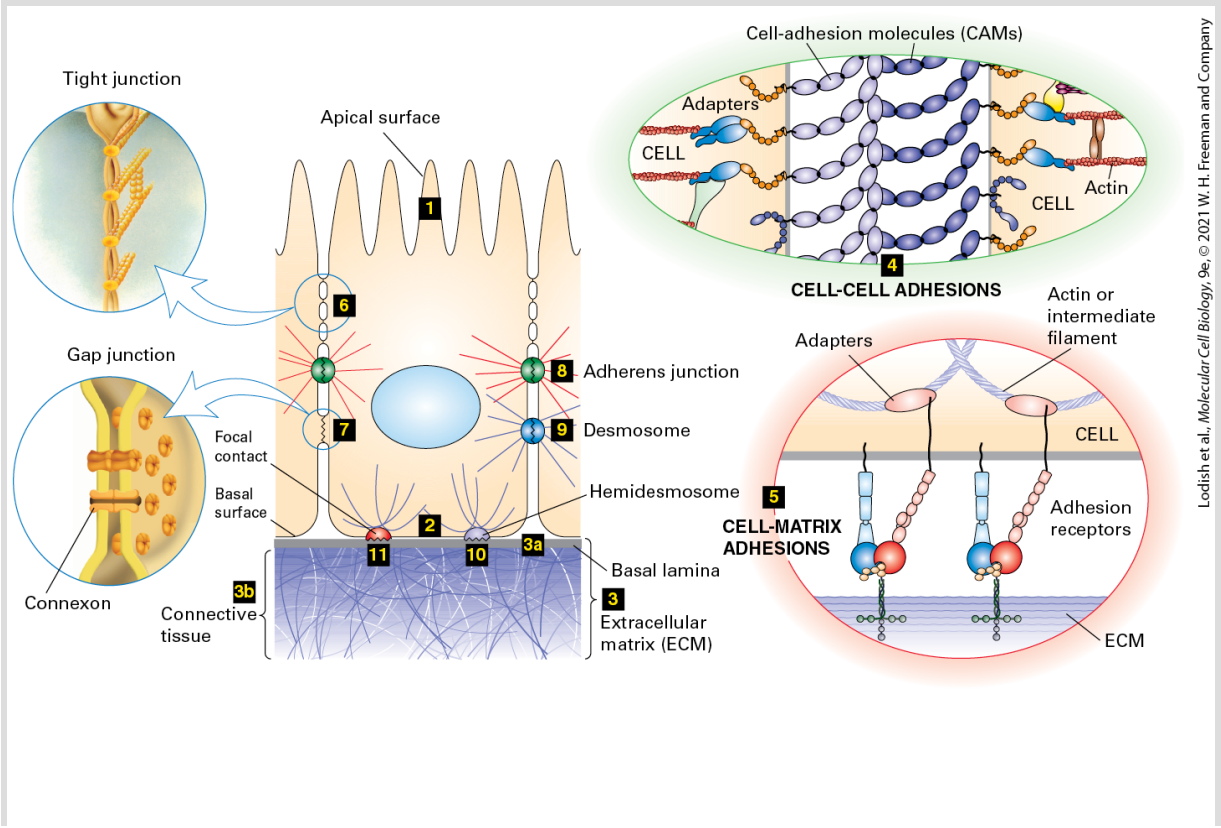


FIGURE 20-1 Overview of major cell-cell and cell-matrix adhesive interactions. Schematic cutaway drawing of a typical epithelial tissue, such as in the inner surface of the intestines. The apical (upper) surface of each cell is packed with fingerlike microvilli (**1**) that project into the intestinal lumen; the basal (lower) surface (**2**) rests on extracellular matrix (ECM). The ECM (**3**) associated with epithelial cells is usually organized into various interconnected layers — such as the basal lamina (**3a**), connecting fibers (not shown), and connective tissue (**3b**) — in which large, interdigitating ECM macromolecules bind to one another and to the cells (**3**). Cell-adhesion molecules (CAMs, also called *cell-cell adhesion receptors*) bind to CAMs on other cells, mediating cell-cell adhesion (**4**), and cell-matrix adhesion receptors bind to various components of the ECM, mediating cell-matrix adhesion (**5**). Both types of adhesion receptors are usually integral membrane proteins whose cytosolic domains often bind to multiple intracellular adapter proteins. These adapters, directly or indirectly, link the CAM to the cytoskeleton (actin or intermediate filaments) and to intracellular signaling pathways (see Figure 20-8). As a consequence, information can be transferred by adhesion receptors and the macromolecules to which they bind from the cell exterior to the intracellular environment and vice versa. In some cases, a complex aggregate of adhesion receptors, adapters, and associated proteins is assembled. Specific, localized aggregates of adhesion receptors form various types of cell junctions, which play important roles in holding tissues together and facilitating communication between cells and their environment. Tight junctions (**6**), lying just under the apical surface, prevent the diffusion of many substances through the extracellular spaces between the cells. Through connexon channels, gap junctions (**7**) allow the movement of small molecules and ions between the cytosols of adjacent cells. The remaining three types of junctions, adherens junctions (**8** and **4**), desmosomes (**9**), hemidesmosomes (**10** and **5**), and focal adhesions (also called *focal contacts*; **11**) link the cytoskeleton of a cell to other cells or to the ECM. See V. Vasioukhin and E. Fuchs, 2001, *Curr. Opin. Cell Biol.* **13**:76–84.

Cell-cell and cell-matrix adhesions not only allow cells to aggregate into distinct tissues, but also provide a means for the bidirectional transfer of information between the exterior and the interior of cells. As we will see, both types of adhesions are intrinsically associated with the cytoskeleton and cellular signaling pathways. As

a result, a cell's surroundings influence its shape and functional properties ("outside-in" effects); likewise, cellular shape and function influence a cell's surroundings ("inside-out" effects). Thus *connectivity* and *communication* are intimately related properties of cells in tissues. Information transfer is important to many biological processes, including cell survival, proliferation, differentiation, and migration. Therefore, it is not surprising that defects that interfere with adhesive interactions and the associated flow of information can cause or contribute to diseases, including a wide variety of neuromuscular and skeletal disorders, heart and vascular defects, clotting abnormalities, cancer, and others.

In this chapter, we examine various types of adhesion molecules found on the surfaces of cells and in the surrounding extracellular matrix. Interactions between these molecules allow the organization of cells into tissues and have profound effects on tissue development, function, and pathology. Many adhesion molecules are members of families or superfamilies of related proteins. While each type of adhesion molecule performs a distinct role, we will focus on the common features shared by members of some of these families to illustrate the general principles underlying their structures and functions. Because of the particularly well-understood nature of the adhesion molecules in tissues that form tight epithelia, as well as their very early evolutionary development, we will initially focus on epithelial tissues, such as the walls of the intestinal tract and the skin. Epithelial cells are normally nonmotile

(immobile), also called *sessile*; however, during development, wound healing, and in certain pathological states (e.g., cancer), epithelial cells can transform into motile cells. Changes in the expression and function of adhesion molecules play a key role in this transformation, as they do in normal biological processes involving cell movement, such as the crawling of white blood cells into sites of infection. We therefore follow the discussion of epithelial tissues with a discussion of adhesion in nonepithelial, developing, and motile tissues.

The evolutionary lineages of plants and animals diverged before multicellular organisms arose (see Figure 1-1). Thus multicellularity and the molecular means for assembling tissues and organs presumably arose independently in animal and plant lineages. Not surprisingly, then, animals and plants exhibit many differences in the organization and development of tissues. For this reason, we first consider the organization of tissues in animals and then deal separately with plants.

20.1 Cell-Cell and Cell–Extracellular Matrix Adhesion: An Overview

There are many different types of cells in the body that dynamically interact with each other in a myriad of ways. These interactions, achieved via adhesion molecules, must be precisely and carefully controlled in time and space and in their physical characteristics (e.g., strength of adhesion) to correctly determine the structures and functions of tissues in a complex organism. It is not surprising, therefore, that cell-cell and cell-matrix adhesion molecules exhibit diverse structures, or that their expression levels vary in different cells and tissues. As a consequence, they mediate the very specific and distinctive cell-cell and cell-matrix interactions that hold tissues together as well as permit essential communication between cells and their environment. We begin this overview with a brief orientation to some of the various types of adhesion molecules present on cells and within the extracellular matrix, their major functions in organisms, and their evolutionary origin. In subsequent sections, we will examine in detail the unique structures and properties of various participants in cell-cell and cell-matrix interactions.

Cell-Adhesion Molecules Bind to One Another and to Intracellular Proteins

Cell-cell adhesion is mediated through membrane proteins called *cell-adhesion molecules* (*CAMs*, also called *cell-cell adhesion receptors*). Many CAMs fall into one of four major families: the cadherins, the immunoglobulin (Ig) superfamily, the integrins, and sugar-binding proteins called *lectins*, one subset of which are called *selectins*. As the schematic structures in Figure 20-2 illustrate, CAMs often comprise multiple, distinct domains, many of which can be found in more than one kind of protein. The functions of these domains vary. Some confer the ability to bind specifically to their partner CAMs on neighboring cells or even to CAMs on the same cell. Some of these domains are present in multiple copies and contribute to the length of the CAMs and thus help define the distance between the plasma membranes of cells bound together by the CAMs. Other membrane proteins, whose structures do not belong to any of the major classes of CAMs in Figure 20-2, are also adhesion receptors and participate in cell-cell and cell-matrix adhesion in various tissues. One example is a set of adhesion G protein–coupled receptors (GPCRs, see Chapter 15). As we will see later, integrins can function both as cell-cell and cell-matrix adhesion receptors. Figure 20-2 depicts an integrin serving as a cell-matrix adhesion receptor that binds to ECM components. Some Ig-superfamily CAMs can play this dual role as well.

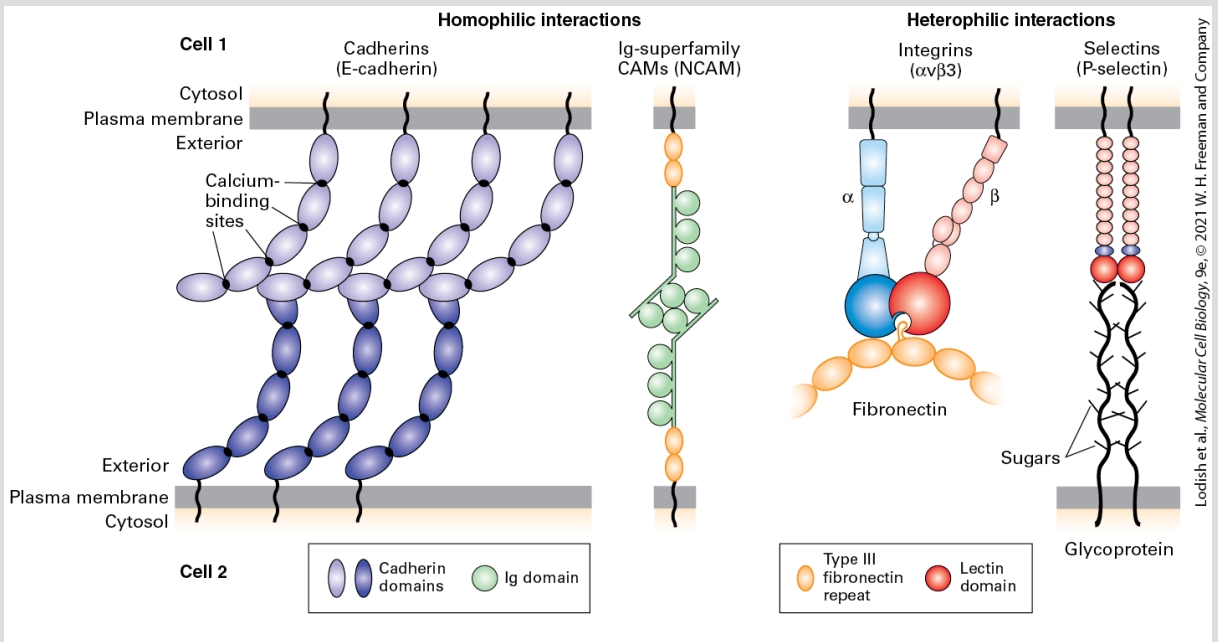


FIGURE 20-2 Major families of cell-adhesion molecules (CAMs) and adhesion receptors. Adhesion receptors can mediate binding, either by homophilic interactions (binding to other molecules of the same type; *left*) or heterophilic interactions (binding to different types of molecules; *right*). E-cadherins on one cell (light blue, cell 1) commonly mediate cell-cell adhesion by forming homophilic cross-bridges with identical E-cadherins on adjacent cells (dark blue, cell 2). E-cadherins also can bind to neighboring E-cadherin molecules on the same cells (see Figures 20-3 and 20-14). Members of the immunoglobulin (Ig) superfamily of adhesion receptors form homophilic linkages (as shown here for NCAM) or heterophilic linkages (to other types of CAMs, not shown). Heterodimeric integrins (e.g., αv and β3 chains) can function as cell-matrix adhesion receptors that bind to very large, multi-adhesive extracellular proteins such as fibronectin, only a small part of which is shown here. Selectins, shown as dimers, contain a carbohydrate-binding lectin domain that recognizes specialized sugar structures on glycoproteins (as shown here) or glycolipids on adjacent cells (heterophilic interactions). Note that CAMs such as E-cadherins, often form higher order oligomers within the plane of the plasma membrane. Many adhesion molecules contain multiple, distinct domains, some of which are found in more than one kind of CAM. The cytoplasmic domains of adhesion receptors are often associated with adapter proteins that link them to the cytoskeleton or to signaling pathways. See R. O. Hynes, 1999, *Trends Cell Biol.* **15**(12):M33–M37; R. O. Hynes, 2002, *Cell* **110**:673–687; and J. Brasch et al., 2012, *Trends Cell Biol.* **22**:299–310.

Through their extracellular domains, CAMs mediate adhesive interactions between cells of the same type (**homotypic adhesion**) or between cells of different types (**heterotypic adhesion**). A CAM on one cell can directly bind to the same kind of CAM on an adjacent cell (**homophilic binding**) or to a different class of CAM (**heterophilic binding**) (see Figure 20-2). CAMs can be broadly distributed along the regions of plasma membranes that contact other cells or clustered in discrete patches or spots that are sometimes called *cell junctions*. Cell-cell adhesions can be tight and long lasting or relatively weak and transient. For example, the associations between neurons in the spinal cord or the metabolic cells in the liver exhibit tight adhesion. In contrast, immune-system cells in the blood often exhibit only brief, weak interactions that allow them to roll along and pass through a blood vessel wall on their way to fight an infection within a tissue.

The cytosolic domains of adhesion receptors recruit sets of multifunctional **adapter proteins** (see Figure 20-1). These adapters act as linkers that directly or indirectly connect adhesion receptors to elements of the cytoskeleton (see Chapters 17 and 18); they can also recruit intracellular molecules that function in signaling pathways (see Chapters 15 and 16) to modify cellular behavior, including gene expression, and the activities of a variety of intracellular proteins including the adhesion receptors themselves. In many cases, a complex aggregate of adhesion receptors, adapter proteins, and other associated proteins is assembled at the inner

surface of the plasma membrane. These complexes facilitate two-way, outside-in and inside-out, communication between cells and their surroundings.

The formation of many cell-cell adhesions entails two types of molecular interactions, called *trans* and *cis* binding interactions (Figure 20-3). **Trans interactions** are also called **intercellular** or *adhesive* interactions, and **cis interactions** are also called **lateral** (in the same cell) interactions. In trans interactions, CAMs on one cell bind to the CAMs on an adjacent cell. In cis interactions, monomeric CAMs on one cell bind to one or more CAMs in the same cell's plasma membrane. For some adhesion molecules the trans interactions between adjacent cells can increase the probability of stable cis interaction formation and thus determine the order in which these interactions occur (trans prior to cis is shown in Figure 20-3). It appears that trans and cis interactions are mutually reinforcing. Furthermore, the association of intracellular molecules with the cytosolic domains of CAMs can dramatically influence the intermolecular interactions of CAMs by promoting their clustering together (e.g., cis interactions) or by altering their conformation in a way that increases the affinity of trans interactions. Just like Velcro, CAMs can generate very tight adhesion when many weak interactions are combined, and this is especially the case when CAMs are concentrated in small, well-defined areas such as cell junctions.

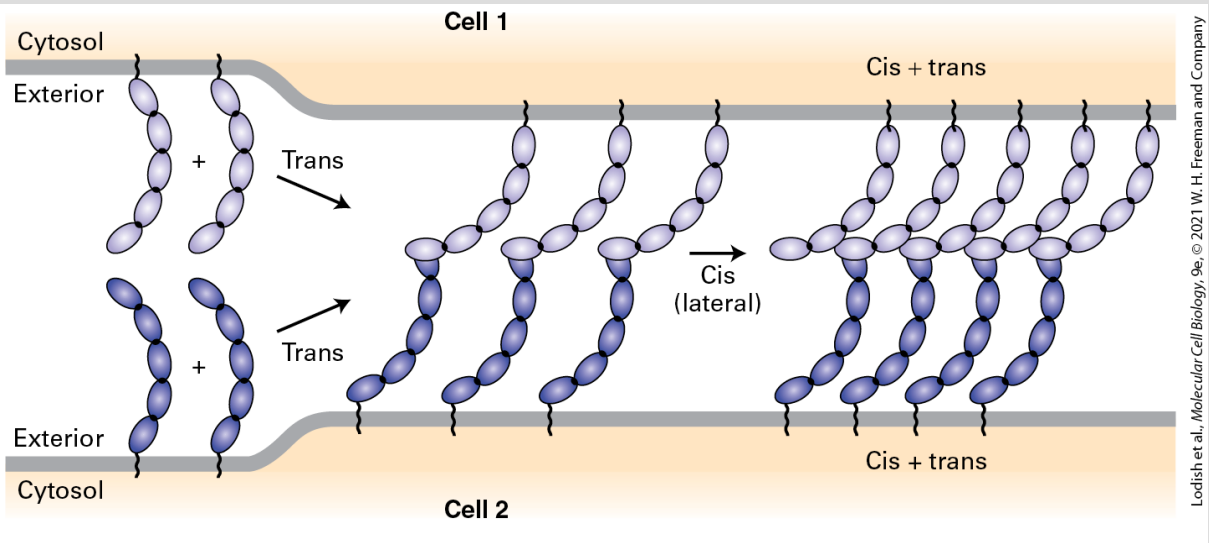


FIGURE 20-3 Model for the generation of cell-cell adhesions. Trans interactions between CAMs on adjacent cells can hold the plasma membranes closely together and permit the subsequent formation of cis (lateral) interactions that collect the CAMs into clusters. The parts of the molecules that participate in these trans and cis interactions vary among the different CAMs. Multiple trans interactions concentrated in space by cis interactions generate a strong, Velcro-like adhesion between the cells. The models shown here are based on CAMs called *classical cadherins* (described in greater detail later in this chapter). See M. S. Steinberg and P. M. McNutt, 1999, *Curr. Opin. Cell Biol.* **11**:554–560; and J. Brasch et al., 2012, *Trends Cell Biol.* **22**:299–310.

Adhesive interactions between cells and the consequences of those interactions on cell structure and function vary considerably, depending on the tissue and the particular CAMs participating. For example, some CAMs require calcium ions to form effective adhesions while others do not. Among the many variables that determine the nature of adhesion between two cells are: the binding affinity of the interacting molecules (thermodynamic properties); the overall “on” and “off” rates of association and dissociation for each interacting molecule (kinetic properties); the spatial

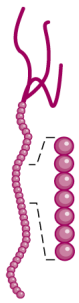
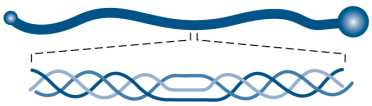
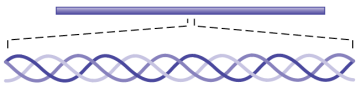
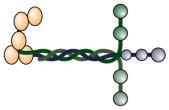
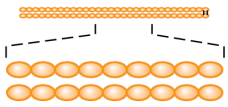
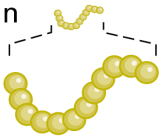
distribution or density of adhesion molecules (ensemble properties); the active versus inactive states of CAMs with respect to adhesion (biochemical properties); and external forces such as stretching and pulling, for example, in muscle or the laminar and turbulent flow of cells and surrounding fluids in the circulatory system (mechanical properties).

The Extracellular Matrix Participates in Adhesion, Signaling, and Other Functions

The ECM is a complex combination of proteins and polysaccharides that is secreted and assembled by cells into a network in which the components bind to one another. The ECM is often involved in holding cells and tissues together. The composition, physical properties, and functions of the ECM are carefully controlled and can vary depending on the tissue type, its location, its physiological state, and chemical modifications of its components, including changes that can occur as a consequence of pathology. These modifications include enzymatic phosphorylation, sulfation and desulfation, cross-linking, cleavage by proteases and glycosidases, and oxidation, as well as nonenzymatic addition of glucose (glycation).

The ECM is usually sensed by cells as a consequence of binding to cell-matrix adhesion receptors on their plasma membranes, which then instruct the cells to behave appropriately in response to their environments or modulate the structure and function of the ECM based on the state of the cells. Different cells can bind to the same patch of ECM via their adhesion receptors and thus be indirectly bound together. ECM components include proteoglycans, a unique type of glycoprotein (a protein with covalently attached sugars); collagens and other proteins that often form fibers; soluble multi-adhesive matrix proteins; and others (Table 20-1). A common feature of many ECM proteins is that they comprise repeating structural domains (sometimes called *repeats*) that form very large proteins. Within one molecule, these repeats can be either very similar (homologous) to other repeats in the protein, distinct from other repeats, or both. The presence of nonidentical repeats that can exhibit distinct binding properties contributes to the multiligand and multifunctional characteristics of many adhesive proteins. For example, multi-adhesive matrix proteins, such as fibronectin and laminin, are long, flexible molecules that contain multiple repeats. They are responsible for binding various types of collagen, other matrix proteins, polysaccharides, and extracellular signaling molecules as well as cell-matrix adhesion receptors. These proteins are important organizers of the extracellular matrix. Through their interactions with adhesion receptors, they also regulate cell-matrix adhesion — and thus cell shape and behavior.

TABLE 20-1 • Extracellular Matrix Proteins

Proteoglycans	Collagens	Multi-adhesive matrix proteins
Perlecan 	Sheet forming (e.g., type IV)  Fibrillar collagens (e.g., types I, II, and III) 	Laminin  Fibronectin  Nidogen/entactin 

Cells contribute to the assembly of the ECM not only by secreting its components, but also by participating directly in the assembly of those components into complex structures containing large fibrils and amorphous macromolecules. Once assembled, the ECM often is not static, but rather highly dynamic in that its chemical, physical, and biological properties can be altered quantitatively or qualitatively as a consequence of cells secreting enzymes, such as proteases, and other molecules into the extracellular space. These alterations in the ECM, which are usually referred to as *remodeling*, can involve covalent chemical modifications (including chemical cross-linking of ECM molecules), partial or essentially complete proteolytic cleavage of ECM components, and addition of newly synthesized ECM molecules. An example we have already encountered, and will discuss in more detail in Section 20.4, is the

role of ECM components in triggering activation of TGF-β by releasing it from its inhibitor (see Figure 16-22).

The relative volumes occupied by cells and their surrounding matrix vary greatly among different animal tissues. Some **connective tissue**, for instance, is mostly matrix with relatively few cells (Figure 20-4a), whereas many other tissues, such as epithelia, are composed of very densely packed cells with relatively little matrix (Figure 20-4b). The density of packing of the molecules within the ECM itself can also vary greatly.

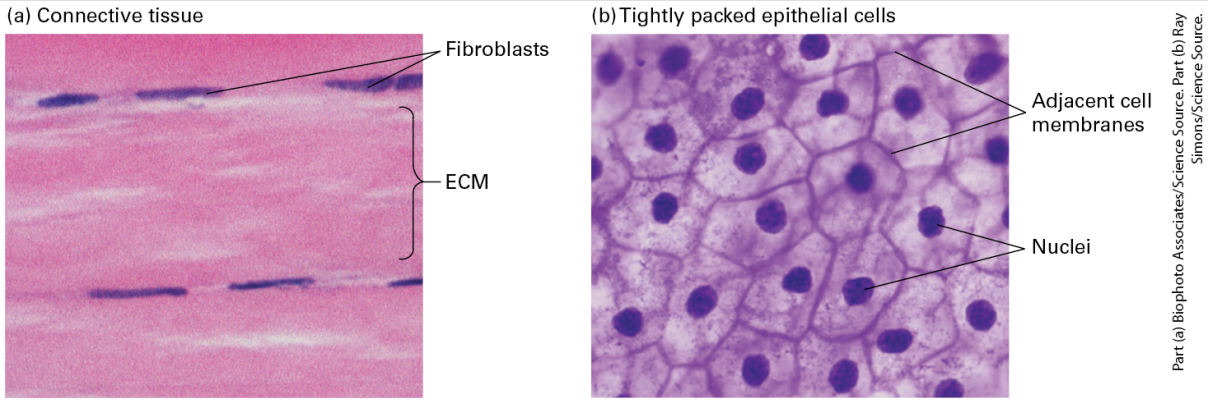
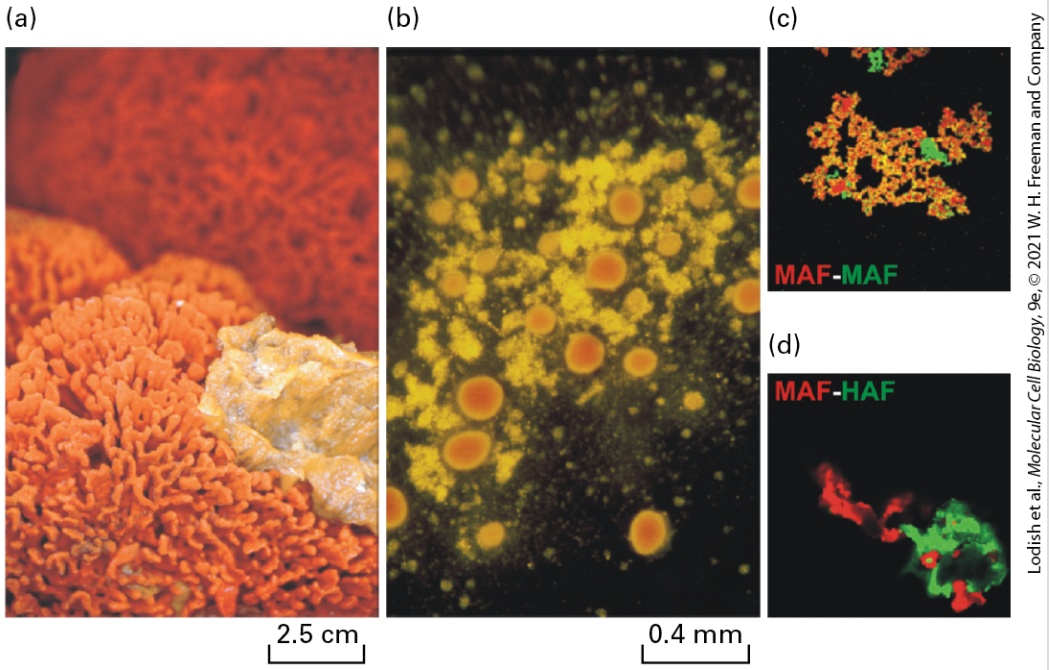


FIGURE 20-4 Variation in the relative density of cells and ECM in different tissues. (a) Dense connective tissue contains mostly extracellular matrix consisting of tightly packed ECM fibers (pink) interspersed with rows of relatively sparse fibroblasts, the cells that synthesized this ECM (purple). (b) Squamous epithelium viewed from the top, showing epithelial cells tightly packed into a quilt-like pattern with the plasma membranes of adjacent cells close to one another and little ECM between the cells (see also Figure 20-10b).

H. V. Wilson's classic studies of adhesion in marine sponge cells showed conclusively that one primary function of the ECM is to

literally hold tissue together. Figures 20-5a and 20-5b, which re-create Wilson's classic work, show that when sponges are mechanically dissociated and individual cells from two sponge species are mixed, the cells of one species will adhere to one another, but not to cells from the other species. This specificity is due in part to species-specific adhesive proteins in the ECM that bind to the cells via adhesion receptors. These adhesive proteins can be purified and used to coat colored beads, which, when mixed, aggregate with one another with a specificity similar to that of intact sponge cells (Figure 20-5c, d).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-5 Mechanically separated marine sponges reassemble through species-specific homotypic cell adhesion. (a) Two intact sponges, *Microciona prolifera* (orange) and *Halichondria panicea* (yellow), growing in the wild. (b) After mechanical disruption and mixing of the individual cells from the two sponge species, their individual cells were allowed to reassociate for about 30 minutes with gentle stirring. The cells aggregated

with species-specific homotypic adhesion, forming clumps of *M. prolifera* cells (orange) and *H. panicea* cells (yellow). In parts (c) and (d) red or green fluorescently labeled beads were coated with the proteoglycan aggregation factor (AF) from the ECM of either *M. prolifera* (MAF) or *H. panicea* (HAF). (c) When beads of both colors were coated with MAF, they all aggregated together, forming yellow aggregates (combination of red and green). (d) MAF (red) and HAF (green) coated beads do not readily form mixed aggregates, but rather assemble into distinct clumps held together by homotypic adhesion. (Magnification 40 ×.)

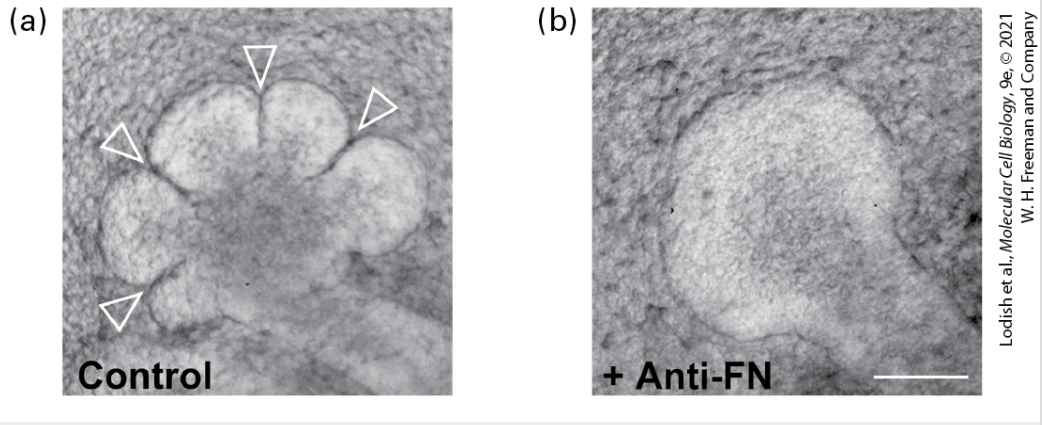
[Parts (a) and (b) republished with permission from Springer, from X. Fernández-Busquets and M. M. Burger, 2003, "Circular Proteoglycans from Sponges: First Members of the Spongican Family," *Cell. Mol. Life Sci*. **60**(1):88–112; permission conveyed through Copyright Clearance Center, Inc. Parts (c) and (d) from J. Jarchow and M. M. Burger, 1998, "Species-Specific Association of the Cell-Aggregation Molecule Mediates Recognition in Marine Sponges," *Cell Adhes. Commun.* **6**(5):405–414; © Taylor and Francis, www.tandfonline.com.]

The ECM plays a multitude of other roles in addition to facilitating cell adhesion (Table 20-2). Different combinations of components tailor the ECM for specific purposes at different anatomic sites, for example, strength in a tendon, strength and rigidity in teeth and bones, cushioning in cartilage, and transparency in the vitreous humor in the eyeball. The composition of the ECM also provides positional and signaling information for cells, letting a cell know where it is and what it should do. ECM remodeling can modulate the interactions of a cell with its environment. Furthermore, the ECM serves as a reservoir for many extracellular signaling molecules that control cell growth and differentiation. In addition, it provides a lattice through or on which cells either can move or are prevented from moving, particularly in the early stages of tissue assembly.

Morphogenesis — the stage of embryonic development in which tissues, organs, and body parts are formed by cell movements and rearrangements — is critically dependent on cell-matrix adhesion as well as cell-cell adhesion. For example, cell-matrix interactions are required for branching morphogenesis (formation of branching structures) to form blood vessels, the air sacs in the lung, mammary and salivary glands, and other structures (Figure 20-6).

TABLE 20-2 • Functions of the Extracellular Matrix

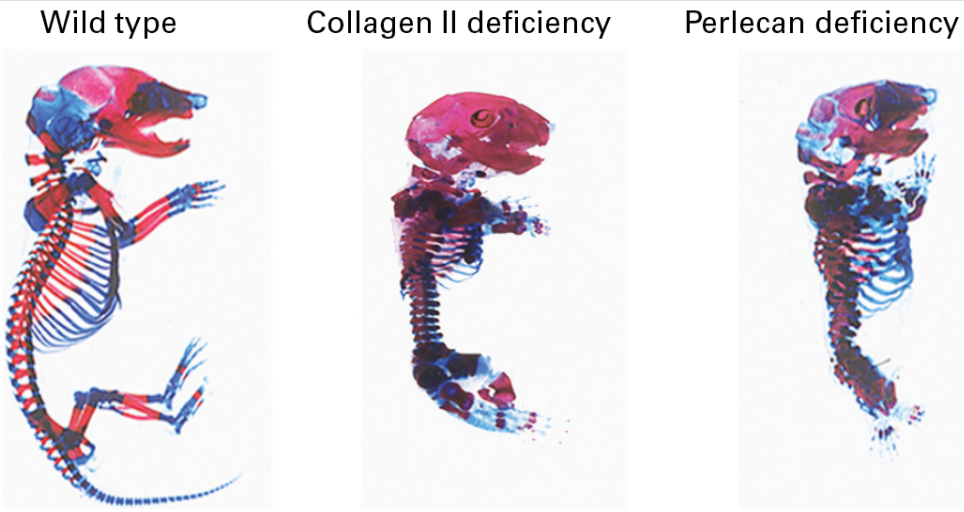
1. Anchoring and surrounding cells to maintain solid-tissue three-dimensional architecture and define tissue boundaries
2. Determining the biomechanical properties (stiffness/elasticity, porosity, shape) of the extracellular environment
3. Controlling cellular polarity, survival, proliferation, differentiation, and fate (e.g., asymmetric division of stem cells; see Chapter 22), and thus embryonic and neonatal development and adult function and responses to the environment and to disease
4. Inhibiting or facilitating cell migration (e.g., serving as either a barrier to movement or, conversely, as a “track” along which cells — or portions of cells — can move)
5. Binding to and acting as a reservoir of growth factors; in some cases, the ECM (a) helps generate an extracellular concentration gradient of the growth factor, (b) serves as a co-receptor for the growth factor, or (c) aids in proper binding of the growth factor to its receptor (ECM component and growth factor jointly serve as a receptor’s combined ligand)
6. Activating cell surface signaling receptors



EXPERIMENTAL FIGURE 20-6 Antibodies to fibronectin block branching morphogenesis in developing mouse tissues. Immature salivary glands were isolated from mouse embryos and allowed to undergo branching morphogenesis in vitro for 10 hours in the absence (a) or presence (b) of an antibody that binds to and blocks the activity of the ECM molecule fibronectin. Anti-fibronectin antibody (Anti-FN) treatment blocked branch formation (arrowheads). Inhibition of fibronectin's adhesion receptor (an integrin) also blocks branch formation (not shown). Scale bar, 100 μm.

[Republished with permission of Nature, from T. Sakai, M. Larsen, and K. M. Yamada, 2003, "Fibronectin Requirement in Branching Morphogenesis," *Nature* **423**(6942):876–881; permission conveyed through Copyright Clearance Center, Inc.]

Disruptions in cell-matrix and cell-cell interactions can have devastating consequences for the development of tissues. Figure 20-7 shows the dramatic changes in the skeletal system of embryonic mice when the genes for either of two key ECM molecules, collagen II and perlecan, are inactivated. Disruptions in adhesion and ECM functions are also characteristic of various pathologies, including cardiovascular, musculoskeletal, kidney, skin, eye, and bone diseases as well as metastatic cancer, in which cancer cells leave their normal locations and spread throughout the body.



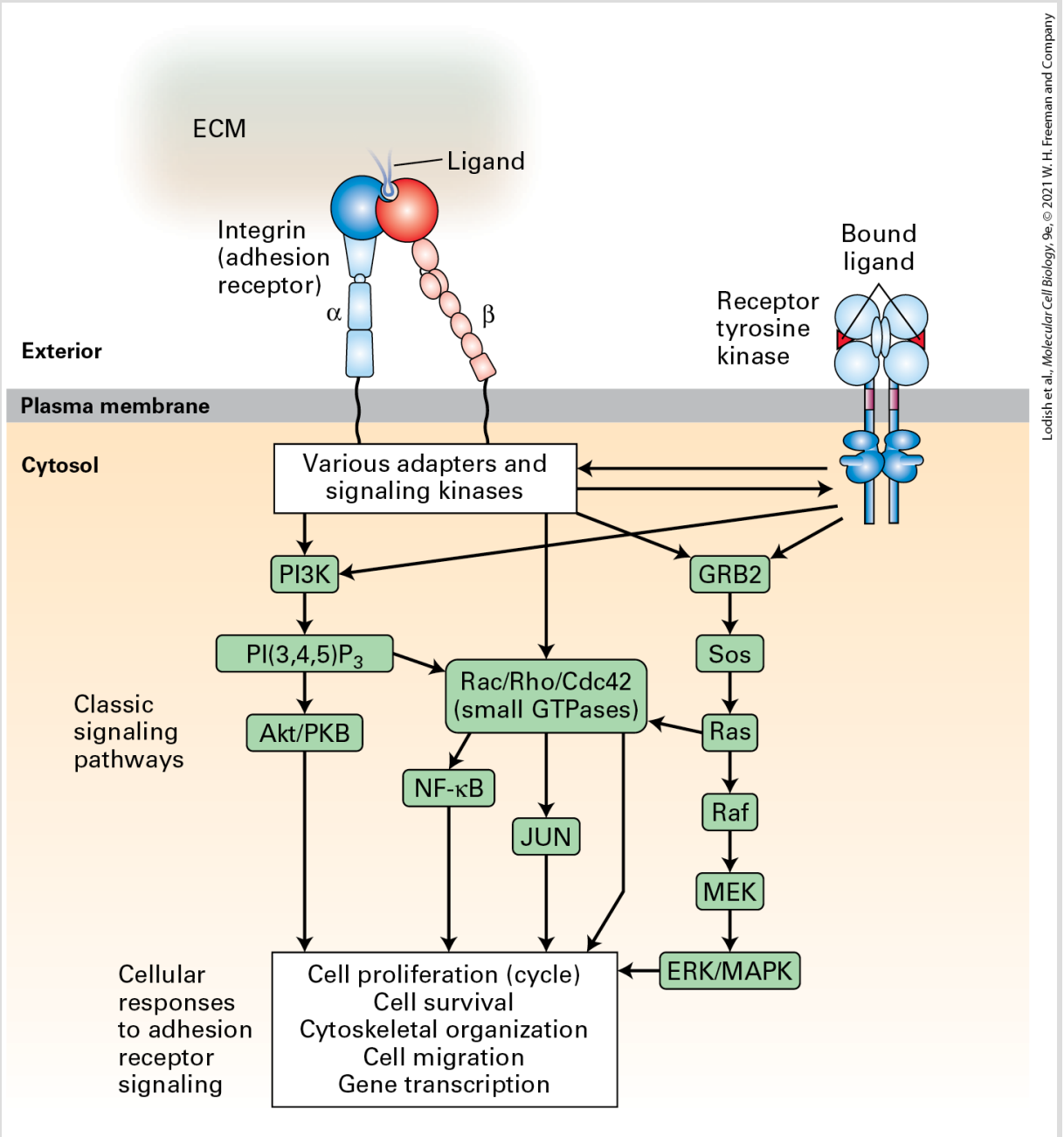
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-7 Inactivating the genes for some ECM proteins results in defective skeletal development in mice. These photographs show skeletons of normal (*left*), collagen II–deficient (*center*), and perlecan-deficient (*right*) murine embryos that were isolated and stained to visualize the cartilage (blue) and bone (red). Absence of these key ECM components leads to forms of short stature (e.g., dwarfism), with many skeletal elements shortened and disfigured.

[Republished with permission from John Wiley & Sons, Inc., from E. Gustafsson et al., 2003, "Role of Collagen Type II and Perlecan in Skeletal Development," *Ann. NY Acad. Sci.* **995**:140–150; permission conveyed through Copyright Clearance Center, Inc.]

Although many CAMs and adhesion receptors were initially identified and characterized because of their adhesive properties, they also play major roles in signaling, using many of the pathways discussed in Chapters 15 and 16. Figure 20-8 illustrates how integrin adhesion receptors physically and functionally interact, via adapters and signaling molecules, with a broad array of intracellular signaling pathways to influence cell survival, gene transcription, cytoskeletal organization, cell motility, and cell proliferation.

Conversely, changes in the activities of signaling pathways inside cells can influence the structures of CAMs and adhesion receptors — for example, by altering adapter binding to the cytosolic portions of the CAMs — and so modulate their ability to interact with other cells and with the ECM. Thus outside-in and inside-out signaling involve numerous interconnected pathways.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-8 Integrin adhesion receptor–mediated signaling pathways control diverse cell functions. Binding of integrins to their ligands induces conformational changes in their cytoplasmic domains, directly or indirectly altering their interactions with cytoplasmic proteins (**outside-in signaling**). These cytoplasmic proteins include adapter proteins (e.g., talins, kindlins, paxillin, vinculin) and signaling kinases [Src-family kinases, focal adhesion kinase (FAK), integrin-linked kinase (ILK)] that transmit signals via diverse signaling pathways, thereby influencing cell proliferation, cell survival, cytoskeletal organization, cell migration, and gene transcription. Components of several signaling pathways, some of

which are associated directly with the plasma membrane, are shown in green boxes. Many of the components of the pathways shown here are shared with other cell-surface-activated signaling pathways (e.g., receptor tyrosine kinases shown on the right) and are discussed in Chapters 15 and 16. In turn, intracellular signaling pathways can, via adapter proteins, modify the ability of integrins to bind to their extracellular ligands (**inside-out signaling**). See W. Guo and F. G. Giancotti, 2004, *Nat. Rev. Mol. Cell Biol.* **5**:816–826; and R. O. Hynes, 2002, *Cell* **110**:673–687.

The Evolution of Multifaceted Adhesion Molecules Enabled the Evolution of Diverse Animal Tissues

Cell-cell and cell-matrix adhesions are responsible for the formation, composition, architecture, and function of animal tissues. Not surprisingly, some adhesion molecules are evolutionarily ancient and are among the most highly conserved proteins in multicellular organisms. Some sponges, the most primitive multicellular organisms, express certain CAMs and multi-adhesive ECM molecules whose structures are strikingly similar to those of the corresponding human proteins. The evolution of metazoans has depended on the evolution of diverse adhesion molecules with novel properties and functions whose levels of expression differ in different types of cells. Some CAMs and adhesion receptors (e.g., cadherins, integrins, and Ig-superfamily CAMs such as L1CAM) and some ECM components (e.g., type IV collagen, laminin, nidogen/entactin, and perlecan-like

proteoglycans) are highly conserved because they play crucial roles in many different organisms, whereas other adhesion molecules are less well conserved. Fruit flies, for example, do not have certain types of collagen or the ECM protein fibronectin, which play important roles in mammals.

The diversity of adhesion molecules arises in large part from two phenomena that can generate the numerous closely related proteins, called **isoforms**, that constitute a protein family. In some cases, the different members of a protein family are encoded by multiple genes that arose from a common ancestor by gene duplication and divergent evolution (see the human β-like globin gene cluster in Chapter 7). In other cases, a single gene produces an RNA transcript that can undergo alternative splicing to yield multiple mRNAs, each encoding a distinct protein isoform (see Figure 7-3 and Section 9.2). Both phenomena contribute to the diversity of some protein families, such as the cadherins. Particular isoforms of an adhesive protein are often expressed in some cell types and tissues, but not others.

Cell-Adhesion Molecules Mediate Mechanotransduction

Mechanotransduction is the reciprocal interconversion of a mechanical force — or stimulus — and biochemical processes. These

interconversions underlie a variety of biological activities, such as signaling, regulated gene expression, cell proliferation, cell migration, and interactions among cells and between cells and the ECM. Evolution has produced a wide variety of **mechanosensors** that respond to a mechanical stimulus by changing their shape and activity. Examples include at least one G protein–coupled receptor that appears to respond directly to extracellular shear stress — a force parallel to the surface of the cell, such as that resulting from the flow of blood in a blood vessel — and a variety of ion channel mechanosensors whose opening is sensitive to mechanical stretch of the plasma membrane. These include certain transient receptor potential (TRP) channels as well as Piezo nonselective cation channels (see Chapter 23), which are found in neurons and non-neuronal cells and that respond to a variety of mechanical stimuli, including touch, airway stretch, blood pressure, and shear and osmotic forces.

Mechanotransduction in the context of cell-cell and cell-matrix interactions usually involves a cell-surface adhesion receptor that transmits mechanical force or biochemical information across the plasma membrane and one or more intracellular or extracellular mechanosensors that respond to the mechanical stimulus by changing shape and activity. For example, tension applied across the length of a multidomain mechanosensor protein, such as the ECM protein fibronectin or the integrin adapter protein **talin**, can separate otherwise tightly packed domains, or even literally pull apart (disassemble) individual domains. As a consequence, one or

more binding sites that were otherwise inaccessible (cryptic) are exposed. Figure 20-9 shows examples of domains that are unfolded by mechanical force, exposing cryptic binding sites. The newly accessible binding sites can then recruit binding partners — in some cases after phosphorylation — and alter cellular or extracellular functions. For example, the stretching of fibronectins by integrins induces domain separation and consequently their assembly into fibrils, which is an early step in the assembly of collagen and other molecules into ECM. The mechanical forces in mechanotransduction can be forces generated within a cell, such as myosin-driven movement of actin filaments (Chapter 17), or outside a cell, such as blood flow, movement of adjacent cells, or contraction or expansion of ECM. As we will see later in this chapter, cell-cell and cell-matrix adhesion sites play particularly important roles in mechanotransduction.

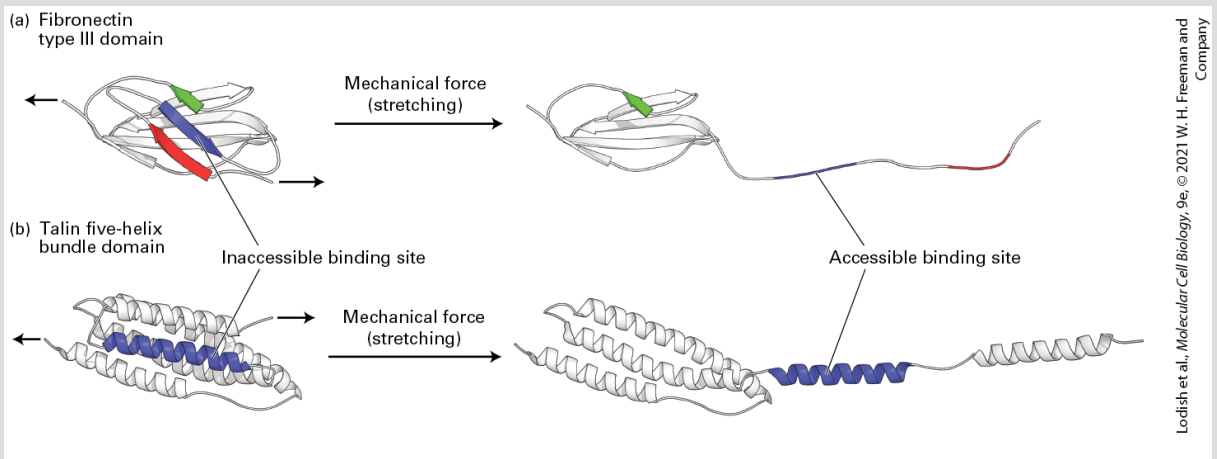


FIGURE 20-9 Models of domains in mechanosensor proteins responding to mechanical forces. (a) Hypothetical model of the partial unfolding of a fibronectin type III domain in the ECM molecule fibronectin when that protein is subjected to mechanical force. Mechanical force generated within the cell by actin movement and mechanotransduced via multiple

integrin adhesion receptors bound to the extracellular dimeric fibronectin can partially unfold the fibronectin. The unfolding is thought to expose a putative, previously hidden binding site on fibronectin (blue segment) that has the potential to form β sheets with other fibronectin molecules, recruiting them to form fibronectin fibrils, and helping assemble the ECM. (b) Hypothetical model of the partial unfolding of a domain (the R1 five-helix bundle) in the intracellular integrin adapter protein talin when it is subjected to mechanical stretching force. This force is generated by actin, which can bind to and pull on the C-terminus of talin while talin's N-terminus is bound to the cytoplasmic tail of integrin's β subunit. The unfolding is thought to expose this domain's otherwise cryptic α-helical vinculin binding site (blue). Vinculin, an actin-binding protein, can then bind to the integrin-talin complex via the exposed site and in turn bind to actin, thus promoting the assembly of multiple actin fibers. The assembly of actin fibers indirectly linked by adapters to integrins strengthens integrin-mediated adhesion and helps to build focal adhesions (see Figures 20-14e and 20-40).

[Part (a) Data from E. P. Gee et al., 2013, *J. Biol. Chem.* **288**:21329–21340; and M. A. Schumacher et al., 2013, *J. Biol. Chem*. **288**:33738–33744. Part (b) Data from M. Yao et al., 2014, *Sci. Rep.* **4**:4610; and E. Papagrigoriou et al., 2004, *EMBO J.* **23**:2942–2951.]

KEY CONCEPTS OF SECTION 20.1

Cell-Cell and Cell–Extracellular Matrix Adhesion: An Overview

- Cell-cell and cell–extracellular matrix (ECM) interactions are critical for assembling cells into tissues, controlling cell shape and function, and determining the developmental fate of cells and tissues. Diseases may result from abnormalities in the structures or expression of adhesion molecules.
- Cell-cell adhesion molecules (CAMs, also called *cell-cell adhesion receptors*) mediate direct intercellular cell-cell adhesions (homotypic and heterotypic), and cell-matrix receptors mediate cell attachment to the ECM (see Figure 20-1). These interactions bind cells into tissues and facilitate communication between cells and their environments.
- The cytosolic domains of cell adhesion receptors bind adapter proteins that mediate interaction with cytoskeletal fibers and intracellular signaling proteins.

- The major families of CAMs are the cadherins, lectins, Ig-superfamily CAMs, and integrins (see Figure 20-2). Members of the integrin and IgCAM superfamilies can also function as cell-matrix adhesion receptors.
- Tight cell-cell adhesions entail both trans (adhesive or intercellular) interactions of like (homophilic) or different (heterophilic) CAMs and cis (lateral) oligomerization of CAMs (see Figure 20-3). The combination of trans and cis interactions produces a Velcro-like adhesion between cells.
- The ECM is a dynamic, complex meshwork of proteins and polysaccharides that contributes to the structure and function of a tissue (see Table 20-2). The major classes of ECM molecules are proteoglycans, collagens, and multi-adhesive matrix proteins, such as fibronectin and laminin.
- CAMs and adhesion receptors, together with their cytoplasmic adapter proteins, play major roles in “outside-in” and “inside-out” signaling, facilitating critically important communication between cells and their surroundings.
- The evolution of adhesion molecules with specialized structures and functions permits cells to assemble into diverse classes of tissues with varying functions.
- Mechanotransduction, the interconversion of a mechanical stimulus or force and biochemical processes, is mediated by adhesion receptors and mechanosensors. Adhesion receptors themselves can also be mechanosensors. Mechanotransduction permits cells to respond to mechanical forces from their environments and to exert mechanical forces on their surroundings.

20.2 Cell-Cell and Cell–Extracellular Matrix Junctions and Their Adhesion Molecules

Cells in epithelial and in nonepithelial tissues use many, but not all, of the same cell-cell and cell-matrix adhesion molecules. We begin our detailed discussion of adhesion with epithelia because of their relatively simple organization, as well as their fundamental role in evolution and development. In this section, we focus on regions of the cell surface that contain clusters of adhesion molecules in discrete patches or spots, called *anchoring junctions*, *tight junctions*, and *gap junctions*. Anchoring and tight junctions play critical roles in mediating cell-cell and cell-matrix adhesion, and all three types of junctions mediate intercellular or cell-matrix communication.

Epithelial Cells Have Distinct Apical, Lateral, and Basal Surfaces

Cells that form epithelial tissues are said to be *polarized* because their plasma membranes are organized into discrete regions. Typically, the distinct surfaces of a polarized epithelial cell are called the **apical** (top), **lateral** (side), and **basal** (base or bottom) surfaces (Figure 20-10; see also Figure 20-1). The apical, lateral, and basal surfaces of epithelial cells can exhibit distinctive characteristics. The area of the apical surface is often greatly expanded by the formation of microvilli — small, fingerlike projections (see Chapter 22). A complex set of evolutionarily conserved molecules, called *polarity regulators*, is required to generate the functional and structural asymmetry of epithelial and other cells (see Chapter 22). Adhesion molecules play essential roles in generating and maintaining these distinct surfaces.

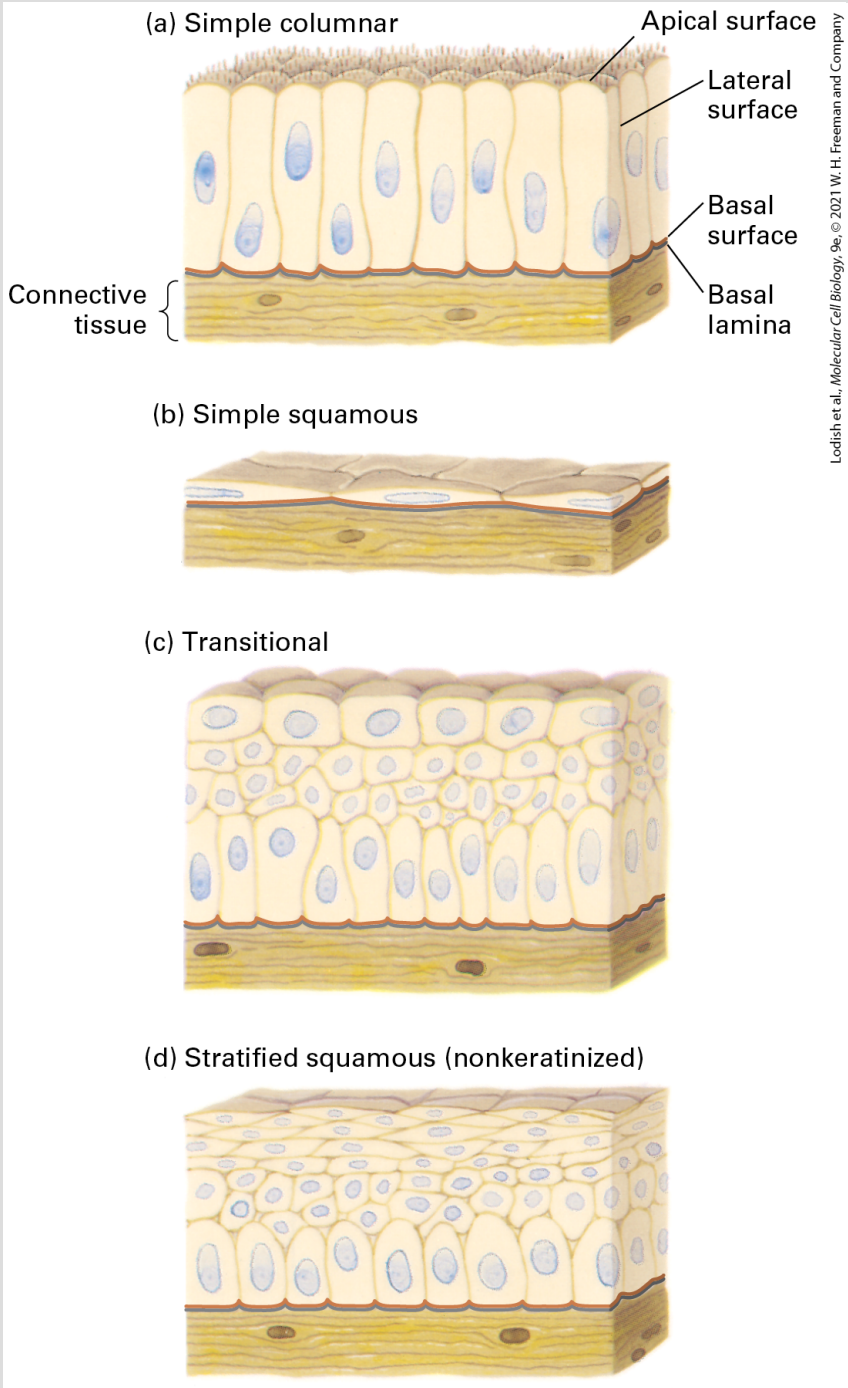


FIGURE 20-10 Principal types of epithelia. (a) Simple columnar epithelia consist of elongated cells, including mucus-secreting cells (in the lining of the stomach and cervical tract) and absorptive cells (in the lining of the small intestine). The thin protrusions at the apical surface are microvilli (see Figure 20-11). (b) Simple squamous epithelia, composed of thin cells, line the blood vessels (endothelial cells/endothelium) and many body cavities. (c) Transitional epithelia, composed of several layers of cells with different shapes, line certain cavities subject to expansion and contraction (e.g., the urinary bladder). (d) Stratified squamous (nonkeratinized) epithelia line surfaces such as the mouth and

vagina; these linings resist abrasion and generally do not participate in the absorption or secretion of materials into or out of the cavity. The basal lamina, a thin fibrous network of collagen and other ECM components, supports all epithelia and connects them to the underlying connective tissue.

Epithelia in different body locations have characteristic morphologies and functions (see Figure 20-10; see also Figure 1-4). Stratified (multilayered) epithelia commonly serve as barriers and protective surfaces (e.g., the skin), whereas simple (single-layered) epithelia often selectively move ions and small molecules from one side of the epithelium to the other. For instance, the simple columnar epithelium lining the stomach secretes hydrochloric acid into the lumen; a similar epithelium lining the small intestine transports products of digestion from the lumen of the intestine across the cells into the blood (see Figure 11-30).

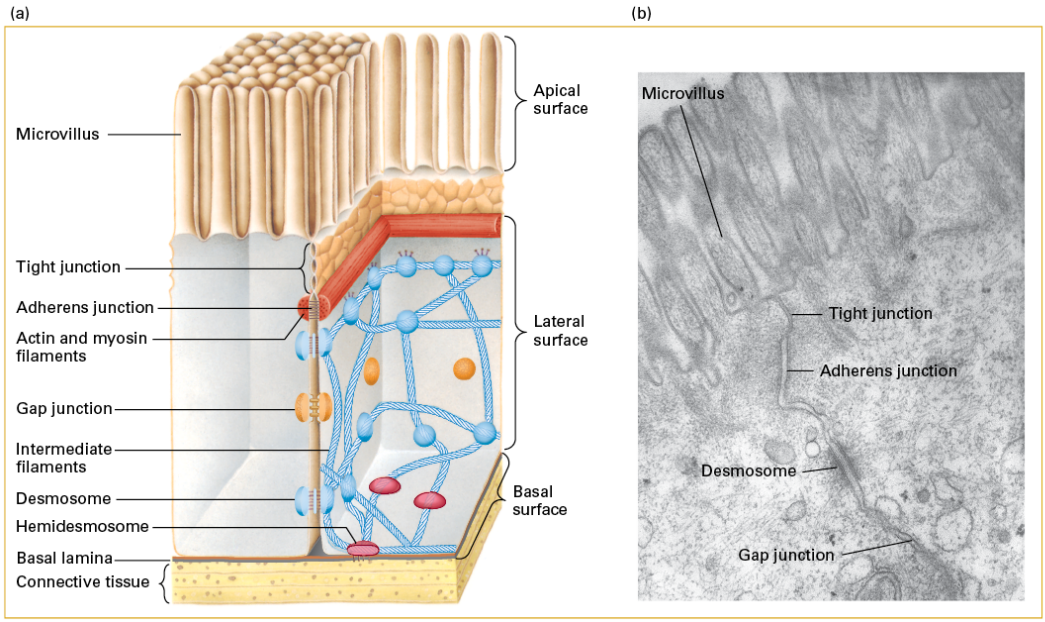
In simple columnar epithelia, adhesive interactions between the lateral surfaces hold the cells together in a two-dimensional sheet, whereas those at the basal surface connect the cells to a specialized underlying extracellular matrix called the **basal lamina**. Often the basal and lateral surfaces are similar in composition and are collectively called the **basolateral** surface. The basolateral surfaces of most simple epithelia are usually on the side of the cell closest to the blood vessels, whereas the apical surface is not in stable, direct contact with other cells or the ECM. In animals with closed circulatory systems, blood flows through vessels whose inner lining is composed of flattened epithelial cells called *endothelial cells*. In general, epithelial cells are sessile, in that adhesion molecules firmly and stably attach them to one another and their associated ECM. One especially important mechanism that generates strong, stable adhesions is the concentration of subsets of these molecules into clusters called *cell junctions*.

Three Types of Junctions Mediate Many Cell-Cell and Cell-Matrix Interactions

All epithelial cells in a sheet are connected to one another and to the ECM by specialized junctions. Although hundreds of individual dispersed adhesion

molecule–mediated interactions are sufficient to cause cells to adhere, the clustered groups of adhesion molecules at cell junctions play special roles in imparting strength and rigidity to a tissue, transmitting information between the extracellular and the intracellular space, controlling the passage of ions and molecules across cell layers, and serving as conduits for the movement of ions and molecules from the cytoplasm of one cell to that of its immediate neighbor. Particularly important to epithelia is the formation of junctions that help form tight seals between the cells, allowing the epithelial sheet to serve as a barrier to the flow of molecules from one side of the sheet to the other.

Three major classes of animal-cell junctions are prominent features of simple columnar epithelia (Figure 20-11 and Table 20-3): **anchoring junctions**, **tight junctions**, and **gap junctions**. Anchoring junctions and tight junctions perform the key task of holding the tissue together. As we shall see, tight junctions also control the flow of solutes through the extracellular spaces between the cells forming an epithelial sheet. Tight junctions are found primarily in epithelial cells, whereas anchoring junctions can be seen in both epithelial and nonepithelial cells. Anchoring junctions and tight junctions in epithelia are organized into three parts: (1) adhesive proteins (CAMs, adhesion receptors) in the plasma membrane that connect one cell to another cell on the lateral surfaces or to the ECM on the basal surfaces; (2) adapter proteins, which connect the CAMs or adhesion receptors to cytoskeletal filaments and signaling molecules; and (3) the cytoskeletal filaments themselves.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-11 Principal types of cell junctions connecting the columnar epithelial cells lining the small intestine. (a) Schematic cutaway drawing of intestinal epithelial cells. The basal surface of the cells rests on a basal lamina, and the apical surface is packed with microvilli that project into the intestinal lumen. Tight junctions, lying just under the microvilli, prevent the diffusion of many substances between the intestinal lumen and internal body fluids (such as the blood) via the extracellular space between cells. Gap junctions allow the movement of small molecules and ions between the cytosols of adjacent cells. The remaining three types of junctions — adherens junctions, desmosomes, and hemidesmosomes — are critical to cell-cell and cell-matrix adhesion and signaling. (b) Electron micrograph of a thin section of epithelial cells in the rat intestine, showing the relative locations of the different junctions.

[Part (b) M. G. Farquhar and G. E. Palade, 1963, *J. Cell Biol.* **17**(2):375–412, Fig. 1; https://doi.org/10.1083/jcb.17.2.375.]

TABLE 20-3 • Cell Junctions

Junction	Adhesion Type	Principal CAMs or Adhesion Receptors	Cytoskeletal Attachment	Intracellular Adapters	Function
Anchoring junctions					
1. Adherens junctions	Cell-cell	Cadherins	Actin filaments	Catenins, vinculin	Shape, tension, signaling, force transmission

2. Desmosomes	Cell-cell	Desmosomal cadherins	Intermediate filaments	Plakoglobin, plakophilins, desmoplakins	Strength, durability, signaling
3. Hemidesmosomes	Cell-matrix	Integrin (α6β4)	Intermediate filaments	Plectin, dystonin/BPAG1	Shape, rigidity, signaling
4. Focal contacts, fibrillar, and 3-D adhesions	Cell-matrix	Integrins	Actin filaments	Talin, kindlin, paxillin, vinculin	Shape, signaling, force transmission, cell movement
Tight junctions	Cell-cell	Occludin, claudins, JAMs	Actin filaments	ZO-1,2,3, PAR3, cingulin	Controlling solute flow, cell polarity, signaling
Gap junctions	Cell-cell	Connexins, innexins, pannexins	Via adapters to other junctions	ZO-1,2,3	Communication, small-molecule transport between cells
Plasmodesmata (plants only)	Cell-cell	Undefined	Actin filaments	NET1A	Communication, molecule transport between cells

Gap junctions permit the rapid diffusion of small, water-soluble molecules between the cytoplasms of adjacent cells. Along with anchoring and tight junctions, gap junctions help a cell communicate with its environment. However, they are structurally very different from anchoring junctions and tight junctions and do not play a key role in strengthening cell-cell and cell-matrix adhesions. Found in both epithelial and nonepithelial cells, gap junctions resemble the distinct cell junctions in plants called *plasmodesmata*, which we discuss in Section 20.6.

Four types of anchoring junctions are present in cells. Two participate in cell-cell adhesion and two participate in cell-matrix adhesion. **Adherens junctions** connect the lateral membranes of adjacent epithelial cells and are usually located near the apical surface, just below the tight junctions (see Figure 20-11). A circumferential

belt of actin and myosin filaments in a complex with the adherens junctions functions as a tension cable that can internally brace the cell and thereby control its shape. Epithelial and some other types of cells, such as smooth muscle and heart cells, are also bound tightly together by **desmosomes**, snap-like points of contact sometimes called *spot desmosomes*. *Hemidesmosomes*, found mainly on the basal surface of epithelial cells, and **focal contacts**, also called **focal adhesions**, anchor an epithelium to components of the underlying ECM, much like nails holding down a carpet. Adherens junctions, desmosomes, and focal adhesions are found in many different types of cells; hemidesmosomes appear to be restricted to epithelial cells.

Bundles of intermediate filaments running parallel to the cell surface or through the cell connect desmosomes and hemidesmosomes, imparting shape and rigidity to the cell, as do actin filaments that connect the cytoskeleton with focal contacts and adherens junctions. The close interaction between these junctions and the cytoskeleton helps transmit shear forces from one region of a cell layer to the epithelium as a whole, providing strength and rigidity to the entire epithelial cell layer. Desmosomes and hemidesmosomes are especially important in maintaining the integrity of skin epithelia. As a consequence, mutations that interfere with hemidesmosomal anchoring in the skin can lead to a condition in which the epithelium becomes detached from its underlying matrix and extracellular fluid accumulates at the basolateral surface, forcing the skin to balloon outward, forming a blister.

Cadherins Mediate Cell-Cell Adhesions in Adherens Junctions and Desmosomes

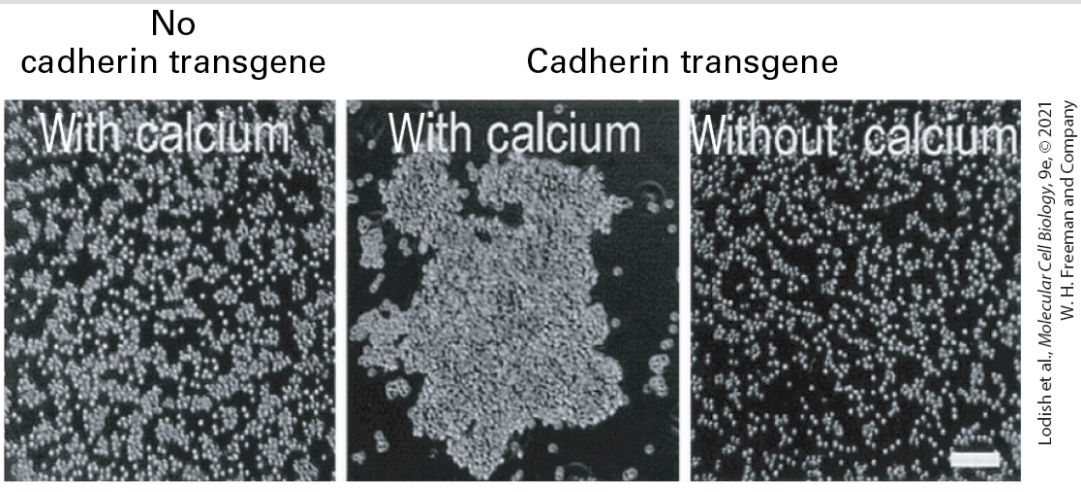
The primary CAMs in adherens junctions and desmosomes belong to the **cadherin** family of glycoproteins that can contain oligosaccharides that are either *N*-linked (see Chapters 13 and 14) or *O*-linked (described in Figure 20-31 in Section 20.4). In vertebrates, this protein superfamily of more than a hundred members can be grouped into at least six subfamilies, including **classical cadherins**, **clustered protocadherins**, and **desmosomal cadherins**, which we will describe below. The

diversity of cadherins arises from the presence of multiple cadherin genes and alternative RNA splicing. It is not surprising that there are many different types of cadherins in vertebrates; many different types of cells in the widely diverse tissues of these animals use cadherins to mediate adhesion and communication, the detailed requirements for which may differ for different types of cells and tissues. Members of the cadherin superfamily can also control cell morphology, such as the assembly and tight packing of microvilli on the apical surfaces of some epithelial cells (see Figures 20-10a and 20-11a). The brain expresses the largest number of different cadherins, including protocadherins, presumably owing to the necessity of forming many specific cell-cell contacts to help establish its complex wiring pattern. Invertebrates, however, are able to function with fewer than 20 cadherins.

Classical Cadherins

The classical cadherins include E-, N-, and P-cadherins, named for the type of tissue in which they were initially identified (epithelial, neural, and placental, respectively). E- and N-cadherins are the most widely expressed, particularly during early differentiation. Sheets of polarized epithelial cells, such as those that line the small intestine or kidney tubules, contain abundant E-cadherin along their lateral surfaces. Although E-cadherin is concentrated in adherens junctions, it is present throughout the lateral surfaces, where it is thought to link adjacent cell membranes.

The results of experiments with L cells, a line of cultured mouse fibroblasts, demonstrated that E-cadherins preferentially mediate homophilic interactions. L cells express no cadherins and adhere poorly to each other and to other types of cells. When the E-cadherin gene was introduced into L cells, the cells were found to adhere preferentially to other cells expressing E-cadherin (Figure 20-12). These engineered cadherin-expressing L cells formed epithelium-like aggregates with one another and with epithelial cells isolated from lungs. Although most E-cadherins exhibit primarily homophilic binding, some mediate heterophilic interactions.



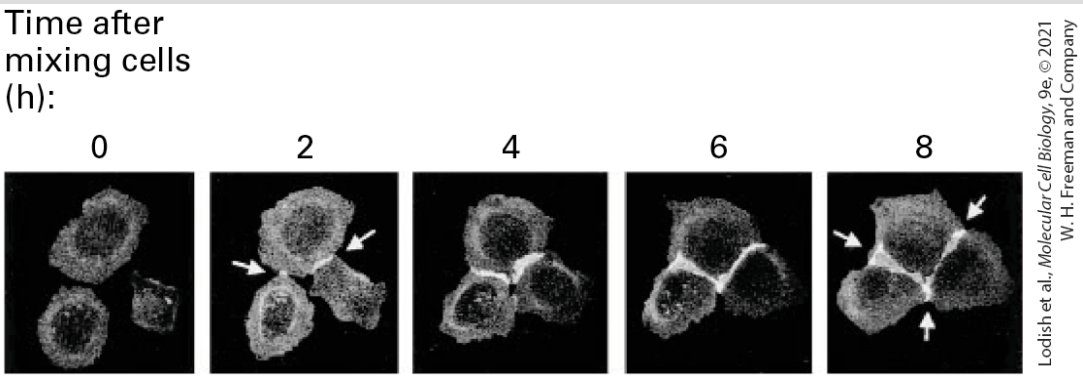
EXPERIMENTAL FIGURE 20-12 E-cadherin mediates Ca^{2+}-dependent adhesion of L cells. Under standard cell-culture conditions, in the presence of calcium in the extracellular fluid, L cells do not aggregate into sheets (*left*). Introduction of a gene that causes the expression of E-cadherin in these cells results in their aggregation into epithelium-like clumps in the presence of calcium (*center*), but not in its absence (*right*). Bar, 60 μm.

[© C. L. Adams et al., 1998, *J. Cell Biol.* **142**(4):1105–1119, Fig. 1E; https://doi.org/10.1083/jcb.142.4.1105.]

The adhesiveness of cadherins depends on the presence of extracellular Ca^{2+}; it is this property (*c*alcium *a*dhering) that gave rise to their name. For example, the adhesion of L cells expressing E-cadherin is prevented when the cells are bathed in a solution that is low in Ca^{2+} (see Figure 20-12). Some adhesion molecules require some minimal amount of Ca^{2+} in the extracellular fluid to function properly, whereas others, such as IgCAMs, are Ca^{2+} independent.

The role of E-cadherin in adhesion can also be demonstrated by experiments with cultured epithelial cells called *Madin-Darby canine kidney (MDCK)* cells (see Figure 4-3). A green fluorescent protein–labeled form of E-cadherin has been used in these cells to show that clusters of E-cadherin mediate the initial attachment of the cells and the subsequent zippering of the cells into sheets (Figure 20-13). In this experimental system, the addition of an antibody that binds to E-cadherin, preventing its homophilic interactions, blocks the Ca^{2+} - dependent attachment of

MDCK cells to one another and the subsequent formation of intercellular adherens junctions.



EXPERIMENTAL FIGURE 20-13 E-cadherin mediates adhesive connections in cultured MDCK epithelial cells. An E-cadherin gene fused to green fluorescent protein (GFP) was introduced into cultured MDCK cells. The cells were then mixed together in a calcium-containing medium, and the distribution of fluorescent E-cadherin was visualized over time (shown in hours). Clusters of E-cadherin mediate the initial attachment and subsequent zippering up of the epithelial cells and the formation of junctions (bicellular junctions are where two cells join and appear as lines; tricellular junctions are the sites of intersection of three cells).

[© C. L. Adams et al., 1998, *J. Cell Biol.* **142**(4):1105–1119, Fig. 2B; https://doi.org/10.1083/jcb.142.4.1105.]

Each classical cadherin molecule contains a single transmembrane domain, a relatively short C-terminal cytosolic domain, and five extracellular cadherin domains (called *EC1–EC5*) (see Figure 20-2). The extracellular domains are necessary for Ca^{2+} binding and cadherin-mediated cell-cell adhesion. Classical cadherin–mediated adhesion entails both cis lateral (on the same cell) clustering and trans adhesive (intercellular) molecular interactions (see Figures 20-3 and 20-14a–c). The binding of three Ca^{2+} ions at each of the sites located between the cadherin repeats (see Figures 20-2 and 20-14a) stabilizes the elongated and curved structure of the extracellular domain. As we shall see shortly, the curved structure of cadherin's extracellular domain is necessary for the proper molecular complementarity that stabilizes cis and trans binding between cadherin molecules. The cis and trans interactions of cadherins, together with their interactions with cytoplasmic adapter and cytoskeletal molecules, permit the zippering up of cadherins into adhesive arrays. Binding of the

EC1 domain of one cadherin molecule to the EC1 domain of another on the adjacent cell is responsible for trans binding (Figure 20-14a; see also Figure 20-3). Although the dissociation constant (K_d) for EC1–EC1 homophilic binding measured using isolated domains in solution is on the order of $10^{-5} - 10^{-4}$ mol/L (relatively weak, or low-affinity, binding), the multiple low-affinity interactions in arrays of intact cadherin molecules on adjacent cells sum to produce a very tight intercellular adhesion.

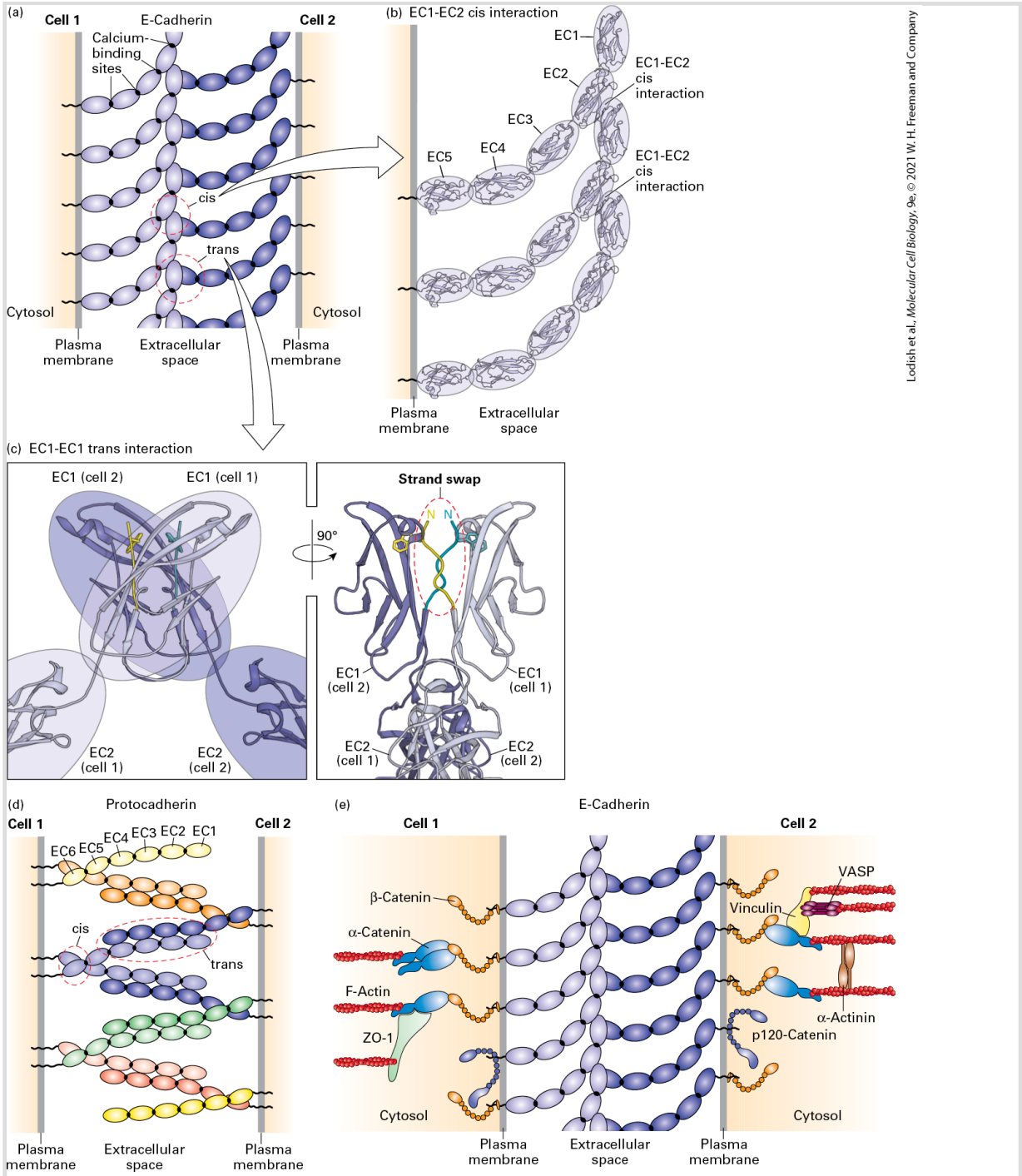


FIGURE 20-14 Trans (intercellular) and cis (on the same cell) interactions of classical cadherins in typical adherens junctions and of protocadherins. (a) The exoplasmic cadherin domains [EC1-EC5, see ovals in (b)] of E-cadherins at adherens junctions on adjacent cells are clustered by homophilic cis and trans interactions (see Figure 20-3). The Ca^{2+}-dependent elongated and curved structure of cadherin's extracellular domains is necessary for stable cis and trans interactions (examples highlighted by dashed circles). (b) EC1-EC2 cis interaction. Cis interactions are defined by the binding of an EC1 domain of one cadherin to an EC2 domain of an adjacent cadherin on the same cell. In (b) and (c) the structure of each extracellular cadherin domain is represented using a ribbon

diagram and highlighted by an oval. (c) EC1-EC1 trans interaction: Two views rotated by 90° of the trans binding of an EC1 domain of one cadherin to an EC1 domain of a cadherin on the adjacent cell. Only the EC1 and a portion of the EC2 domains are shown. Note that a small segment of polypeptide at the N-terminus of each of the two EC1 domains [highlighted in yellow (cell 1) and blue (cell 2)] swings out and replaces the equivalent segment from its binding partner (strand swap, dashed oval on the right). The strand swap places the side chain of a tryptophan residue on each of the segments into a binding pocket on the adjacent EC1 domain — an interaction that substantially stabilizes the trans binding. (d) The clustered protocadherins at neuron-to-neuron interfaces form a lattice of cis and trans interactions with their exoplasmic cadherin domains (ovals, EC1-EC6) that differs from that in classical cadherins (part a). Different protocadherin isoforms are represented by different colors (lighter colors expressed by Cell 1, darker colors expressed by Cell 2). The dimeric, homo- (same color) and heterophilic (different colors) cis interactions form between the EC6 domain of one molecule and EC5-EC6 domains of its partner (cis dashed circle). The dimeric, homophilic trans interactions form between the EC1-EC4 domains of one molecule with the EC1-EC4 domains of the same isoform (same color) on the adjacent cell. (e) The cytosolic domains of the classical E-cadherins bind directly or indirectly to multiple adapter proteins (e.g., β-catenin), which both connect the junctions to actin filaments (F-actin) of the cytoskeleton and participate in intracellular signaling pathways. Somewhat different sets of adapter proteins are illustrated in the two cells to emphasize the variety of adapters in adherens junctions. Some of these adapters, such as vinculin and ZO-1, can interact with several different CAMs. See V. Vasioukhin and E. Fuchs, 2001, *Curr. Opin. Cell Biol.* **13**:76–84; and J. Brasch, et al., 2012, *Trends Cell Biol.* **22**:299–310.

[Parts (a–c) Data from O. J. Harrison et al., 2011, *Structure* **19**: 244–256, PDB ID 3q2w. Part (d) Data from J. Brasch et al., 2019, *Nature* **569**: 280–283, PDB ID 6e6b.]

Determination of the structures of the extracellular domains of cadherins, together with analyses of the structures and binding properties of many mutants of the key binding domains, have provided a clear picture of the cis and trans interactions that underlie classical cadherin–mediated cell adhesion. The key features of cadherin cis and trans binding interactions are (1) the calcium-dependent curvature of the five extracellular cadherin domains that permits proper relative orientations of the EC1 and EC2 domains (see Figures 20-2 and 20-14a); (2) for cis interactions, the binding of one side of an EC1 domain to a complementary surface on the EC2 domain of an adjacent molecule on the same membrane (see Figures 20-2 and 20-14a and b); and (3) for trans interactions, the binding of a different surface of the EC1 domain to an EC1 domain from a cadherin molecule on the adjacent cell (Figure 20-14c). The trans EC1–EC1 binding is stabilized when a small segment of the protein at the N-terminus of each of the two EC1 domains swings out and replaces the equivalent segment from its binding partner (strand swap; see Figure 20-14c).

Another group of CAMs, called *clustered protocadherins*, provides an excellent example of how evolution can employ common molecular building blocks — in this case, cadherin domains — for distinct structural and functional purposes. Clustered protocadherins play a role in establishing the complex wiring diagram of neurons in mammalian brains. Multiple protocadherin genes encode 52 different clustered protocadherin proteins (isoforms) in humans (58 in mice). The random combinations of 10–15 different isoforms that are expressed in any individual neuron provide a bar code that helps a neuron distinguish itself from other neurons and thus helps generate neuronal wiring diagrams. The domain architecture of these protocadherins with six extracellular cadherin domains (EC1–EC6) differs from that of classical cadherins with five domains, and the organization of the cis and trans binding (see Figure 20-14d) differs markedly from that of classical cadherins (see Figure 20-14a–c). In clustered protocadherins, cis interactions are mediated by the EC6 domain of one molecule binding to the EC5 and EC6 domains of an adjacent molecule (Figure 20-14d). These cis interactions can be both homophilic and heterophilic, allowing the same or different isoforms to associate with each other into cis dimers [different colors represent different isoforms in part (d)]. In contrast, trans interactions are mediated by strictly homophilic, antiparallel binding of the EC1-EC4-domains (see Figure 20-14d). Trans binding only occurs when the sequences of all four of these EC domains perfectly match those of a corresponding protocadherin isoform on the adjacent cell. The lattices of intermixed protocadherin isoforms that form an interface between the membranes of neurons are responsible for their influence on neuronal wiring.

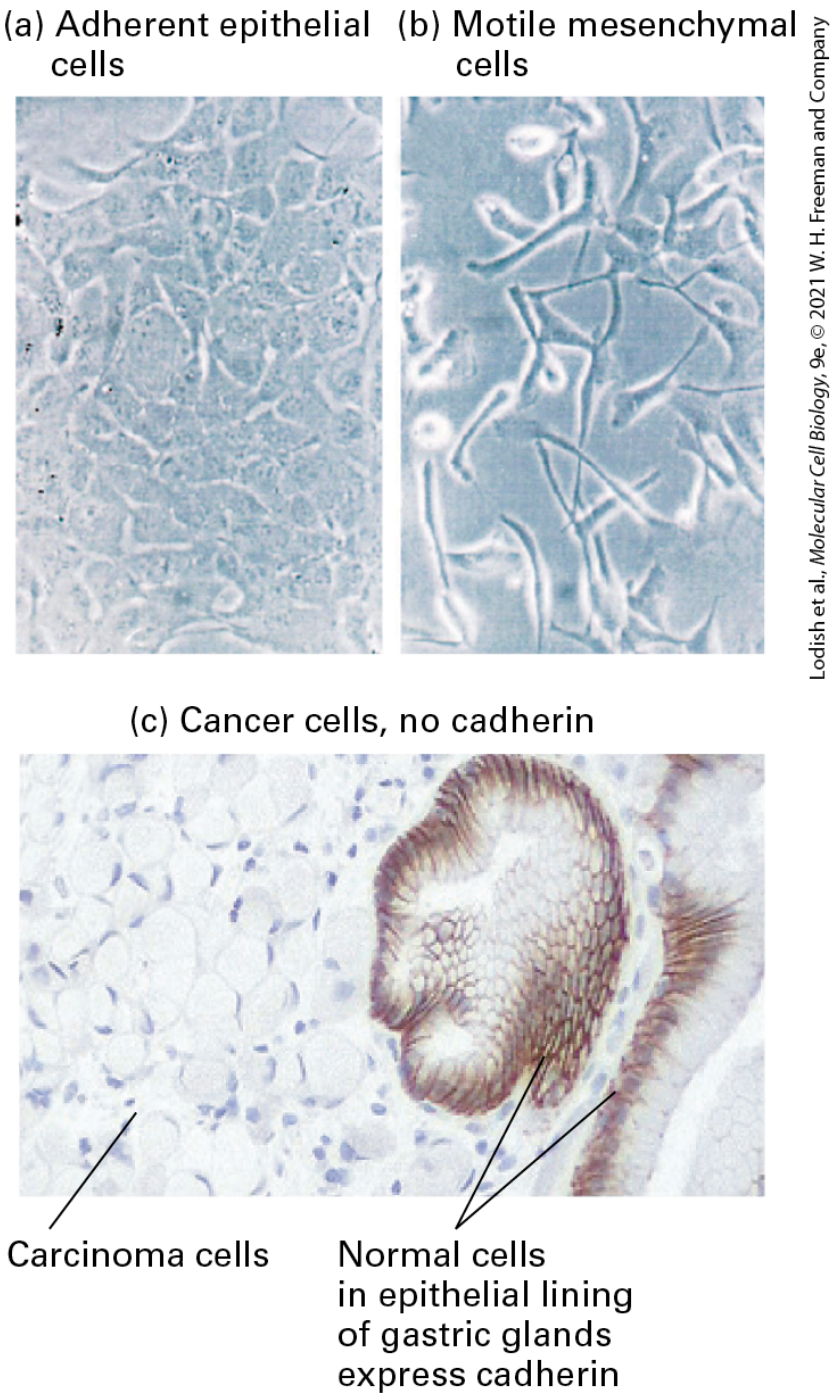
The cytosolic domains of CAMs have profound effects on the binding properties and functions of these adhesive molecules. For example, the C-terminal cytosolic domain of classical cadherins is linked to the actin cytoskeleton by adapter proteins (Figure 20-14e). These linkages are essential for strong adhesion. Indeed, cadherins together with their adaptor proteins in adherens junctions participate in mechanotransduction; a moderate increase in the mechanical tension at the junctions generated either intracellularly by actin and myosin, or by extracellular forces such as those transferred from adjacent cells or shear force due to blood flow, induces the formation of larger clusters of cadherins and stronger intercellular

adhesion. An important mechanosensor is the adapter protein α-catenin, which is attached to the cytoplasmic domain of cadherin via the protein β-catenin and also binds directly to F-actin (see Figure 20-14e). When the cadherin/β-catenin/α-catenin/F-actin complex is subjected to mechanical stress, the C-terminal actin-binding domain of α-catenin undergoes a conformational change to bind more tightly to the actin. In addition, a tension-induced conformational change in α-catenin's adhesion modulation domain uncovers otherwise cryptic binding sites for additional proteins, including the actin-binding adaptor vinculin. As a consequence, additional F-actin filaments can be recruited (see Figure 20-14e) and these further enlarge and strengthen the adherens junction. We will see vinculin playing a similar role in integrin-mediated adhesion later in this chapter.

Disruption of the interactions between classical cadherins and α-catenin or β-catenin dramatically reduces cadherin-mediated cell-cell adhesion. This disruption occurs spontaneously in tumor cells, which sometimes fail to express α-catenin, and can be induced experimentally by depleting the cytosolic pool of accessible β-catenin. The cytosolic domains of cadherins also interact with other adaptor molecules including p120-catenin, which can influence the stability of the cadherin. Interestingly, β-catenin plays a dual role: it not only mediates cytoskeletal attachment, but also serves as a signaling molecule, translocating to the nucleus and altering gene transcription in the Wnt signaling pathway (see Figure 16-26).

Classical cadherins play a critical role during tissue differentiation. Each classical cadherin has a characteristic tissue distribution. In the course of differentiation, the amounts or types of cell-surface cadherins change, affecting many aspects of cell-cell adhesion, cell migration, and cell division. For instance, the normal reorganization of tissues during morphogenesis is often accompanied by the conversion of nonmotile epithelial cells into motile cells, called *mesenchymal* cells, that are precursors for other tissues. This **epithelial-to-mesenchymal transition (EMT)** is associated with a reduction in the expression of E-cadherin (Figure 20-15a, b). The EMT is also associated with pathology, as in the conversion of epithelial cells into malignant carcinoma cells. For example, certain ductal breast tumors and hereditary diffuse gastric cancer (Figure 20-15c) characteristically involve a loss of E-

cadherin activity. It is well known that animal cell-cell contact can inhibit cell proliferation. During tissue development, once dividing epithelial cells have formed a well-defined, tightly bound epithelium, they have no need for further cell division unless they are damaged or receive a signal to undergo the EMT. It is now clear that one mechanism used to inhibit proliferation of epithelial cells in epithelia is E-cadherin- and catenin-mediated regulation of the Hippo pathway that controls cell proliferation (see Chapter 21).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-15 E-cadherin activity is lost during the epithelial-to-mesenchymal transition and during cancer progression. A protein called *Snail* that suppresses the expression of E-cadherin is associated with the epithelial-to-mesenchymal transition (EMT). (a) Normal epithelial MDCK cells grown in culture. (b) Expression of the *snail* gene in MDCK cells causes them to undergo an EMT. (c) Distribution of E-cadherin detected by immunohistochemical staining (dark brown) in a thin section of tissue from a patient with hereditary diffuse gastric cancer. E-

Cadherin is seen at the intercellular borders of normal stomach gastric gland epithelial cells; no E-cadherin is seen at the borders of underlying invasive carcinoma cells.

[Parts (a) and (b) republished with permission from Elsevier, from A. Martinez Arias, 2001, "Epithelial Mesenchymal Interactions in Cancer and Development," *Cell* **105**(4):425–431; permission conveyed through Copyright Clearance Center, Inc. Part (c) republished with permission from John Wiley & Sons, Inc., from F. Carneiro et al., 2004, "Model of the Early Development of Diffuse Gastric Cancer in E-Cadherin Mutation Carriers and Its Implications for Patient Screening," *J. Pathol*. **203**(2):681–687; permission conveyed through Copyright Clearance Center, Inc.]

The firm epithelial cell-cell adhesions mediated by cadherins in adherens junctions permit the formation of a second class of intercellular junctions in epithelia — tight junctions, to which we will turn shortly.

Infection with rhinoviruses (RV) is the most frequent cause of the common cold, and infection with virulent class C rhinoviruses (RV-C) can cause more severe illnesses, including exacerbation of asthma. To enter cells and replicate, RV-C must bind to cell-surface receptors. A cadherin family member called *CDHR3*, which is highly expressed in epithelial cells in the human airway, is a receptor for RV-C. Pathogens such as RV-C often evolve to co-opt proteins that have normal functions in their target (host) tissues. Genetic studies have shown that a naturally occurring mutation in humans that changes a cysteine to tyrosine (C → Y) in the EC5 domain of CDHR3 is associated with increased wheezing illnesses and hospitalizations for childhood asthma. In cultured cells, this C → Y mutation increases the cell-surface expression of CDHR3 and the binding and replication of RV-C. Treatments that disrupt the RV-C/cadherin (CDHR3) interaction have the potential to prevent or treat respiratory diseases caused by RV-C. ■

Desmosomal Cadherins

Desmosomes (Figure 20-16) contain two specialized cadherins, *desmoglein* and *desmocollin*, whose cytosolic domains are distinct from those in the classical cadherins. The cytosolic domains of desmosomal cadherins bind to adapter proteins

such as plakoglobin (similar in structure to β-catenin) and plakophilins, and these bind to a member of the plakin family of adapters, called *desmoplakin*. These adapters form the thick cytoplasmic plaques that are characteristic of desmosomes. The desmoplakins directly mediate plaque binding to intermediate filaments.

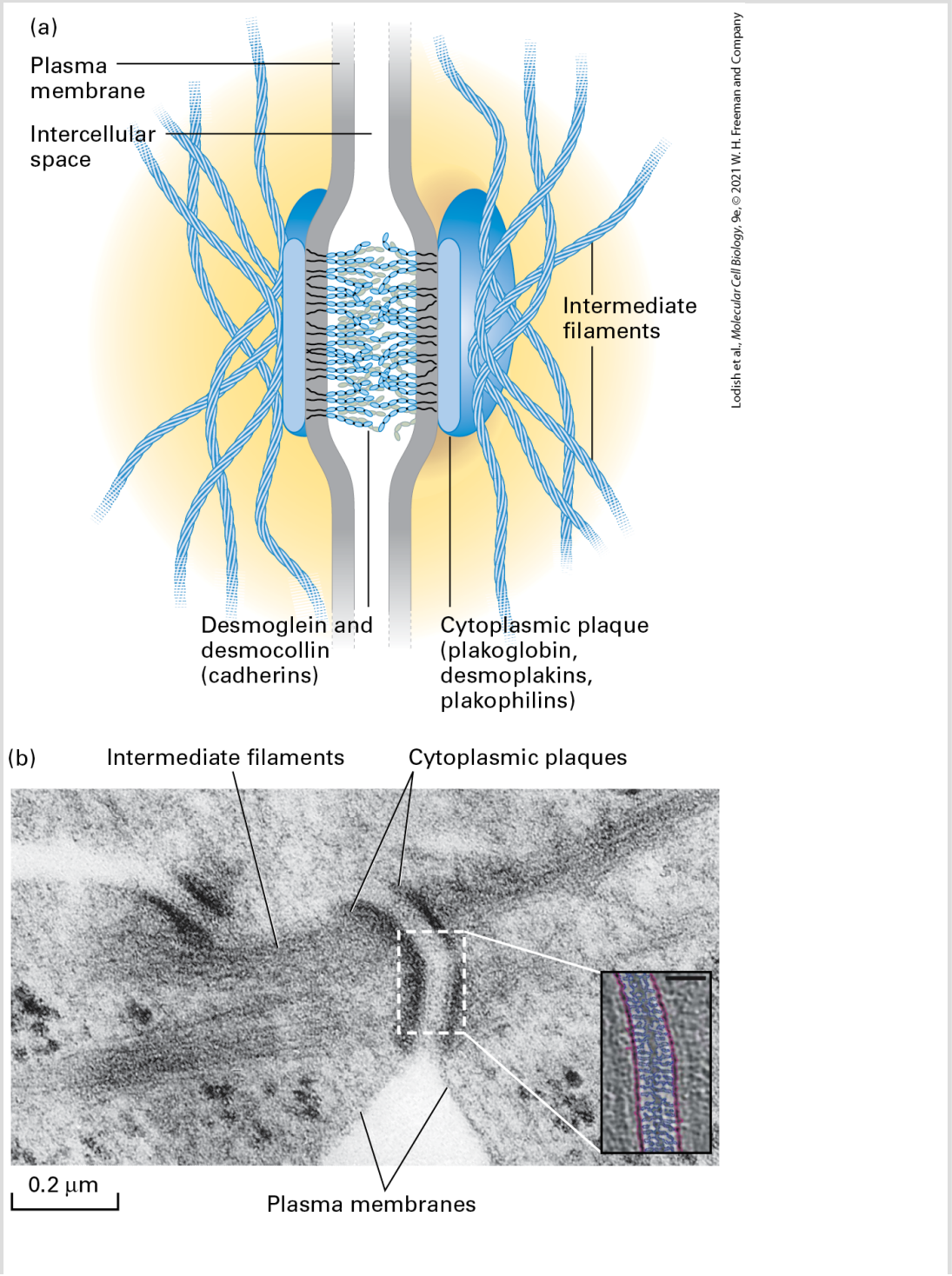


FIGURE 20-16 Desmosomes. (a) Model of a desmosome between epithelial cells with attachments to the sides of intermediate filaments. The key CAMs in desmosomes are the desmosomal cadherins desmoglein and desmocollin. Adapter proteins bound to the cytoplasmic domains of these cadherins include plakoglobin, desmoplakins, and plakophilins. See B. M. Gumbiner, 1993, *Neuron* **11**:551–564; L. A. Staehelin and B. E. Hull, 1978, *Sci. Am.* **238**:140; and D. R. Garrod, 1993, *Curr. Opin. Cell Biol.* **5**:30–40. (b) Electron micrograph of a thin section of a desmosome connecting two cultured differentiated human keratinocytes. Bundles of intermediate filaments radiate from the two darkly staining cytoplasmic plaques that line the inner surface of the adjacent plasma membranes. *Inset:* Electron microscopic tomograph of a desmosome linking two human epidermal cells (plasma membranes, pink; desmosomal cadherins, blue; bar, 35 nm).

[Part (b) republished with permission from Nature, from A. Al-Amoudi et al., 2007, "The Molecular Architecture of Cadherins in Native Epidermal Desmosomes," *Nature* **450**(7171): 832–837; permission conveyed through Copyright Clearance Center, Inc.]

The cadherin desmoglein was identified through studies of an unusual but revealing skin disease called *pemphigus vulgaris*, an autoimmune disease. Patients with autoimmune disorders synthesize self-attacking (or "auto") antibodies that bind to a normal body protein. In pemphigus vulgaris, the auto-antibodies disrupt adhesion between epithelial cells, causing blisters of the skin and mucous membranes. The predominant auto-antibodies in patients were shown to be specific for desmoglein; indeed, the addition of such antibodies to normal skin induces the formation of blisters and disruption of cell adhesion. ■

Integrins Mediate Cell-Matrix Adhesions, Including Those in Epithelial-Cell Hemidesmosomes

To be stably anchored to solid tissues and organs, simple columnar epithelial sheets must be firmly attached via their basal surfaces to the underlying ECM (basal lamina). This attachment occurs via adhesion receptors called **integrins** (see Figure 20-2), which are located both within and outside of anchoring junctions called *hemidesmosomes* (see Figure 20-11a). Hemidesmosomes comprise integral membrane proteins linked via cytoplasmic adapter proteins (e.g., plakins) to keratin-based

intermediate filaments. The principal ECM adhesion receptor in epithelial hemidesmosomes is integrin α6β4.

Integrins function as cell-cell and cell-ECM adhesion receptors in a wide variety of epithelial and nonepithelial cells, mediating many cell-matrix and cell-cell interactions (Table 20-4). In vertebrates, at least 24 integrin heterodimers, composed of 18 types of α subunits and 8 types of β subunits in various αβ heterodimeric combinations, are known. A single type of β chain can interact with any one of several different types of α chains, forming distinct integrins that bind different ligands. This phenomenon of *combinatorial diversity* allows a relatively small number of components to serve a large number of distinct functions. Although most cells express several distinct integrins that bind the same or different ligands, many integrins are expressed predominantly in certain types of cells. Not only do many integrins bind more than one ligand, but there are ligands that can bind to any one of several different integrins.

TABLE 20-4 • Selected Vertebrate Integrins

Subunit Composition	Primary Cellular Distribution	Ligands
α1β1	Many types	Mainly collagens
α2β1	Many types	Mainly collagens; also laminins
α3β1	Many types	Laminins
α4β1	Hematopoietic cells	Fibronectin; **VCAM-1**
α5β1	Fibroblasts	Fibronectin
α6β1	Many types	Laminins
αLβ2	T lymphocytes	**ICAM-1, ICAM-2**
αMβ2	Monocytes	Serum proteins (e.g., C3b, fibrinogen, factor X); **ICAM-1**
αIIbβ3	Platelets	Serum proteins (e.g., fibrinogen, von Willebrand factor, vitronectin); fibronectin

α6β4	Epithelial cells	Laminin

NOTE: The integrins are grouped into subfamilies having a common β subunit. Ligands shown in red are CAMs; all others are ECM or serum proteins. Some subunits can have multiple spliced isoforms with different cytosolic domains.

SOURCE: Data from R. O. Hynes, 1992, *Cell* **69**(1):11.

All integrins appear to have evolved from two ancient general subgroups: those that bind proteins containing the tripeptide sequence Arg-Gly-Asp, usually called the **RGD motif** (fibronectin is one such protein), and those that bind laminin. Several integrin α subunits contain a distinctive, inserted domain, the *I-domain*, which can mediate binding of certain integrins to various collagens in the ECM. Some integrins with I-domains are expressed exclusively on leukocytes (white blood cells) and red and white blood cell precursor (hematopoietic) cells. I-domains also recognize CAMs on other cells, including members of the Ig superfamily (e.g., ICAMs, VCAMs), and thus participate in cell-cell adhesion.

Integrins typically exhibit low affinities for their ligands, with dissociation constants (K_d) between 10^{-6} and 10^{-7} mol/L. However, the multiple weak interactions generated by the binding of hundreds or thousands of integrin molecules to their ligands on cells or in the ECM allow a cell to remain firmly anchored to its ligand-expressing target.

Parts of both the α subunit and the β subunit of an integrin molecule contribute to the primary extracellular ligand-binding site (see Figure 20-2). Ligand binding to integrins also requires the simultaneous binding of divalent cations. Like that of other adhesion molecules, the cytosolic region of integrins interacts with adapter proteins, which in turn bind to the cytoskeleton and to intracellular signaling molecules (see Figure 20-8). Most integrins are linked via adapters to the actin cytoskeleton, including two of the integrins that connect the basal surface of epithelial cells to the basal lamina via the ECM molecule laminin. Some integrins, however, interact with intermediate filaments. The cytosolic domain of the β4 chain in the α6β4 integrin in hemidesmosomes (see Figure 20-1), which is much longer

than the cytosolic domains of other integrin β chains, binds to specialized adapter proteins, which in turn interact with keratin-based intermediate filaments (see Table 20-4). Other integrins (e.g., α3β1) are the adhesion receptors in the focal contacts linking the epithelial basal lamina with the actin cytoskeleton (see Figure 20-1).

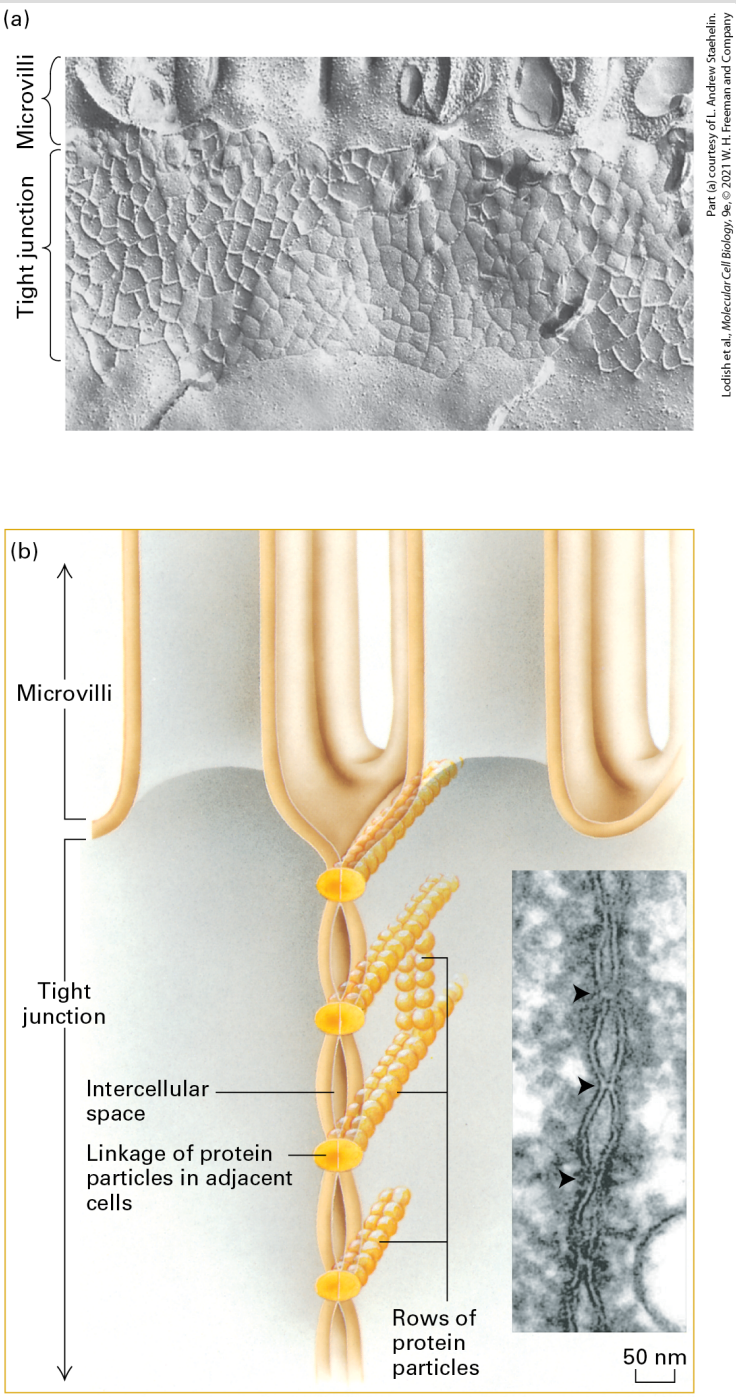
As we will see, the diversity of integrins and their ECM ligands allows integrins to participate in a wide array of key biological processes, including the inflammatory response and the migration of cells to their correct locations during morphogenesis. The importance of integrins in diverse processes is highlighted by the defects exhibited by knockout mice engineered to have mutations in various integrin subunit genes. These defects include major abnormalities in development, blood vessel formation, leukocyte function, inflammation, bone remodeling, and blood clotting. Despite their differences, all these processes depend on integrin-mediated interactions between the cytoskeleton and either the ECM or CAMs on other cells.

In addition to their adhesion function, integrins can mediate outside-in and inside-out signaling (see Figure 20-8). The engagement of integrins by their extracellular ligands can, through adapter proteins bound to the integrin's cytosolic region, influence the cytoskeleton and intracellular signaling pathways (outside-in signaling). Conversely, intracellular signaling pathways can alter the structure of integrins and consequently their abilities to adhere to their extracellular ligands and mediate cell-cell and cell-matrix interactions (inside-out signaling). We will see a detailed example of integrin-mediated signaling later in this chapter. Integrin-mediated signaling pathways influence processes as diverse as cell survival, cell proliferation, and programmed cell death (see Chapter 21).

Tight Junctions Seal Off Body Cavities and Restrict Diffusion of Membrane Components

For polarized epithelial cells to function as barriers and mediators of selective transport, extracellular fluids surrounding their apical and basolateral membranes must be kept separate. Tight junctions between adjacent epithelial cells are usually

located in a band surrounding the cell just below the apical surface (Figure 20-17; see also Figure 20-11). These specialized junctions form a barrier that seals off body cavities such as the intestinal lumen and separates the blood from the cerebral spinal fluid of the central nervous system (i.e., the blood-brain barrier).



Part (a) courtesy of L. Andrew Staehelin.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-17 Tight junctions. (a) Freeze-fracture preparation of tight-junction zone between two intestinal epithelial cells. The fracture plane passes through the plasma membrane of one of the two adjacent cells (see also Figure 20-11). A honeycomb-like network of ridges and grooves below the microvilli constitutes the tight-junction zone. (b) Schematic drawing shows how a tight junction might be formed by the linkage of rows of protein particles in adjacent cells. In the inset micrograph of an ultrathin sectional view of a tight junction, the adjacent cells can be seen in close contact where the rows of proteins interact. See L. A. Staehelin and B. E. Hull, 1978, *Sci. Am.* **238**:140.

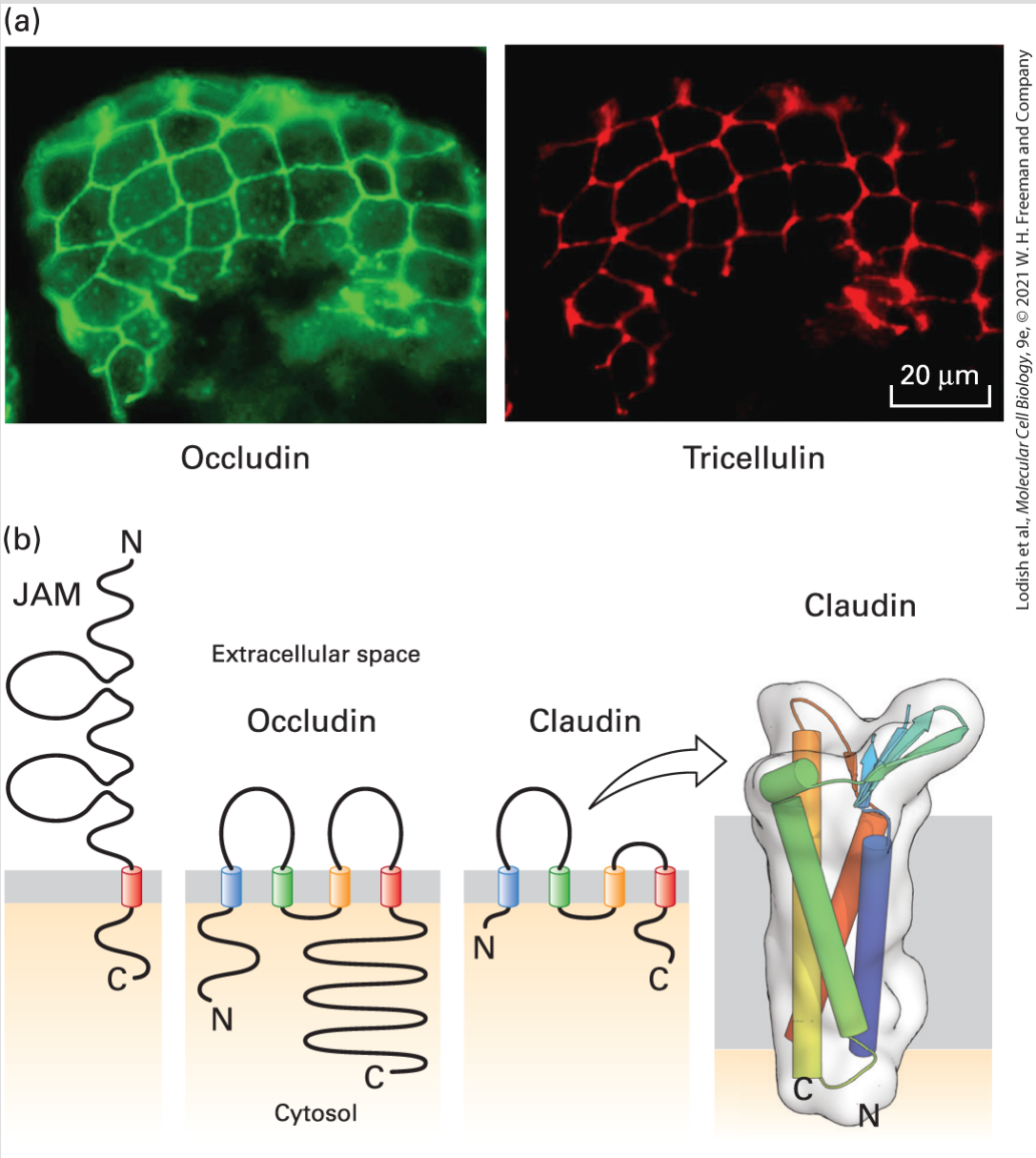
[Part (b) republished with permission from Nature, from S. Tsukita, M. Furuse, and M. Itoh, 2001, "Multifunctional Strands in Tight Junctions," *Nat. Rev. Mol. Cell Biol*. **2**(4):285–293; permission conveyed through Copyright Clearance Center, Inc.]

Tight junctions prevent the diffusion of macromolecules and, to varying degrees, small water-soluble molecules and ions across an epithelium via the spaces between cells. They also help establish and maintain the polarity of epithelial cells by preventing the diffusion of membrane proteins and glycolipids between the apical and the basolateral regions of the plasma membrane, ensuring that these regions contain different membrane components. Indeed, the lipid compositions of the apical and basolateral regions of the plasma membrane's exoplasmic leaflet (see Chapter 10) are distinct. Essentially all cell surface glycolipids are restricted to the exoplasmic face of the apical membrane, as are all proteins linked to the membrane by a glycosylphosphatidylinositol (GPI) anchor (see Figure 10-19). In contrast, the apical and basolateral regions of the plasma membrane's cytosolic leaflet have uniform membrane composition in epithelial cells; their lipids and proteins can apparently diffuse laterally from one region of the membrane to the other.

Tight junctions are composed of thin bands of plasma-membrane proteins that completely encircle the cell and are in contact with similar thin bands on adjacent cells. When thin sections of the tight junctions in cells are viewed in an electron microscope, the lateral surfaces of adjacent cells appear to touch each other at intervals and even to fuse in the zone just below the apical surface (see Figure 20-11b). In freeze-fracture preparations, tight junctions appear as an interlocking network of ridges and grooves in the plasma membrane (Figure 20-17a). Very high magnification reveals that rows of protein particles 3–4 nm in diameter form the ridges seen in freeze-fracture micrographs of tight junctions. In the model shown in

Figure 20-17b, the tight junction is formed by a double row of these particles, one row donated by each cell. Treatment of an epithelium with the protease trypsin destroys the tight junctions, supporting the proposal that proteins are essential structural components of these junctions.

There are a number of integral membrane proteins found in tight junctions; two of the best studied are *occludin* and *claudin* (from the Latin *claudere*, "to close"). Each of these proteins has four membrane-spanning α helices (Figure 20-18) and are representative of a group of such cell-surface proteins called *tetraspanins*. The mammalian claudin gene family encodes as many as 27 homologous proteins that exhibit distinct tissue-specific patterns of expression and, as described below, exhibit distinct properties. A group of *junction adhesion molecules (JAMs)* have also been found to contribute to homophilic adhesion and other functions of tight junctions. JAMs and another junctional protein, the *coxsackievirus and adenovirus receptor (CAR)*, contain a single transmembrane α helix and belong to the Ig superfamily of CAMs. The extracellular domains of rows of occludin, claudin, and JAMs in the plasma membrane of one cell apparently form extremely tight links with similar rows of the same proteins in an adjacent cell, creating a tight seal. Ca^{2+}-dependent cadherin-mediated adhesion also plays an important role in tight-junction formation, stability, and function.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-18 Proteins that compose tight junctions. (a) Immunofluorescence localization of occludin (green) and tricellulin (red) in mouse intestinal epithelium. Note that tricellulin is predominantly concentrated in tricellular junctions. (b) The junction adhesion molecule (JAM) has a single transmembrane domain and an extracellular region with two immunoglobulin domains, whereas occludin and claudins contain four transmembrane helices. The transmembrane helices of claudin-15 form a four-helix bundle, and the extracellular loops contain a five-stranded β sheet (seen edgewise in this view) and α helices. These extracellular secondary structures participate in cis interactions that form a single-file row of claudins, which in turn interact via other cis contacts with an adjacent row of claudins in the same membrane of each cell. There are also trans interactions between the rows of claudins in adjacent cells (see also Figure 20-20).

[Part (a) J. Ikenouchi et al., 2005, *J. Cell Biol.* **171**(6):939–945, Fig. 3A; https://doi.org/10.1083/jcb.200510043. Part (b) structure of claudin-15 data from H. Suzuki et al., 2014,

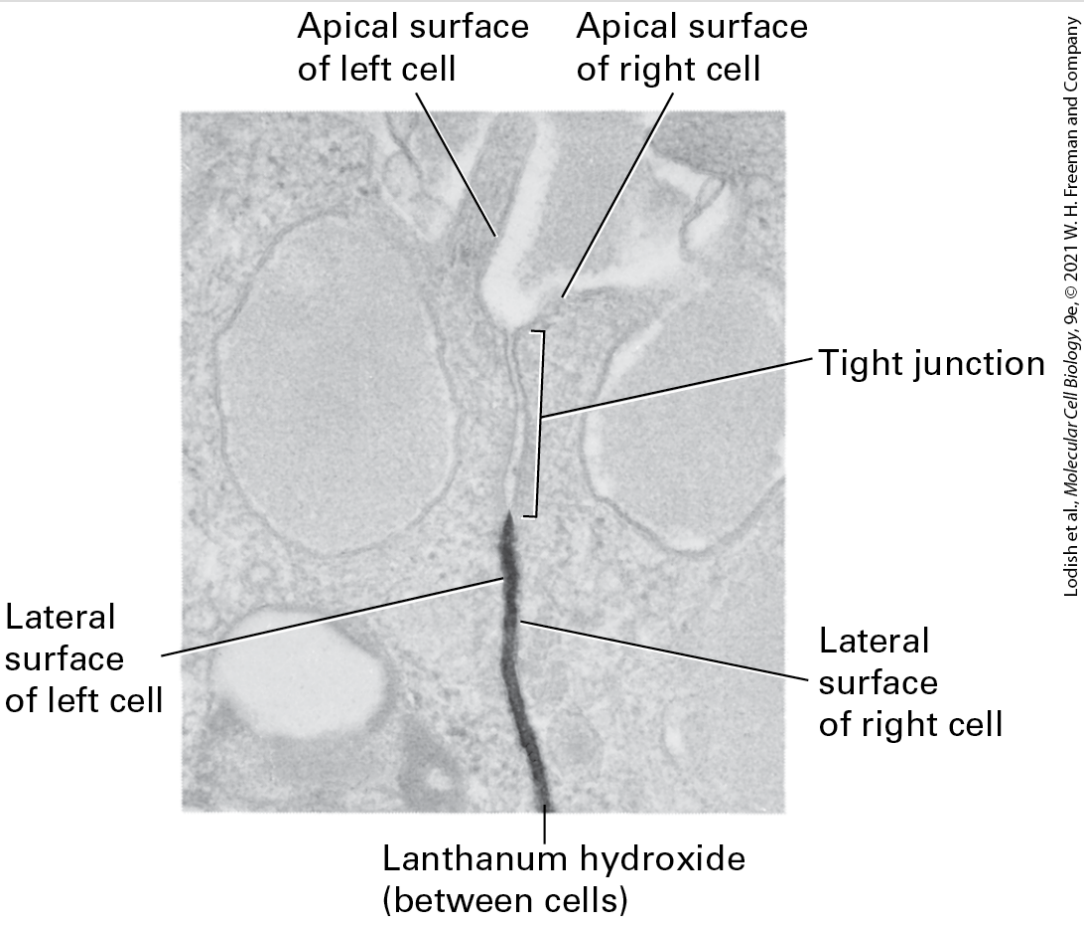
Science **344**:304–307, PDB ID 4p79.]

At the intersection of three cells connected to one another by tight junctions (see Figure 20-13 and Figure 20-18a), two additional transmembrane proteins are incorporated into the tight junctions: *tricellulin*, which has four membrane-spanning helices; and *angulins*, which have a single transmembrane helix and one extracellular immunoglobulin domain and appear to be required for the assembly of tricellulin where the cells intersect. As with tight junctions, there appear to be tension-sensitive proteins that localize at tricellular vertices in at least some adherens junctions.

As is the case for adherens junctions and desmosomes, cytosolic adapter proteins and their connections to the cytoskeleton are critical components of tight junctions and can mediate signaling. For example, when one specific claudin, claudin-6, is engaged in intercellular tight junctions, its cytoplasmic domain can initiate a kinase signaling cascade that results in the activation of nuclear receptors (e.g., the retinoic acid receptor γ and the estrogen receptor α) to control cell behavior (see Chapter 8). Also, the long C-terminal cytosolic segment of occludin binds to PDZ domains in some large, multidomain adapter proteins. **PDZ domains** are about 80 to 90 amino acids long and are found in various cytosolic proteins; they mediate binding to other cytosolic proteins or to the C-termini of particular plasma-membrane proteins. Cytosolic proteins containing a PDZ domain often have more than one of them. In the human genome, there are more than 250 PDZ domains in hundreds of proteins. Proteins with multiple PDZ domains can serve as scaffolds on which to assemble proteins into larger functional complexes. Several multiple-PDZ-domain–containing adapter proteins are associated with tight junctions, including the *zonula occludens (ZO)* proteins ZO-1, ZO-2, and ZO-3, which not only interact with occludin, claudin, and other adapter and signaling proteins but also mediate association with actin fibers. These interactions appear to stabilize the linkage between occludin and claudin molecules that is essential for maintaining the integrity of tight junctions. ZO proteins, which may form phase-separated biomolecular condensates at tight

junctions, can also function as adapters in adherens junctions (see Figure 20-14) and gap junctions.

The role of tight junctions as permeability barriers has been demonstrated in many experiments. For example, in one simple experiment, lanthanum hydroxide (an electron-dense colloid of high molecular weight) is injected into the pancreatic blood vessel of an experimental animal; a few minutes later, the pancreatic epithelial acinar cells are fixed and prepared for microscopy. As shown in Figure 20-19, the lanthanum hydroxide diffuses from the blood into the space that separates the lateral surfaces of adjacent acinar cells, but it cannot penetrate past the tight junction.



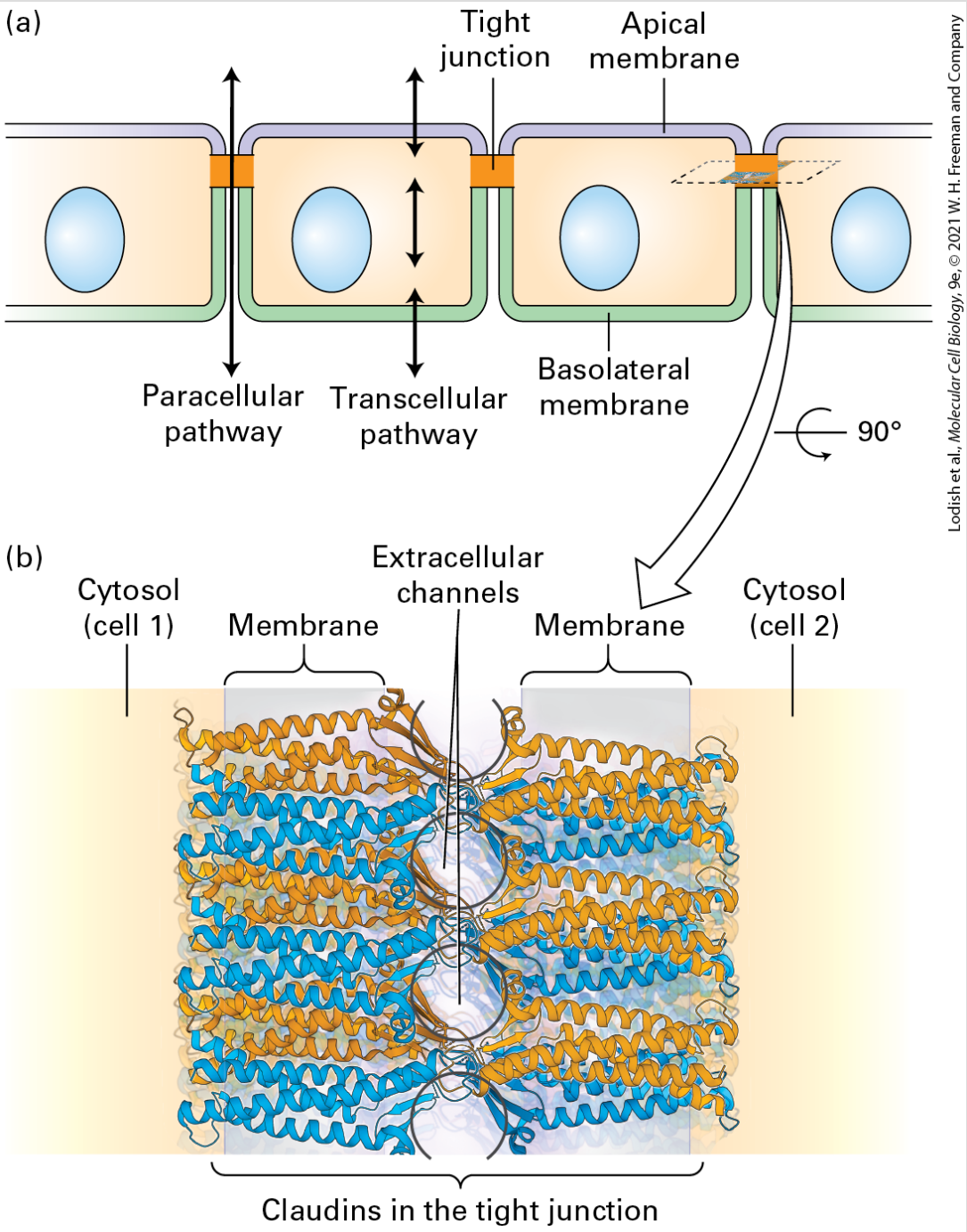
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-19 Tight junctions prevent passage of large molecules through extracellular spaces between epithelial cells. Tight junctions in the pancreas are impermeable to the

large water-soluble colloid lanthanum hydroxide (dark stain) administered from the basolateral side of the epithelium.

[From D. S. Friend and N. B. Gilula, 1972, *J. Cell Biol.* **53**(3):758–776.]

As a consequence of tight junctions, many nutrients cannot move across an epithelium between cells; instead, their transport is achieved in large part through the **transcellular pathway** via specific membrane-bound transport proteins (Figure 20-20a; see also Figure 11-30). The barrier to diffusion provided by tight junctions, however, is not absolute; some can exhibit size- and ion-selective permeability, depending on the type of claudins incorporated into the tight junction. For some, but not all, claudins the intercellular trans association of their closely packed extracellular domains results in the formation of paracellular channels (pores) (Figure 20-20b). Thus certain small molecules and ions can move from one side of the epithelium to the other through the **paracellular pathway** (see Figure 20-20). The importance of selective permeability is highlighted by the evolutionary conservation of the molecules that establish it and the diseases that arise when it is disrupted. For example, murine embryos cannot develop properly if selective permeability is disrupted because proper fluid balance on the two sides of epithelia cannot be maintained. Similarly, the kidneys depend on proper tight-junction permeability to establish the ion gradients necessary for normal regulation of body fluids and waste removal. Owing at least in part to the varying properties of the different types of claudin molecules located in different tight junctions, the permeability of the tight junctions to ions, small molecules, and water varies enormously among different epithelial tissues.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-20 Transcellular and paracellular pathways of transepithelial transport. (a) Transcellular transport requires the cellular uptake of molecules on one side of the cell and subsequent release on the opposite side by mechanisms for crossing membranes (see Chapter 11). In paracellular transport molecules move extracellularly through channels in the tight junctions, whose permeability to small molecules and ions depends on the composition of the junctional components (mainly claudins) and the physiological state of the epithelial cells. A cross-sectional plane through the tight junction linking the right-most two cells is indicated by dashed lines. Part (b) shows a perpendicular view of this plane

(90 ° rotation). (b) A computational model of claudin-15 in a tight junction. No other tight-junctional proteins are included in this model. This perpendicular view of the cross-sectional plane illustrated in part (a) shows the transmembrane helices of claudin-15 molecules spanning the tight junctional membranes. In each of the two cells, there are adjacent, single-file rows of claudins (orange and blue) that form trans intercellular contacts with the rows of claudins in the adjacent cell. Extracellular channels that are oriented perpendicular to this cross-sectional plane (outlined by gray circles) are formed by the extracellular domains of claudin-15. These channels permit the paracellular movement of small molecules and ions across the epithelium. There are negatively charged side chains of aspartic acid residues that line the channels and appear to contribute to the cation specificity of claudin-15-dependent channels.

[Part (a) Data from S. Tsukita et al., 2001, *Nat. Rev. Mol. Cell Biol.* **2**:285. Part (b) Information from P. Samanta et al., 2018, *J. Gen. Physiol.* **150**:949–968.]

The permeability of tight junctions can be altered by intracellular signaling pathways, especially G protein and cyclic AMP–coupled pathways (see Chapter 15). The regulation of tight-junction permeability is often studied by measuring ion flux (electrical resistance, called *transepithelial resistance*) or the movement of radioactive or fluorescent molecules across monolayers of MDCK or other epithelial cells.

The importance of paracellular transport is apparent in several human diseases. In hereditary hypomagnesemia, a defect in the *claudin16* gene prevents the normal paracellular flow of magnesium in the kidney. This defect results in an abnormally low blood level of magnesium, which can lead to convulsions. Furthermore, a mutation in the *claudin14* gene causes hereditary deafness, apparently by altering transport around hair-cell epithelia in the cochlea of the inner ear.

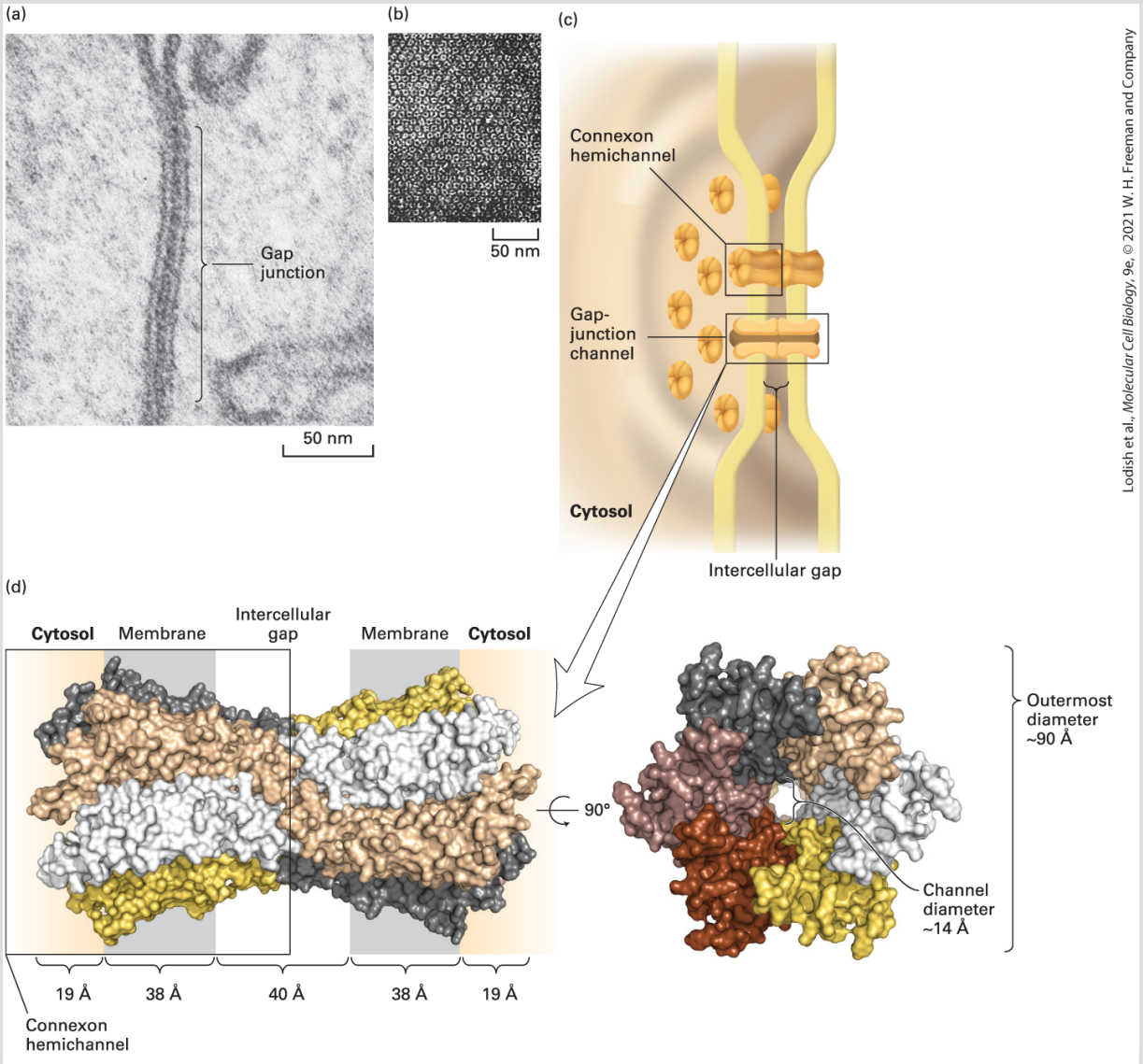
Some pathogens have evolved to exploit the molecules in tight junctions. Some use junctional proteins as co-receptors to attach to cells prior to infecting them (e.g., hepatitis C virus uses claudin-1 and occludin, together with two other co-receptors, to enter liver cells). Others break down the tight-junction barrier and cross epithelia via paracellular movement, and still others produce toxins that alter barrier function. For example, toxins produced by *Vibrio cholerae*, the enteric bacterium that causes cholera, alter the permeability barrier of the intestinal epithelium by altering the composition or activity of tight junctions. *Vibrio cholerae* also releases a protease that disrupts tight junctions by degrading the extracellular domain of occludin.

Other bacterial toxins can affect the ion-pumping activity of membrane transport proteins in intestinal epithelial cells. Toxin-induced changes in tight-junction permeability (increased paracellular transport) and in protein-mediated ion pumping (increased transcellular transport) can result in massive losses of internal body ions and water into the gastrointestinal tract, which in turn leads to diarrhea and potentially lethal dehydration (see Chapter 11). ■

It is clear that tight junctions are important for maintaining tissue integrity and function. Thus these junctions should and can respond to dynamic changes in epithelia, for example, when they are subjected to either external forces (such as stretching of the bladder by filling with urine) or internal forces (actin-myosin-induced changes in tension). It is clear that when there are changes in the mechanical forces at the tight junction there are coordinated changes in the composition of junctional adaptor proteins and paracellular permeability. The molecular mechanisms underlying mechanotransduction at tight junctions are not as well understood as those at adherens junctions, with some studies suggesting that ZO-1 may serve as a mechanosensor.

Gap Junctions Composed of Connexins Allow Small Molecules to Pass Directly Between the Cytosols of Adjacent Cells

Early electron micrographs of tissues revealed sites of cell-cell contact with a characteristic intercellular gap (Figure 20-21a). This feature, which was found in virtually all animal cells that contact other cells, prompted early morphologists to call these regions gap junctions. In retrospect, the most important feature of these junctions is not the 2–4 nm gap itself, but a well-defined set of cylindrical particles that cross the gap and compose pores connecting the cytosols of adjacent cells (Figure 20-21b, c). As we will see in Section 20.6, plant cells also assemble pores that connect the cytosols of adjacent cells, but those channels, called *plasmodesmata*, differ considerably in structure from gap junctions.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-21 Gap junctions. (a) In this thin section through a gap junction connecting two mouse liver cells, the two plasma membranes are closely associated for a distance of several hundred nanometers, separated by a gap of 2–3 nm. (b) Numerous roughly hexagonal particles are visible in this perpendicular view of the cytosolic face of a region of plasma membrane enriched in gap junctions. Each hexagonal particle aligns with a similar particle on an adjacent cell, forming a channel connecting the two cells. (c) Schematic model of a gap junction connecting two plasma membranes. Both membranes contain connexon hemichannels, cylinders of six dumbbell-shaped connexin molecules. Two connexons join in the gap between the cells to form a gap-junction channel, 1.4–2.0 nm in diameter, that connects the cytosols of the two cells. (d) Structure of recombinant human Cx26 gap junction as determined by x-ray crystallography (3.5-Å resolution). *Left:* Space-filling model of a side view of the complete structure of two attached connexons oriented as in part (c). Each of the six connexins that comprise a connexon has four transmembrane helices and is shown in a distinct color. The structures of the loops connecting the transmembrane helices are not well defined and not shown. *Right:* View from the cytosol perpendicular to the membrane bilayers, looking down on the connexon with its central pore. The diameter of the pore's channel is ~14 Å, and it is lined by

many polar/charged amino acids. See S. Nakagawa, S. Maeda, and T. Tsukihara, 2010, *Curr. Opin. Struct. Biol.* **20**(4):423–430.

[Part (a) courtesy of Daniel Goodenough. Part (b) D. L. Caspar et al., 1977, **74**:605–628, Fig. 2b; https://doi.org/10:1083/jcb.74.2.605. Part (c) Data from L. A. Staehelin and B. E. Hull, 1978, *Sci. Am.* **238**:140. Part (d) Data from S. Maeda et al., 2009, *Nature* **458**:597–602, PDB ID 2zw3.]

In many animal tissues, anywhere from a few to thousands of gap-junction particles cluster together in patches (e.g., along the lateral surfaces of epithelial cells; see Figure 20-11). When the plasma membrane is purified and then sheared into small fragments, some pieces mainly containing patches of gap junctions are generated. Owing to their relatively high protein content, these fragments have a higher density than the bulk of the plasma membrane and can be purified by equilibrium density-gradient centrifugation (see Figure 4-37). When these preparations are viewed perpendicular to the membrane, the gap junctions appear as arrays of hexagonal particles that enclose water-filled channels (see Figure 20-21b).

The effective pore size of gap junctions can be measured by injecting a cell with a fluorescent dye covalently linked to membrane bilayer–impermeable molecules of various sizes and observing with a fluorescence microscope whether the dye passes into neighboring cells. Gap junctions between mammalian cells permit the passage of molecules as large as 1.2 nm in diameter. In insects, these junctions are permeable to molecules as large as 2 nm in diameter. Generally speaking, molecules smaller than 1200 Da pass freely and those larger than 2000 Da do not pass; the passage of intermediate-sized molecules is variable and limited. Thus ions, many low-molecular-weight precursors of cellular macromolecules, products of intermediary metabolism, and small intracellular signaling molecules can pass from cell to cell through gap junctions.

In neural tissue, some neurons are connected by gap junctions through which ions pass rapidly, thereby allowing very rapid transmission of electrical signals. Impulse transmission through these connections, called *electrical synapses*, is almost a thousand times as rapid as at chemical synapses (see Chapter 23). Gap junctions are also present in many non-neural tissues, where they help to integrate the electrical

and metabolic activities of many cells. In the heart, for instance, gap junctions rapidly pass ionic signals among cardiac muscle cells, which are tightly bound together via desmosomes. Thus gap junctions contribute to the electrically stimulated coordinate contraction of cardiac muscle cells during a heartbeat. As discussed in Chapter 15, some extracellular hormonal signals induce the production or release of small intracellular signaling molecules called *second messengers* (e.g., cyclic AMP, IP_3, and Ca^{2+}) that regulate cellular metabolism. Because many second messengers can be transferred between cells through gap junctions, hormonal stimulation of one cell has the potential of triggering a coordinated response by that cell as well as many of its neighbors. Such gap-junction-mediated signaling plays an important role, for example, in the secretion of digestive enzymes by the pancreas and in the coordinated muscular contractile waves (peristalsis) in the intestine. Another vivid example of gap-junction-mediated transport is the phenomenon of *metabolic coupling*, or *metabolic cooperation*, in which a cell transfers nutrients or intermediary metabolites to a neighboring cell that is itself unable to synthesize them. Gap junctions play critical roles in the development of egg precursors (oocytes) in the ovaries by mediating the movement of both metabolites and signaling molecules, such as cyclic GMP, between an oocyte and its surrounding granulosa cells, as well as between neighboring granulosa cells.

A current model of the structure of the gap junction is shown in Figure 20-21c–d. Vertebrate gap junctions are composed of **connexins**, a family of structurally related transmembrane proteins with molecular weights between 26,000 and 60,000. Each vertebrate hexagonal particle consists of 12 noncovalently associated connexin molecules: six form a cylindrical hemichannel, called a *connexon*, in one plasma membrane that is joined to a connexon in the adjacent cell membrane, forming a continuous aqueous channel (diameter ~14 Å) between the cells. Each individual connexin molecule has four membrane-spanning α helices with a topology similar to that of claudin (see Figure 20-18), resulting in 24 transmembrane α helices in each connexon hemichannel.

A completely different family of proteins, the *innexins*, forms the gap junctions in invertebrates. A third family of innexin-like proteins, called *pannexins*, has been found in both vertebrates and invertebrates. Pannexins form hexamer hemichannels (pannexons) whose opening can be regulated by changes in membrane potential or mechanical stress. When open, pannexons permit direct exchange of small molecules (such as ATP) and ions between the intracellular and extracellular spaces. Pannexons are thought to play key roles in release of ATP from cells into the extracellular space. Extracellular ATP (as well as ADP and AMP) can function as an intercellular messenger or transmitter by binding to and activating the cell-surface purinergic receptors P1, P2X, or P2Y on target cells.

There are 21 different connexin genes in humans, and different sets of connexins are expressed in different cell types. The existence of this diversity, together with the generation of mutant mice with inactivating mutations in connexin genes, has highlighted the importance of connexins in a wide variety of cellular systems. Some cells express a single connexin that forms homotypic connexons. Most cells, however, express at least two connexins; these different proteins can assemble into heteromeric connexons, which in turn form heterotypic gap-junction channels. Diversity in channel composition leads to differences in channel permeability. For example, channels made from a 43-kDa connexin isoform, Cx43 — the most ubiquitously expressed connexin — are more than a hundred times as permeable to ADP and ATP as those made from Cx32 (32 kDa).

The permeability of gap junctions is regulated by post-translational modification of connexins (e.g., phosphorylation) and is sensitive to changes in environmental conditions such as intracellular pH and Ca^{2+} concentration, membrane potential, and the intercellular potential between adjacent interconnected cells ("voltage gating"). The N-termini of connexins appear to be especially important in the gating mechanism. Thus, as is the case for many ion channels (see Chapter 11), the channel in some gap junctions can be either opened or closed. One example of the physiological regulation of gap junctions occurs during mammalian childbirth. The smooth muscle cells in the mammalian uterus must contract strongly and synchronously during labor to expel the fetus. To facilitate this coordinated activity,

immediately before and during labor there is an approximately five- to tenfold increase in the amount of the major connexin in these cells, Cx43, and an increase in the number and size of gap junctions, which is reversed rapidly postpartum (following childbirth).

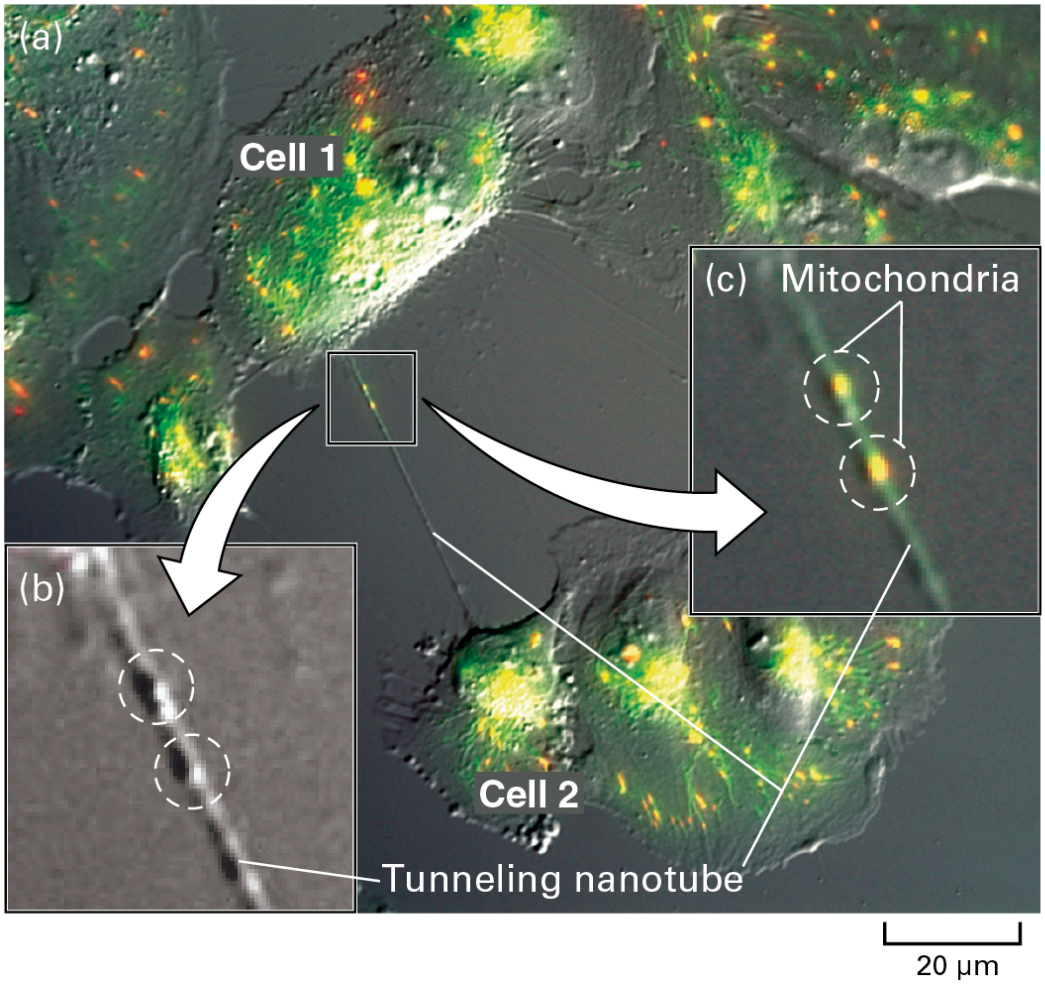
The assembly of connexins, their trafficking within cells, and the formation of functional gap junctions apparently depend on N-cadherin and its associated adapter proteins (e.g., α- and β-catenins, ZO-1, and ZO-2) as well as desmosomal proteins (plakoglobin, desmoplakin, and plakophilin-2). PDZ domains in ZO-1 and ZO-2 bind to the C-terminus of Cx43 and mediate its interaction with catenins and N-cadherin. The relevance of these relationships is particularly evident in the heart, which depends on gap junctions for rapid coordinated electrical coupling and on adjacent adherens junctions and desmosomes for mechanical coupling between cardiomyocytes to achieve the intercellular integration of electrical activity and movement required for normal cardiac function. It is noteworthy that ZO-1 serves as an adapter for adherens (see Figure 20-14), tight, and gap junctions, suggesting that this and other adapters can help integrate the formation and functions of these diverse junctions.

Mutations in connexin genes cause at least eight human diseases, including neurosensory deafness (Cx26 and Cx31), cataracts or heart malformations (Cx43, Cx46, and Cx50), and the X-linked form of Charcot-Marie-Tooth disease (Cx32), which is marked by progressive degeneration of peripheral nerves. ■

Tunneling Nanotubes Can Mediate Metabolic Coupling and Transfer Organelles Between Animal Cells

Tunneling nanotubes can serve as an alternative to gap junctions for moving molecules between cells. Tunneling nanotubes are tubelike projections of the plasma membrane that form a continuous channel connecting the cytosols of animal

cells (Figure 20-22) and can transfer chemical and electrical signals between cells in a manner analogous to structures in plants called *plasmodesmata* (described in Section 20.6). Tunneling nanotubes are typically unbranched, straight tubes and can have a wide variety of diameters (50–300 nm) and lengths (extending between cells from $<10\ \mu m$ to $>100\ \mu m$, they can be longer than several cell diameters). All tunneling nanotubes have actin filaments passing through the central channel, and in some types of cells they also contain microtubules. Remarkably, functional mitochondria can travel between cells by passing through tunneling nanotubes in cell culture (see Figure 20-22) and in vivo, thereby rescuing receiving cells that have mitochondrial defects or deficiencies. Thus the concept of metabolic coupling described for gap junctions can be extended to include the movement of small molecules and organelles through tunneling nanotubes. Pathogens may also use tunneling nanotubes to spread between cells.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-22 Microscopic visualization of a tunneling nanotube and mitochondria in cultured human cells. (a) Cultured human retinal pigment epithelial cells (ARPE-19 cell line) were incubated with a fluorescent dye (JC-1) that specifically stains mitochondria and then examined by a combination of conventional bright-field microscopy (see Chapter 4) to visualize the cells and fluorescence microscopy to visualize mitochondria (green intracellular fluorescence). A typical tunneling nanotube can be seen connecting cells 1 and 2. Inset (b) shows a higher magnification of the bright-field-only image with two bulges in the tunneling nanotube highlighted by dashed circles. Inset (c) shows a higher magnification of the same region of the combination image indicating two likely mitochondria within the tunneling nanotube at the positions of those bulges.

[D. Wittig et al., 2012, "Multi-Level Communication of Human Retinal Pigment Epithelial Cells via Tunneling Nanotubes." *PLoS One* **7**(3):e33195; https://doi.org/10.1371/journal.pone.0033195.]

KEY CONCEPTS OF SECTION 20.2

Cell-Cell and Cell–Extracellular Matrix Junctions and Their Adhesion Molecules

- Epithelial cells have distinct apical, basal, and lateral surfaces. Microvilli projecting from the apical surfaces of many epithelial cells considerably expand the cells' surface areas.
- Three major classes of cell junctions — anchoring junctions, tight junctions, and gap junctions — assemble epithelial cells into sheets and mediate communication between them (see Figures 20-1 and 20-11). Anchoring junctions can be further subdivided into adherens junctions, focal contacts, desmosomes, and hemidesmosomes.
- Adherens junctions and desmosomes are cadherin-containing anchoring junctions that bind the membranes of adjacent cells, giving strength and rigidity to the entire tissue.
- Cadherins are cell-adhesion molecules (CAMs) responsible for Ca^{2+}-dependent interactions among cells in epithelial and other tissues. They promote strong cell-cell adhesion by mediating both lateral (cis) and adhesive intercellular (trans) interactions.
- Adapter proteins that bind to the cytosolic domain of cadherins and other cell-cell and cell matrix adhesion receptors mediate the association of cytoskeletal and signaling molecules with the plasma membrane (see Figures 20-8 and 20-14). Strong cell-cell adhesion depends on the linkage of the interacting adhesion receptors to the cytoskeleton.
- Hemidesmosomes are integrin-containing anchoring junctions that attach cells to elements of the underlying extracellular matrix.
- Integrins are a large family of αβ heterodimeric cell-surface proteins that mediate both cell-cell and cell-matrix adhesions and inside-out and outside-in signaling in numerous tissues.
- Tight junctions block the diffusion of proteins and some lipids in the plane of the plasma membrane, contributing to the polarity of epithelial cells. They also limit and regulate the extracellular (paracellular) flow of water and solutes from one side of the epithelium to the other (see Figure 20-20). Two key integral membrane proteins found in tight junctions are occludin and claudin.
- Gap junctions are constructed of multiple copies of connexin proteins, which are assembled into a transmembrane channel that connects the cytosols of two adjacent cells (see Figure 20-21). Small molecules and ions can pass through gap junctions, permitting metabolic and electrical coupling of adjacent cells.
- Tunneling nanotubes are tubelike projections of plasma membranes that form a continuous channel connecting the cytosols of nearby animal cells (see Figure 20-22).

20.3 The Extracellular Matrix I: The Basal Lamina

In animals, the extracellular matrix (ECM) has multiple functions (see Table 20-2). The ECM helps organize cells into tissues and coordinates their cellular functions by activating intracellular signaling pathways that control cell growth, proliferation, and gene expression. The ECM can directly influence cell and tissue structure and function. In addition, it can serve as a repository for inactive or inaccessible signaling molecules (e.g., growth factors) that are released to function when the ECM is disassembled or remodeled by hydrolyases, such as proteases. Indeed, hydrolyzed fragments of ECM macromolecules can themselves have independent biological activity. The ensemble of proteins that compose the ECM itself and associated proteins that covalently modify (e.g., chemically cross-link, phosphorylate, cleave), bind to, or otherwise regulate the composition and structure of the ECM is called the *matrisome*. Proteomic (Chapter 3) and genomic analyses suggest that there are approximately 1030 and 1110 genes that encode the human and mouse matrisomes, respectively. Dysfunction of matrisome components can cause a wide variety of diseases that affect many different tissues and organs. It is noteworthy that there are ECM components, as well as extracellular domains of plasma-membrane proteins, that are phosphorylated on serine, threonine, or tyrosine side chains. Kinases that are present in the luminal compartments

of the secretory pathway and some that are apparently secreted into the extracellular space catalyze these phosphorylations.

Many functions of the ECM and, indeed, some features of the assembly of the ECM require transmembrane adhesion receptors, including the integrins, that bind directly to ECM components and that also interact, through adapter proteins, with the cytoskeleton. Adhesion receptors bind to three types of molecules abundant in the ECM of all tissues (see Table 20-1):

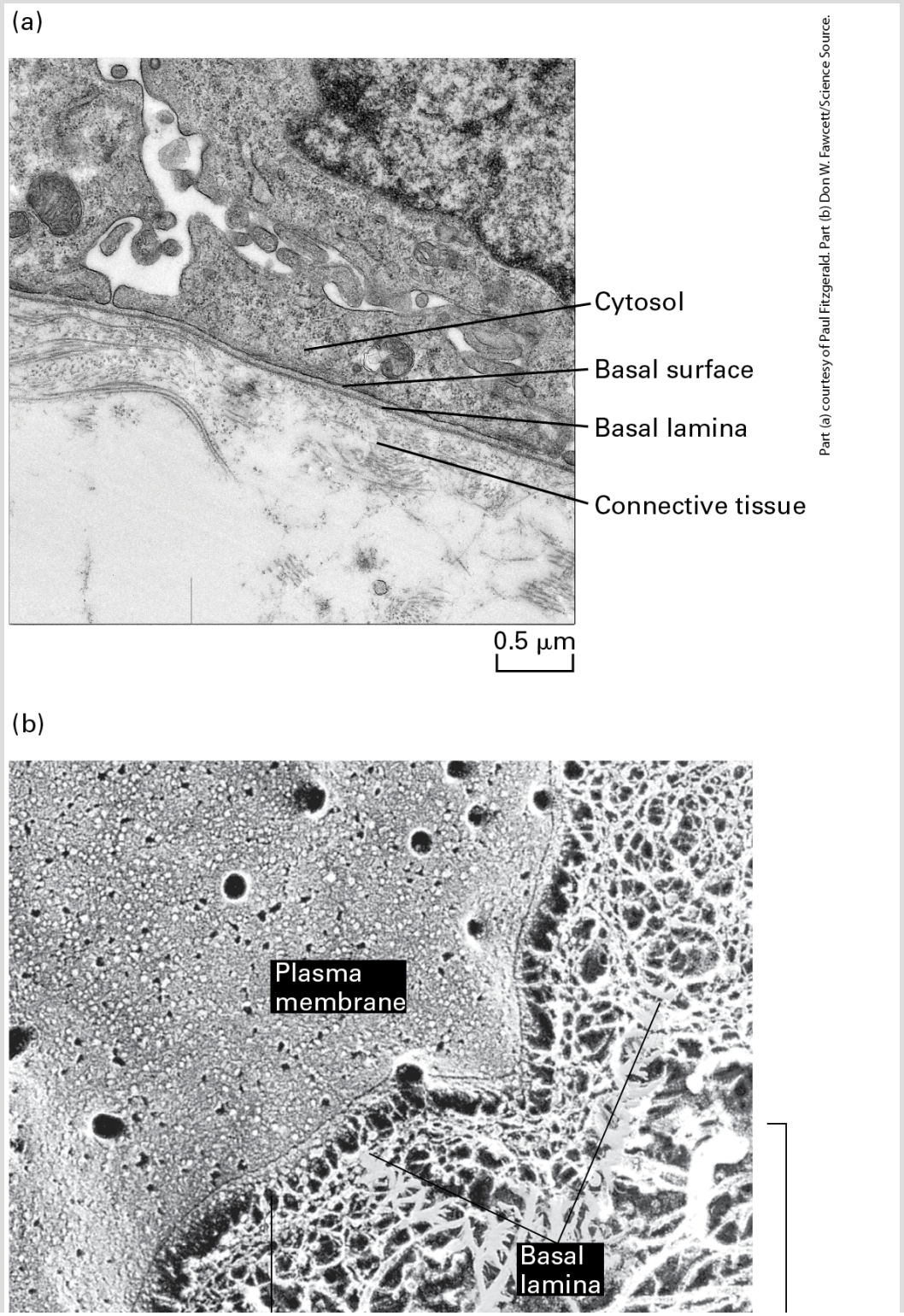
- **Proteoglycans**, a group of glycoproteins that cushion cells and bind a wide variety of extracellular and cell-surface molecules
- **Collagen** fibers, which provide structural integrity, mechanical strength, and resilience
- Soluble **multi-adhesive matrix proteins**, such as laminin and fibronectin, which bind to and cross-link adhesion receptors and other ECM components

We begin our description of the structures and functions of these major ECM components in the context of the basal lamina: the specialized sheet of ECM that plays a particularly important role in determining the overall architecture and function of epithelial tissues. In the following section, we discuss the ECM molecules commonly found in nonepithelial tissues, including connective tissue.

The Basal Lamina Provides a Foundation for Assembly of Cells into Tissues

In animals, most organized groups of cells in epithelial and nonepithelial tissues are underlain or surrounded by the basal lamina, a sheet-like meshwork of ECM components usually no more than 60–120 nm thick (Figure 20-23). The basal lamina is structured differently in different tissues. In columnar and other epithelia such as intestinal lining and skin, it is a foundation on which only one surface of the cells rests. In other tissues, such as muscle or fat, the basal lamina surrounds each cell. Basal laminae play important roles in regeneration after tissue damage and in embryonic development. For instance, the basal lamina helps four- and eight-celled embryos adhere together in a ball (see Figures 22-2 and 22-3). In the development of the nervous system, neurons migrate along ECM pathways that contain basal laminal components. In higher animals, two distinct basal laminae are employed to form a tight barrier that limits diffusion of molecules between the blood and the brain (blood-brain barrier), and in the kidney, a specialized basal lamina serves as a selectively permeable blood filter. In muscle, the basal lamina helps protect the cell membranes from damage during contraction and relaxation. Thus the basal lamina is important for organizing cells into tissues and distinct compartments, repairing tissues, forming permeability barriers, and guiding migrating cells

during development. It is therefore not surprising that its components have been highly conserved throughout evolution.



Part (a) courtesy of Paul Fitzgerald. Part (b) Don W. Fawcett/Science Source.

FIGURE 20-23 A basal lamina separates epithelial cells and some other cells from connective tissue. (a) Transmission electron micrograph of a thin section of cells (*top*) and underlying connective tissue (*bottom*). The electron-dense layer of the basal lamina can be seen to follow the undulations of the basal surfaces of the cells. (b) Electron micrograph of a quick-freeze deep-etch preparation of skeletal muscle, showing the plasma membrane, basal lamina, and surrounding connective-tissue collagen fibers. In this preparation, the basal lamina is revealed as a meshwork of filamentous proteins that associates with the plasma membrane and the thicker collagen fibers of the connective tissue.

Most of the ECM components in the basal lamina are synthesized by the cells that rest on it. Four ubiquitous protein components, each of which comprises multiple distinct repeating domains, are found in basal laminae (Figure 20-24):

- *Type IV collagen*, trimeric molecules with both rodlike and globular domains that form a two-dimensional network
- *Laminins*, a family of multi-adhesive, cross-shaped proteins that form a fibrous two-dimensional network with type IV collagen and that also bind to integrins and other adhesion receptors
- *Perlecan*, a large proteoglycan whose protein contains multiple copies of seven different structural domains (total of 48 domains) and which binds to and cross-links many ECM components and cell-surface molecules

- *Nidogen* (also called *entactin*), a rodlike molecule that cross-links type IV collagen, perlecan, and laminin, which helps incorporate other components into the ECM and also stabilizes basal laminae

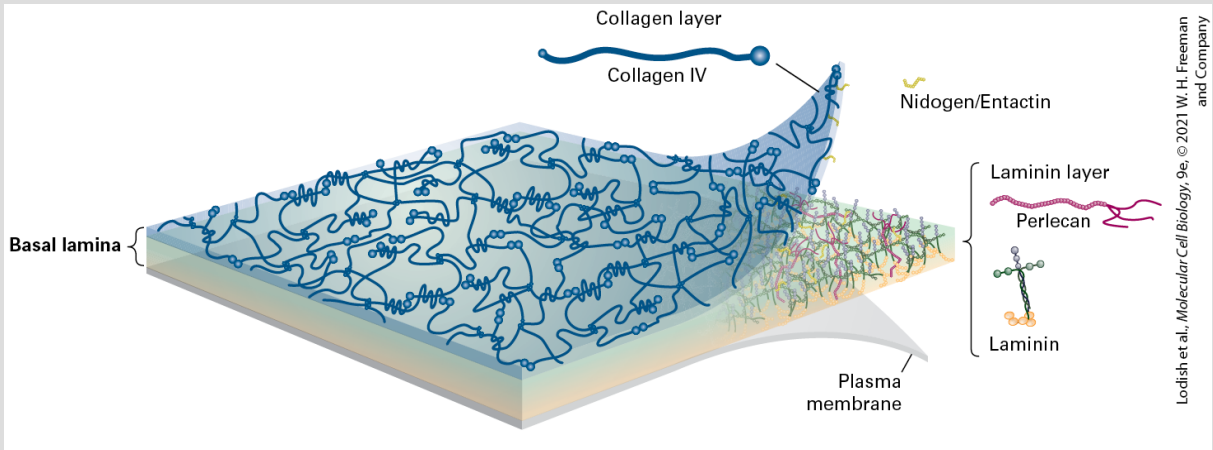


FIGURE 20-24 Major protein components of the basal lamina. Type IV collagen and laminin each form two-dimensional networks (see Figures 20-25 and 20-27), which are cross-linked by nidogen/entactin, perlecan, and laminin molecules, and which interact via laminins with the plasma membranes of adjacent cells.

Other ECM molecules, such as members of the evolutionarily ancient family of glycoproteins called *fibulins*, are incorporated into various basal laminae, depending on the tissue and the particular functional requirements of the basal lamina.

As depicted in Figure 20-1, one side of the basal lamina is linked to cells by adhesion receptors, including integrins in hemidesmosomes, which bind to laminin in the basal lamina. The other side of the basal lamina is anchored to the adjacent connective tissue by a layer of collagen fibers embedded in a proteoglycan-rich

matrix. In stratified squamous epithelia (e.g., skin; see Figure 20-10d), this linkage between the dermis and epidermis is mediated by anchoring fibrils of type VII collagen. Loss-of-function mutations in the gene encoding type VII collagen results in the rare and debilitating blistering disease called *dystrophic epidermolysis bullosa*. Together, the basal lamina and the anchoring collagen fibrils form the structure called the *basement membrane*.

Laminin, a Multi-Adhesive Matrix Protein, Helps Cross-Link Components of the Basal Lamina

Laminin, the principal multi-adhesive matrix protein in basal laminae, is a heterotrimeric protein comprising α, β, and γ chains. At least 16 laminin isoforms in vertebrates are assembled from 5 α, 3 β, and 3 γ chains, with each chain numbered to reflect the chain composition: laminin-111 (α1β1γ1) or laminin-511 (α5β1γ1). Each laminin isoform exhibits a distinctive pattern of tissue- and developmental stage–specific expression. As shown in Figure 20-25, many laminins are large, cross-shaped proteins (molecular weight of about 820,000), although some are Y- or rod-shaped. Globular domains at the N-terminus of each subunit bind to one another and thus mediate the self-assembly of laminins into mesh-like networks. Five globular *LG domains* at the C-terminus of the laminin α subunit mediate Ca^{2+}-dependent binding to cell-surface laminin receptors,

including certain integrins (see Table 20-4) as well as sulfated glycolipids, syndecan, and dystroglycan, which will be described further in Section 20.4. Some of these interactions are via negatively charged carbohydrates on the receptors. LG domains are found in a wide variety of other proteins and can mediate binding to steroids and proteins as well as carbohydrates. Laminin is the principal basal laminal ligand of integrins.

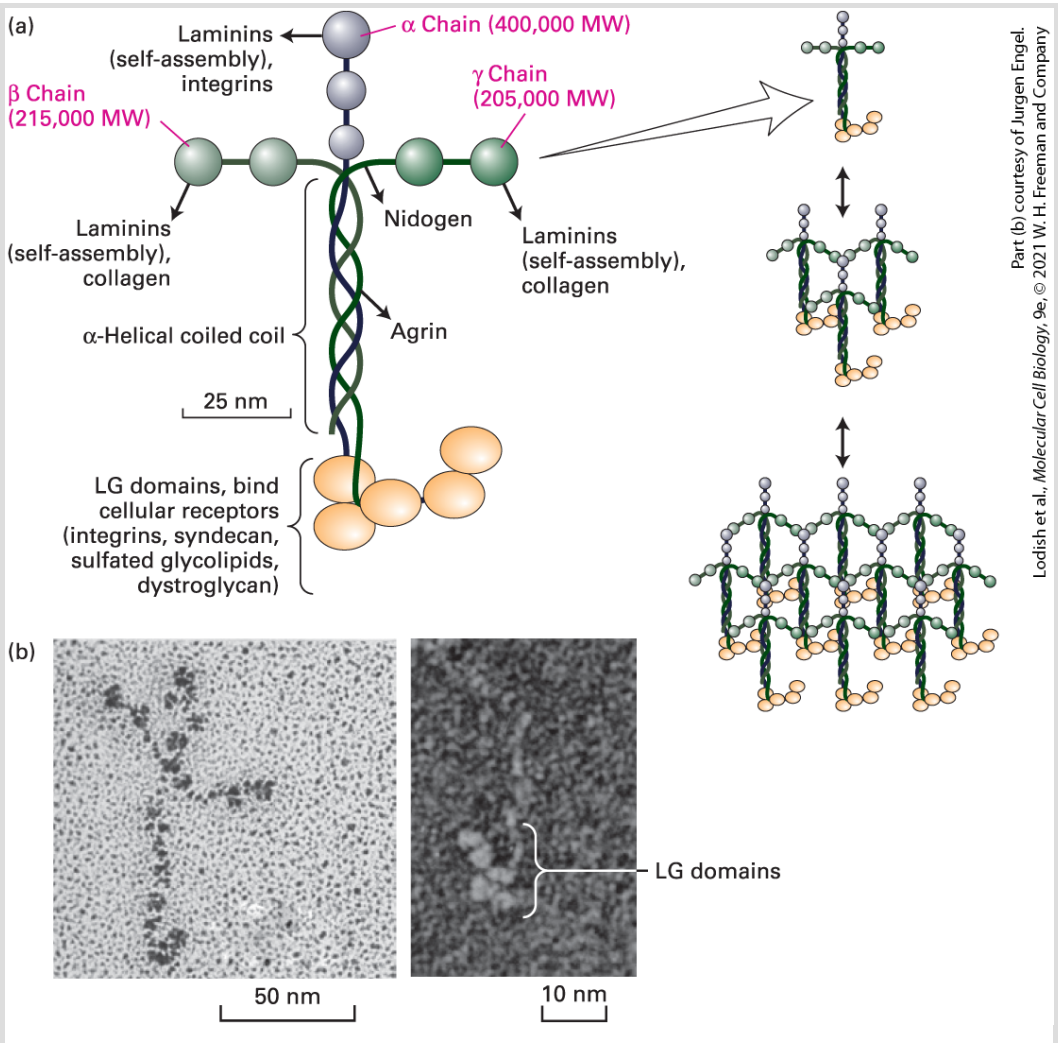


FIGURE 20-25 Laminin is a heterotrimeric multi-adhesive matrix protein found in all basal laminae. (a) Schematic model of cross-shaped laminin molecule showing the general shape, location of globular domains, and coiled-

coil region in which laminin's three chains are covalently linked by several disulfide bonds. Different regions of laminin bind to adhesion receptors and various matrix components (indicated by arrows and brackets). *Right:* Laminins assemble into a lattice via interactions between their N-terminal globular domains. See G. R. Martin and R. Timpl, 1987, *Annu. Rev. Cell Biol.* **3**:57–85; M. Durbeej, 2010, *Cell Tissue Res.* **339**:259–268; and S. Meinen et al., 2007, *J. Cell Biol.* **176**:979–993. (b) Electron micrographs of an intact laminin molecule, showing its characteristic cross shape (*left*), and the carbohydrate-binding LG domains near the C-terminus (*right*).

[*Left*: Republished with permission from Elsevier, from J. Engel et al., 1981, "Shapes, Domain Organizations and Flexibility of Laminin and Fibronectin, Two Multifunctional Proteins of the Extracellular Matrix," *J. Mol. Biol*. **150**(1):97–120; permission conveyed through Copyright Clearance Center, Inc. *Right*: Republished with permission from Elsevier, from R. Timpl et al., 2000, "Structure and Function of Laminin LG Modules," *Matrix Biol.* **19**(4):309–317; permission conveyed through the Copyright Clearance Center, Inc.]

Sheet-Forming Type IV Collagen Is a Major Structural Component of the Basal Lamina

Type IV collagen is, together with laminin, a principal structural component of all basal laminae and can bind to adhesion receptors, including some integrins. Collagen IV is one of at least 28 types of collagen in humans that participate in the formation of distinct ECMs in various tissues (Table 20-5). There are also at least 20 additional collagen-like proteins (such as host defense collagens) in the human proteome. Although collagen isoforms differ in certain

structural features and in their tissue distribution, all collagens are trimeric proteins made from three polypeptides, each encoded by one of at least 43 genes in humans, usually called *collagen α chains*. The three α chains in a collagen molecule can be identical (forming a homotrimer) or different (forming a heterotrimer). All or parts of the three-stranded collagen molecule can twist together into a special triple helix called a *collagenous* triple helix. When there is more than one triple-helical segment, these segments are joined by nonhelical regions of the protein, as we will see shortly for type IV collagen. Within a helical segment, each of the three α chains twists into a left-handed helix, and the three chains then wrap around one another to form a right-handed triple helix (Figure 20-26).

TABLE 20-5 • Selected Collagens

Type	Molecule Composition	Structural Features	Representative Tissues
Fibrillar Collagens			
I	$[\alpha 1(\mathrm{I})]_2[\alpha 2(\mathrm{I})]$	300-nm-long fibrils	Skin, tendon, bone, ligaments, dentin, interstitial tissues
II	$[\alpha 1(\mathrm{II})]_3$	300-nm-long fibrils	Cartilage, vitreous humor
III	$[\alpha 1(\mathrm{III})]_3$	300-nm-long fibrils; often with type I	Skin, muscle, blood vessels

V	$[\alpha 1(\mathrm{V})]_2[\alpha 2(\mathrm{V})]$, $[\alpha 1(\mathrm{V})]_3$	390-nm-long fibrils with globular N-terminal extension; often with type I	Cornea, teeth, bone, placenta, skin, smooth muscle
Fibril-Associated Collagens			
VI	$[\alpha 1(\mathrm{VI})][\alpha 2(\mathrm{VI})][\alpha 3(\mathrm{VI})]$	Lateral association with type I; periodic globular domains	Most interstitial tissues
IX	$[\alpha 1(\mathrm{IX})][\alpha 2(\mathrm{IX})][\alpha 3(\mathrm{IX})]$	Lateral association with type II; N-terminal globular domain; bound GAG	Cartilage, vitreous humor
Sheet-Forming and Anchoring Collagens			
IV	$[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$	Two-dimensional network	All basal laminae
VII	$[\alpha 1(\mathrm{VII})]_3$	Long fibrils	Below basal lamina of the skin
XV	$[\alpha 1(\mathrm{XV})]_3$	Core protein of chondroitin sulfate proteoglycan	Widespread; near basal lamina in muscle
Transmembrane Collagens			
XIII	$[\alpha 1(\mathrm{XIII})]_3$	Integral membrane protein	Hemidesmosomes in skin
XVII	$[\alpha 1(\mathrm{XVII})]_3$	Integral membrane protein	Hemidesmosomes in skin

Host Defense Collagens

Collectins	Oligomers of triple helix; lectin domains	Blood, alveolar space
C1q	Oligomers of triple helix	Blood (complement)
Class A scavenger receptors	Homotrimeric membrane proteins	Macrophages

SOURCE: Data from K. Kühn, 1987, in R. Mayne and R. E. Burgeson, eds., *Structure and Function of Collagen Types*, Academic Press, p. 2; M. Van Der Rest and R. Garrone, 1991, *FASEB J.* **5**:2814.

The collagen triple helix can form because of an unusual abundance of three amino acids in the α chains: glycine, proline, and a modified form of proline called *hydroxyproline* (see Figure 2-15). They make up the characteristic repeating sequence motif Gly-X-Y, where X and Y can be any amino acid but are often proline in position X and hydroxyproline in position Y, and less often lysine and hydroxylysine. Glycine is essential because its small side chain, a hydrogen atom, is the only one that can fit into the crowded center of the three-stranded helix (see Figure 20-26b). Hydrogen bonds help hold the three chains together. Although the rigid peptidyl-proline and peptidyl-hydroxyproline linkages are not compatible with formation of a classic single-stranded α helix, they stabilize the distinctive collagenous triple helix. The hydroxyl group in

hydroxyproline in the Y position helps hold its ring in a conformation that stabilizes the three-stranded helix.

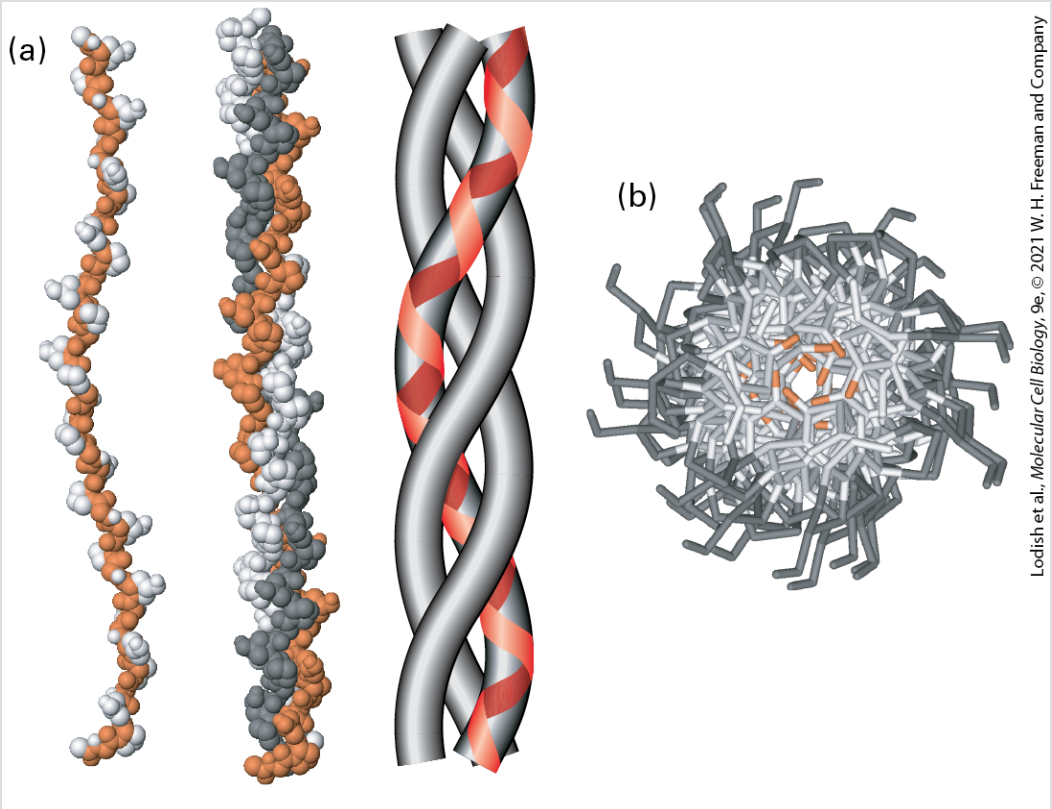


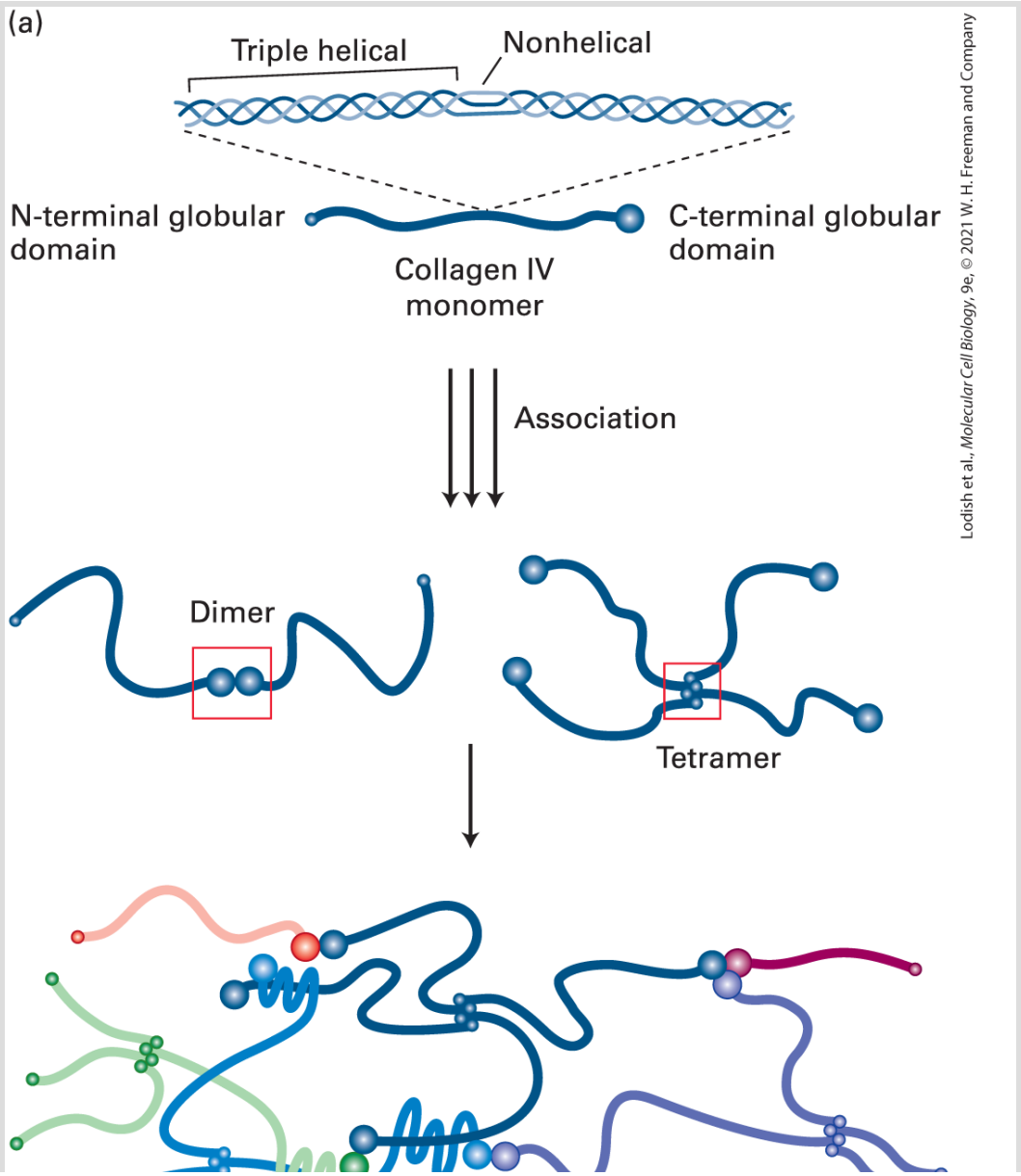
FIGURE 20-26 The collagen triple helix. (a) *Left:* Side view of the crystal structure of a polypeptide fragment whose sequence is based on repeating sets of three amino acids, Gly-X-Y, characteristic of collagen α chains. *Center:* Each chain is twisted into a left-handed helix, and three chains wrap around one another to form a right-handed triple helix. *Right:* A schematic model illustrates the triple-helical nature of the structure and shows the left-handed twist of the individual collagen α chains (red line). (b) View down the axis of the triple helix. The proton side chains of the glycine residues (orange) point into the very narrow space between the polypeptide chains in the center of the triple helix. In collagen mutations in which other amino acids replace glycine, the proton in glycine is replaced by larger groups that disrupt the packing of the chains and destabilize the triple-helical structure.

[Data from R. Z. Kramer et al., 2001, *J. Mol. Biol.* **311**:131, PDB ID 1bkv.]

There are several distinct cell-surface receptors for collagen IV and other types of collagen (other collagens are discussed in the next section). These cell-surface receptors include certain integrins, discoidin domain receptors 1 and 2 (which are tyrosine kinase receptors), glycoprotein VI (on platelets), leukocyte-associated Ig-like receptor-1, members of the mannose receptor family, and a modified form of the protein CD44. They can play critical roles in helping to assemble the ECM and in integrating cellular activity with the ECM.

The unique properties of each collagen isoform are due mainly to differences in (1) the number and lengths of the collagenous triple-helical segments; (2) the segments that flank or interrupt the triple-helical segments and that fold into other kinds of three-dimensional structures; and (3) covalent modification of the α chains (e.g., hydroxylation, glycosylation, oxidation, cross-linking). For example, the chains in type IV collagen are designated IVα chains. Mammals express six homologous IVα chains, which assemble into three different heterotrimeric type IV collagens with distinct properties. All subtypes of type IV collagen, however, form a 400-nm-long triple helix (Figure 20-27) that is interrupted about 24 times by nonhelical segments and flanked by a large globular domain at the C-terminus of the chain and a smaller globular domain at the N-terminus. The nonhelical regions introduce flexibility into the molecule. Through both lateral associations and interactions entailing the globular N- and C-termini, type IV collagen molecules assemble into a branching and covalently cross-linked, irregular two-dimensional

fibrous network that forms a lattice on which, together with the laminin lattice, the basal lamina is built (see Figures 20-24 and 20-27). Multiple, unusual sulfilimine (—S=N—) or thioether bonds between hydroxylysine (or lysine) and methionine residues covalently cross-link some adjacent C-terminal domains of type IV collagen molecules, and thus contribute to the stability of the network.



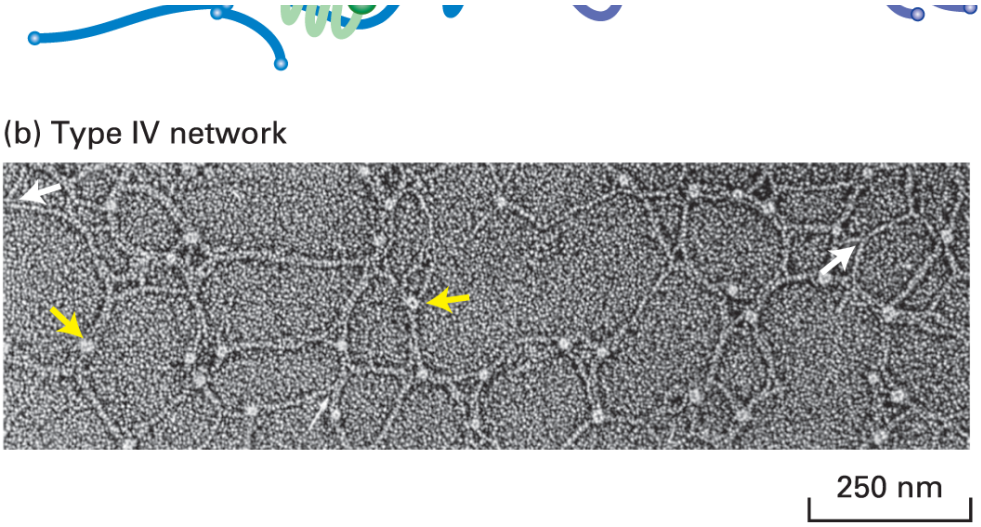


FIGURE 20-27 Structure and assembly of type IV collagen. (a) Schematic representation of type IV collagen. This 400-nm-long molecule has a small noncollagenous globular domain at the N-terminus and a large globular domain at the C-terminus. The collagenous triple helix is interrupted by nonhelical segments that introduce flexible kinks into the molecule. Lateral interactions between triple-helical segments, as well as head-to-head and tail-to-tail interactions between the globular domains, form dimers, tetramers, and higher order complexes, yielding a sheet-like network. See A. Boutaud et al., 2000, *J. Biol. Chem.* **275**:30716. (b) Electron micrograph of type IV collagen network formed in vitro. The lacy appearance results from the flexibility of the molecule, the side-to-side binding between triple-helical segments (white arrows), and the interactions between C-terminal globular domains (yellow arrows).

[Part (b) P. D. Yurchenco and G. C. Ruben, 1987, *J. Cell Biol.* **105**(6):2559–2568, Fig. 1c; https://doi.org/10.1083/jcb.105.6.2559.]

In the kidney, a double basal lamina called the *glomerular basement membrane* separates the epithelium that lines the urinary space from the endothelium that lines the surrounding blood-filled capillaries. Defects in this structure, which is responsible for

ultrafiltration of the blood and initial urine formation, can lead to renal failure. For instance, mutations that alter the C-terminal globular domain of certain IVα chains are associated with progressive renal failure as well as sensorineural hearing loss and ocular abnormalities, a condition known as *Alport's syndrome*. In *Goodpasture's syndrome*, a relatively rare autoimmune disease, antibodies bind to the α3 chains of type IV collagen found in the glomerular basement membrane and lungs. This binding sets off an immune response that causes cellular damage, resulting in progressive renal failure and pulmonary hemorrhage. ■

Perlecan, a Proteoglycan, Cross-Links Components of the Basal Lamina and Cell-Surface Receptors

Perlecan, the major secreted proteoglycan in basal laminae, consists of a large multidomain core protein (~470 kDa) to which polysaccharides are covalently attached. The core protein is made up of multiple repeats of five distinct domains, including laminin-like LG domains (3 copies), EGF-like domains (12 copies), and Ig domains (22 copies). The many globular repeats give it the appearance of an approximately 200-nm-long string of pearls when visualized by electron microscopy; hence the name *perlecan*. Perlecan contains three types of covalent polysaccharide chains: *N*-linked chains (see Chapter 14), *O*-linked chains, and

glycosaminoglycans (GAGs) (*O*-linked sugars and GAGs are discussed further in Section 20.4). GAGs are long, linear polymers of repeating disaccharides. Glycoproteins containing covalently attached GAG chains are called *proteoglycans*. Both the protein and the GAG components of perlecan contribute to its ability to incorporate into and define the structure and function of basal laminae. Because its multiple domains and its polysaccharide chains have distinct binding properties, perlecan binds to dozens of other molecules, including other ECM components (e.g., laminin, nidogen/entactin), cell-surface receptors, and polypeptide growth factors. Simultaneous binding to these molecules results in perlecan-mediated cross-linking. Perlecan can be found in basal laminae and in non–basal laminal ECM. The adhesion receptor dystroglycan can bind perlecan directly, via perlecan's LG domains, and indirectly, via its binding to laminin. In humans, mutations in the perlecan gene can lead either to dwarfism or to muscle abnormalities, apparently due to dysfunction of the neuromuscular junction that controls muscle firing.

KEY CONCEPTS OF SECTION 20.3

The Extracellular Matrix I: The Basal Lamina

- The matrisome is the ensemble of proteins that compose the ECM itself and associated proteins that covalently modify (e.g., chemically cross-link, phosphorylate, cleave) the ECM.
- The basal lamina, a thin meshwork of ECM molecules, separates most epithelia and other organized groups of cells from adjacent connective tissue. Together, the basal

lamina and the immediately adjacent collagen network form a structure called the *basement membrane*.

- Four ECM proteins are found in all basal laminae (see Figure 20-24): laminin (a multi-adhesive matrix protein), type IV collagen, perlecan (a proteoglycan), and nidogen/entactin.
- Adhesion receptors such as integrin anchor cells to the basal lamina, which in turn is connected to other ECM components (see Figure 20-1). Laminin in the basal lamina is the principal ligand of α6β4 integrin (see Table 20-4).
- Laminin and other multi-adhesive matrix proteins are multidomain molecules that bind multiple adhesion receptors and ECM components.
- The large, flexible molecules of type IV collagen interact end to end and laterally to form a mesh-like scaffold to which other ECM components and adhesion receptors can bind (see Figures 20-24 and 20-27).
- Type IV collagen is a member of the collagen family of proteins, which is distinguished by the presence of repeating tripeptide sequences of Gly-X-Y that give rise to the collagen triple-helical structure (see Figure 20-26). Different collagens are distinguished by the length and chemical modifications of their α chains and by the presence or absence of segments that interrupt or flank their triple-helical regions.
- Perlecan, a large, multidomain, secreted proteoglycan that is present primarily in basal laminae, binds many ECM components and adhesion receptors. Proteoglycans consist of membrane-associated or secreted core proteins covalently linked to one or more specialized polysaccharide chains called *glycosaminoglycans (GAGs)*.

20.4 The Extracellular Matrix II: Connective Tissue

Connective tissue, such as tendon and cartilage, differs from other solid tissues in that, rather than being filled primarily with cells, most of its volume is made up of extracellular matrix packed with insoluble protein fibers (see Figure 20-4). ECM in connective tissue has several key components, some of which are found in other types of tissues as well:

- *Collagens*, trimeric molecules that are often bundled together into fibers (fibrillar collagens)
- *Glycosaminoglycans (GAGs)*, specialized linear polysaccharide chains of specific repeating disaccharides that can be highly hydrated and confer diverse binding and physical properties (e.g., resistance to compression)
- *Proteoglycans*, glycoproteins containing one or more covalently bound GAG chains
- *Multi-adhesive proteins*, large multidomain proteins often comprising many copies ("repeats") of a few distinctive domains that bind to and cross-link a variety of adhesion receptors and ECM components
- *Elastin*, a protein that forms the amorphous core of elastic fibers

Collagen is the most abundant fibrous protein in connective tissue. Rubberlike elastin fibers, which can be stretched and relaxed, are

also present in deformable sites (e.g., skin, tendons, heart). The fibronectins, a family of multi-adhesive matrix proteins, form their own distinct fibrils in the ECM of most connective tissues. Although several types of cells are found in connective tissues, the various ECM components are produced largely by cells called **fibroblasts**. In this section, we explore the structure and function of the various ECM components in connective tissue, and we see how the ECM is degraded and remodeled by a variety of specialized proteases.

Fibrillar Collagens Are the Major Fibrous Proteins in the ECM of Connective Tissues

About 80–90 percent of the collagen in the body consists of **fibrillar collagens** (types I, II, and III), located primarily in connective tissues (see Table 20-5). Because of its abundance in tendon-rich tissue such as rat tail, type I collagen is easy to isolate and was the first collagen to be characterized. Its fundamental structural unit is a long (300 nm), thin (1.5 nm diameter) triple helix (see Figure 20-26) consisting of two α1(I) chains and one α2(I) chain, each 1050 amino acids in length. The triple-stranded molecules pack tightly together and wrap around one another, forming *microfibrils* that associate into higher order polymers called collagen *fibrils*, which in turn often aggregate into larger bundles called collagen *fibers* (Figure 20-28).

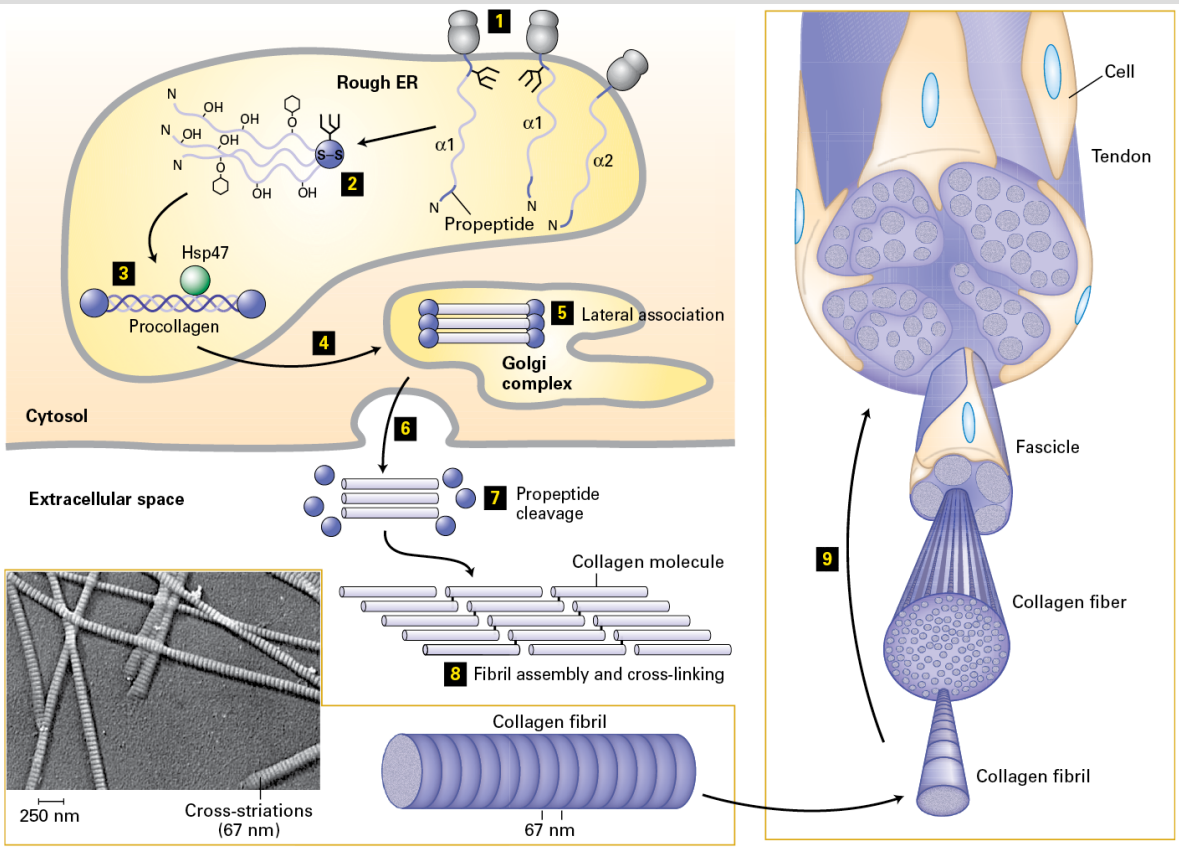


FIGURE 20-28 Biosynthesis of fibrillar collagens. Step 1: Procollagen α chains are synthesized on ribosomes associated with the endoplasmic reticulum (rough ER), and in the ER, asparagine-linked oligosaccharides are added to the C-terminal propeptide. Step 2: Propeptides associate to form trimers and are covalently linked by disulfide bonds, and selected residues in the Gly-X-Y triplet repeats are covalently modified [certain prolines and lysines are hydroxylated; galactose or galactose-glucose (hexagons) are attached to some hydroxylysines; prolines are cis → trans isomerized]. Step 3: The modifications facilitate zipper-like formation and stabilization of triple helices, and binding by the ER chaperone protein Hsp47, which may stabilize the helices or prevent premature aggregation of the trimers, or both. Steps 4 and 5: The folded procollagens are packaged into large transport vesicles with the help of specialized membrane proteins in the ER (not shown) and then transported to and through the Golgi complex, where some lateral association into small bundles takes place. The chains are then secreted (step 6), the N- and C-terminal propeptides are removed (step 7), and the trimers assemble into fibrils and are covalently cross-linked (step 8). The 67 nm staggering of the trimers gives the fibrils a striated appearance in electron micrographs (*inset*). Step 9: The fibrils can assemble into larger

and larger bundles, some of which form the tendons that attach muscle to bone. See A. V. Persikov and B. Brodsky, 2002, *Proc. Nat'l. Acad. Sci. USA* **99**:1101–1103.

[Electron micrograph republished with permission from John Wiley & Sons, Inc., from J. Gross, 1953, "Evaluation of Structural and Chemical Changes in Connective Tissue," *Ann. NY Acad. Sci*. **56**(4):674–683; permission conveyed through Copyright Clearance Center, Inc.]

Classes of collagen that are less abundant, but nevertheless important, include *fibril-associated collagens*, which link the fibrillar collagens to one another or to other ECM components; *sheet-forming* and *anchoring collagens*, which form two-dimensional networks in basal laminae (type IV) and connect the basal lamina in skin to the underlying connective tissue (type VII); *transmembrane collagens*, which function as adhesion receptors; and *host defense collagens*, which help the body recognize and eliminate pathogens. Interestingly, several collagens (e.g., types IX, XVIII, and XV) are also proteoglycans with covalently attached GAGs (see Table 20-5).

Fibrillar Collagen Is Secreted and Assembled into Fibrils Outside the Cell

Fibrillar collagens are secreted proteins, produced primarily by fibroblasts in the ECM. Collagen biosynthesis and secretion follow the normal pathway for a secreted protein, described in detail in Chapters 13 and 14. The collagen α chains are synthesized as longer precursors, called *pro-α chains*, by ribosomes attached to the endoplasmic reticulum (ER). The pro-α chains undergo a series of

covalent modifications and fold into triple-helical *procollagen* molecules before their transport from the ER to the Golgi complex and subsequent release from cells (see Figure 20-28). Specialized membrane proteins in the ER allow the packaging of the very large collagen molecules into large transport vesicles for transfer to the Golgi complex.

After the secretion of procollagen from the cell, extracellular peptidases remove the N-terminal and C-terminal propeptides. The resulting molecules, which consist almost entirely of a triple-stranded helix because of long stretches of the characteristic collagen repeating sequence motif Gly-X-Y, associate laterally to generate fibrils with a diameter of 50–200 nm. In fibrils, adjacent collagen molecules are displaced from one another by 67 nm, about one-quarter of their length. This staggered array produces a striated effect that can be seen in both light and electron microscopic images of collagen fibrils (see Figure 20-28, *inset*). The unique properties of the fibrillar collagens are mainly due to the formation of fibrils.

Short segments at either end of the fibrillar collagen α chains that are not composed of the repeating sequence motif Gly-X-Y, and thus are not triple-helical, are of particular importance in the formation of collagen fibrils. Lysine and hydroxylysine side chains in these segments are covalently modified by extracellular lysyl oxidases to form aldehydes in place of the amine group at the end of the side

chain. These reactive aldehyde groups form covalent cross-links with lysine, hydroxylysine, and histidine residues in adjacent molecules. The cross-links stabilize the side-by-side packing of collagen molecules and generate a very strong fibril. The removal of the terminal propeptides and covalent cross-linking take place in the extracellular space to prevent the potentially catastrophic assembly of large fibrils within the cell.

The post-translational modifications of pro-α chains are crucial for the formation of mature collagen molecules and their assembly into fibrils. Defects in these modifications have serious consequences, which ancient mariners frequently experienced. For example, ascorbic acid (vitamin C) is an essential cofactor for the hydroxylases responsible for adding hydroxyl groups to proline and lysine residues in pro-α chains. In cells deprived of ascorbate, as in the disease *scurvy*, the pro-α chains are not hydroxylated sufficiently to form stable triple-helical procollagen at normal body temperature, and the procollagen that forms cannot assemble into normal fibrils. Without the structural support of collagen, blood vessels, tendons, and skin become fragile. Fresh fruit in the diet can supply sufficient vitamin C to support the formation of normal collagen. Historically, British sailors were provided with limes to prevent scurvy, leading to their being called “limeys.” Mutations in lysyl hydroxylase genes can also cause connective-tissue defects. ■

Type I and II Collagens Associate with Nonfibrillar Collagens to Form Diverse Structures

Collagens differ in the structures of the fibers they form and in how these fibers are organized into networks. Of the predominant types of collagen found in connective tissues, type I collagen forms long fibers, whereas networks of type II collagen are more mesh-like. In tendons, for instance, the long type I collagen fibers connect muscles to bones and must withstand enormous forces. Because type I collagen fibers have great tensile strength, tendons usually can be stretched without being broken. Indeed, gram for gram, type I collagen is stronger than steel. Two quantitatively minor fibrillar collagens, type V and type XI, co-assemble into fibers with type I collagen, thereby regulating the structures and properties of the fibers. Incorporation of type V collagen, for example, results in smaller-diameter fibers.

Type I collagen fibrils are also used as the reinforcing rods in the construction of bone. Bones and teeth are hard and strong because they contain large amounts of dahllite, a crystalline calcium- and phosphate-containing mineral. Most bones are about 70 percent mineral and 30 percent protein, the vast majority of which is type I collagen. Bones form when certain cells (chondrocytes and osteoblasts) secrete collagen fibrils that are then mineralized by deposition of small dahllite crystals.

In many connective tissues, particularly skeletal muscle, proteoglycans and a fibril-associated collagen called *type VI collagen* are noncovalently bound to the sides of type I fibrils and may bind the fibrils together to form thicker collagen fibers (Figure 20-29a). Type VI collagen is unusual in that the molecule consists of a relatively short triple helix with globular domains at both ends. The lateral association of two type VI monomers generates an *antiparallel* dimer. The end-to-end association of these dimers through their globular domains forms type VI *microfibrils.* These microfibrils have a beads-on-a-string appearance, with about 60-nm-long triple-helical regions separated by 40-nm-long globular domains.

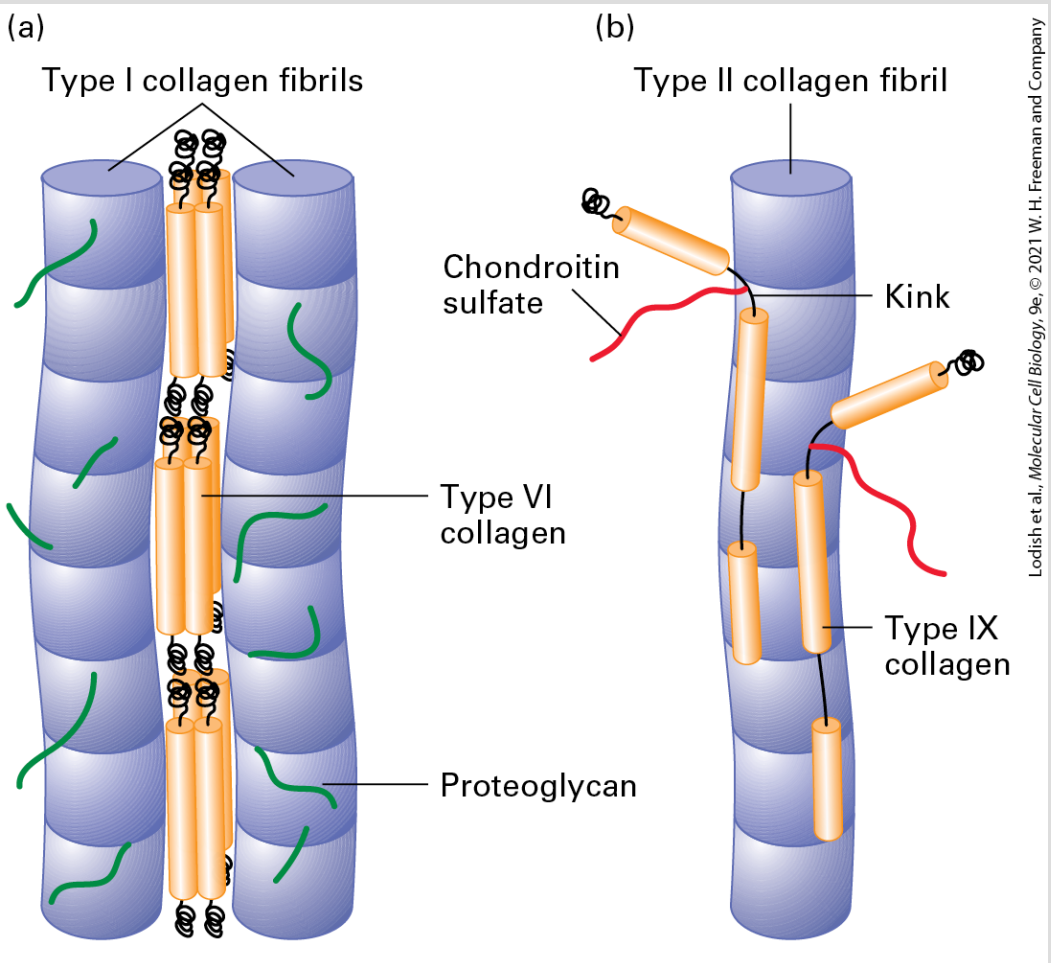


FIGURE 20-29 Interactions of fibrillar collagens with fibril-associated collagens. (a) In tendons, type I fibrils are all oriented in the direction of the stress applied to the tendon. Proteoglycans and type VI collagen bind noncovalently to type I fibrils, coating the surface. The microfibrils of type VI collagen, which contain globular and triple-helical segments, bind to type I fibrils and link them together into thicker fibers. See R. R. Bruns et al., 1986, *J. Cell Biol.* **103**:393. (b) In cartilage, type IX collagen molecules are covalently bound at regular intervals along type II fibrils. A chondroitin sulfate chain, covalently linked to the α2(IX) chain at the flexible kink, projects outward from the fibril, as does the globular N-terminal region. See L. M. Shaw and B. R. Olsen, 1991, *Trends Biochem. Sci.* **16**:191.

The fibrils of type II collagen, the major collagen in cartilage, are smaller in diameter than type I fibrils and are oriented randomly in a viscous proteoglycan matrix. The rigid collagen fibrils impart strength to the matrix and allow it to resist large deformations. Type II fibrils are cross-linked to matrix proteoglycans by type IX collagen, another fibril-associated collagen. Type IX collagen and several related types have two or three triple-helical segments connected by flexible kinks and a globular N-terminal segment (Figure 20-29b). The globular N-terminal segment of type IX collagen extends from the type II fibril at the end of one of its helical segments, as does a chondroitin sulfate GAG chain (chondroitin sulfate is described below) that is sometimes linked to one of the type IX chains. These protruding nonhelical structures are thought to anchor the type II fibril to proteoglycans and other components of the matrix. The interrupted triple-helical structure of type IX and related collagens prevents them from assembling into fibrils, although they can associate with fibrils formed from other collagen types and form covalent cross-links to them.

Mutations affecting type I collagen and its associated proteins cause a variety of human diseases. Certain mutations in the genes encoding the type I collagen α1(I) or α2(I) chains lead to *osteogenesis imperfecta*, or brittle-bone disease. Because every third position in a collagen α chain must be a glycine for the triple helix to form (see Figure 20-26), mutations of glycine to almost any other amino acid are deleterious, resulting in poorly formed and unstable helices. A

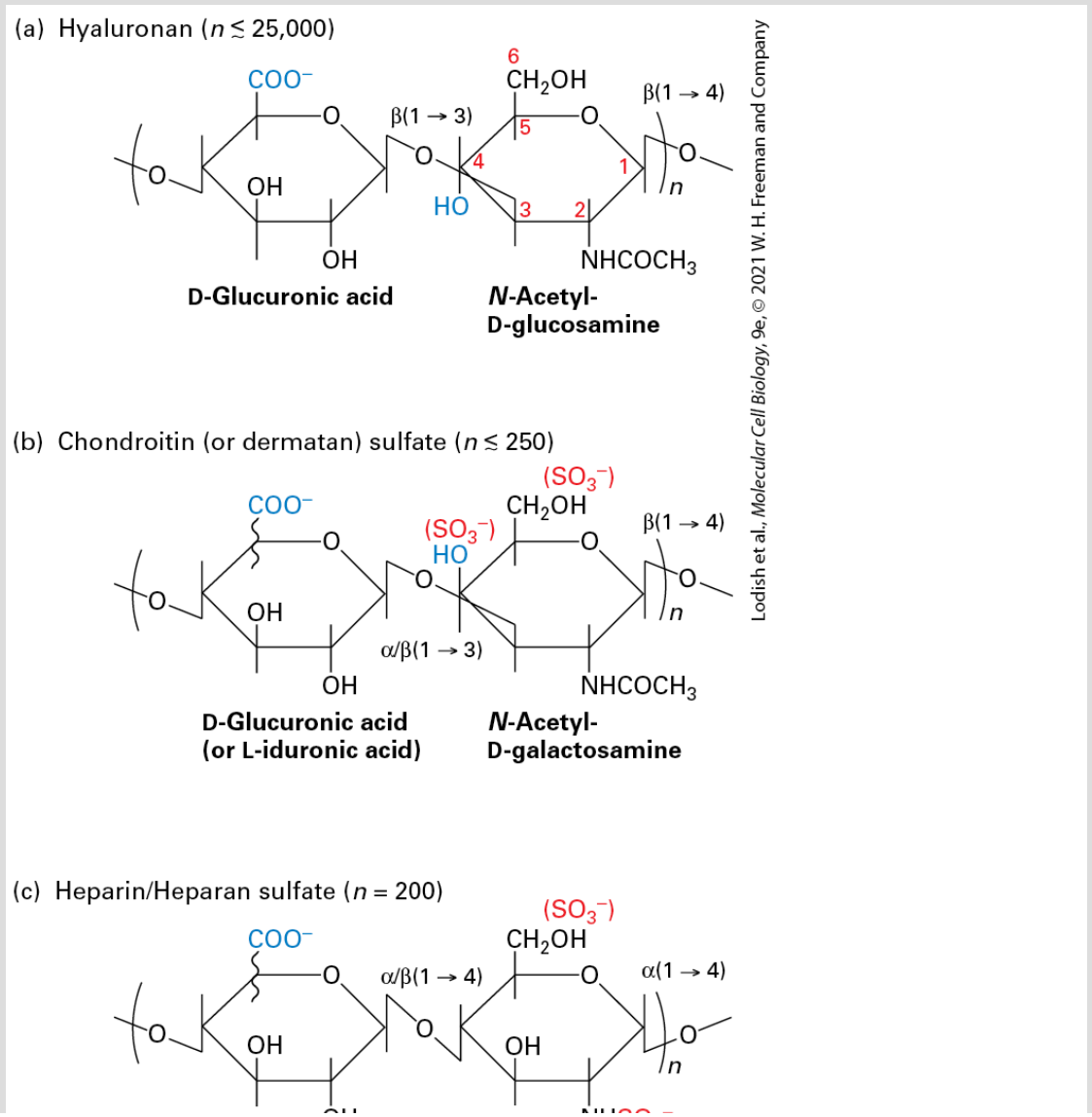
defect in only one of the three α chains in a collagen molecule can disrupt the whole molecule's triple-helical structure and function. A mutation in a single copy (allele) of either the α1(I) gene or the α2(I) gene, both located on autosomes, can cause this disorder. Thus it normally shows autosomal dominant inheritance (see Chapter 6).

Absence or malfunctioning of fibril-associated collagen in muscle tissue due to mutations in the type VI collagen genes cause dominant or recessive congenital muscular dystrophies with generalized muscle weakness, respiratory insufficiency, muscle wasting, and muscle-related joint abnormalities. Skin abnormalities have also been reported with type VI collagen disease. ■

Proteoglycans and Their Constituent GAGs Play Diverse Roles in the ECM

As we saw with perlecan in the basal lamina, proteoglycans play an important role in cell-matrix adhesion. Proteoglycans are a subset of secreted or cell-surface glycoproteins containing covalently linked, specialized polysaccharide chains called **glycosaminoglycans (GAGs)**. GAGs are long linear polymers of specific repeating disaccharides. Usually one sugar is either a uronic acid (D-glucuronic acid or L-iduronic acid) or D-galactose; the other sugar is *N*-acetylglucosamine or *N*-acetylgalactosamine (Figure 20-30). One or both of the sugars contain at least one anionic group (carboxylate

or sulfate). Thus each GAG chain bears many negative charges. GAGs are classified into several major types based on the nature of the repeating disaccharide unit: heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate, and hyaluronan (see Figure 20-30). A hypersulfated form of heparan sulfate called *heparin*, produced mostly by mast cells, plays a key role in allergic reactions. It is also used medically as an anticlotting drug because of its ability to activate a natural clotting inhibitor called *antithrombin III*.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

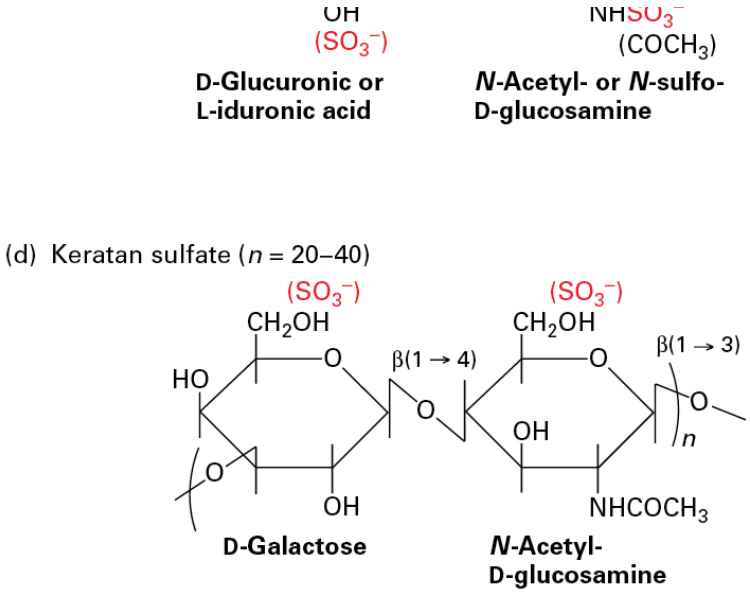


FIGURE 20-30 The repeating disaccharides of glycosaminoglycans (GAGs). Each of the four classes of GAGs is formed by polymerization of monomeric units into repeats of a particular disaccharide and subsequent modifications, including addition of sulfate groups and inversion (epimerization) of the carboxyl group on carbon 5 of D-glucuronic acid to yield L-iduronic acid. The squiggly lines represent covalent bonds that are oriented either above (D-glucuronic acid) or below (L-iduronic acid) the ring. Heparin is generated by hypersulfation of heparan sulfate, whereas hyaluronan is not sulfated.

With the exception of hyaluronan, which is discussed below, all the major GAGs occur naturally as components of proteoglycans. Like other secreted and transmembrane glycoproteins, proteoglycan core proteins are synthesized in the endoplasmic reticulum, and the GAG chains are assembled on and covalently attached to these cores in the Golgi complex. To generate heparan or chondroitin sulfate chains, a three-sugar linker is first attached to the hydroxyl side chains of certain serine residues in a core protein (Figure 20-31a);

thus these GAGs are ***O*-linked oligosaccharides**, several examples of which are shown in Figure 20-31.

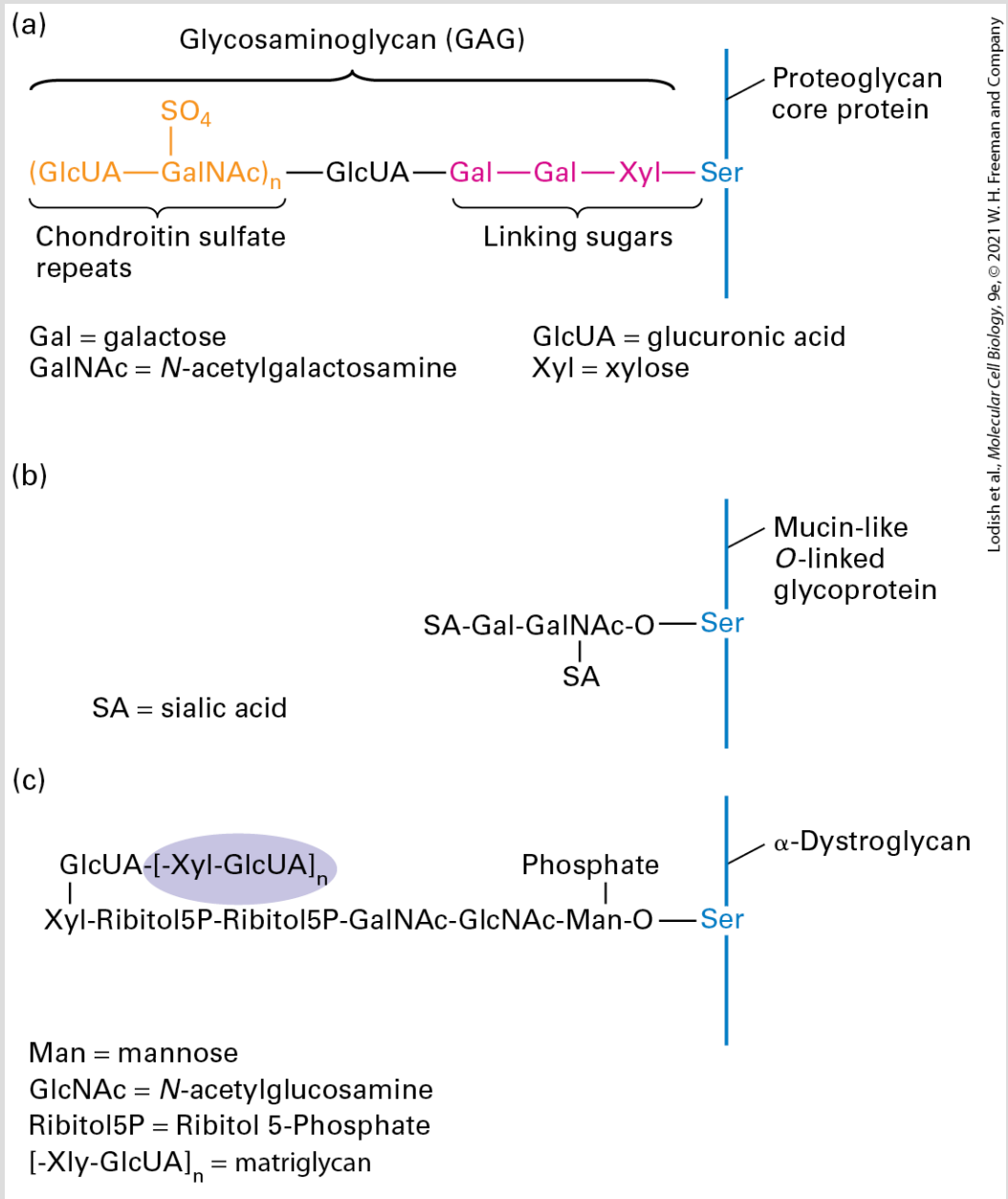


FIGURE 20-31 Hydroxyl *O*-linked oligosaccharides. (a) Synthesis of a GAG, in this case chondroitin sulfate, is initiated by transfer of a xylose (Xyl) residue to a serine residue in the core protein, most likely in the Golgi complex, followed by sequential addition of two galactose (Gal) residues. Glucuronic acid (GlcUA)

and *N*-acetylgalactosamine (GalNAc) residues are then added sequentially to these linking sugars and some of the GalNAc monomers are sulfated, forming the chondroitin sulfate chain. Heparan sulfate chains are connected to core proteins by the same three-sugar linker. Keratan sulfate GAGs are covalently attached to proteins via *N*-linked rather than *O*-linked connections. (b) Mucin-like *O*-linked chains are covalently bound to glycoproteins via an *N*-acetylgalactosamine (GalNAc). (c) In the integral membrane protein α-dystroglycan on cell surfaces, a complex polysaccharide is attached to hydroxyl side chains via an O-mannose (O-Man). A polymer of the GlcUA-Xyl disaccharide ($[-\mathrm{Xyl\text{-}GlcUA}]_n$) called *matriglycan* is attached to a phosphorylated *O*-linked mannose via an oligosaccharide containing *N*-acetylglucosamine (GlcNAc), ribitol 5-phosphate (Ribito5P), Xyl, and GlcUA. Matriglycan provides a binding site for ECM molecules, such as laminin and perlecan (described in detail in Section 20.5).

O-linked oligosaccharides are attached to the hydroxyl side chains of specific serine and theonine residues in a variety of secreted proteins and the exoplasmic (extracellular) domains of some membrane proteins. For example, mucin-like *O*-linked chains are covalently bound to glycoproteins via an *N*-acetylgalactosamine (GalNAc) monosaccharide to which are covalently attached a variety of other sugars, often including sialic acid (SA; Figure 20-31b). The high concentration of long *O*-linked chains on mucin glycoproteins on the apical plasma membranes of epithelial cells apparently contributes to the bending of the membranes and the formation of microvilli (see Figures 20-11, 17-1, and 17-2). The only *O*-linked glycoproteins in baker's yeast are mannose-linked (O-Man). In other eukaryotes (not in nematode worms or plants), certain specialized *O*-linked sugars are bound to proteins via mannose

monosaccharides. Some proteins, including classical cadherins and protocadherins (Section 20.2), have relatively short O-mannose linked oligosaccharides. Others have longer oligosaccharides attached via the O-mannose, such as that attached to the cell-surface protein dystroglycan (Figure 20-31c; described in detail in Section 20.5).

In contrast to protein-linked heparan or chondroitin sulfate chains, the linkers for the addition of keratan sulfate chains are oligosaccharide chains attached to asparagine residues; such ***N*-linked oligosaccharides** are present in many glycoproteins (see Chapters 13 and 14), although only a subset carry GAG chains. All GAG chains are elongated by the alternating addition of sugar monomers to form the disaccharide repeats characteristic of a particular GAG; the chains are often modified subsequently by the covalent linkage of small molecules such as sulfate. The mechanisms responsible for determining which proteins are modified with GAGs, the sequence of disaccharides to be added, the sites to be sulfated, and the lengths of the GAG chains are unknown. The ratio of polysaccharide to protein in all proteoglycans is much higher than that in most other glycoproteins.

Function of GAG Chain Modifications

Similar to the sequence of amino acids in proteins, the arrangement of the sugar residues in GAG chains and the modification of specific

sugars in those chains can determine their function and that of the proteoglycans that contain them. For example, groupings of certain modified sugars in the GAG chains of heparin sulfate proteoglycans can control the binding of growth factors to certain receptors or the activities of proteins in the blood-clotting cascade.

In the past, the chemical and structural complexity of proteoglycans posed a daunting barrier to analyzing and understanding their structures and their many diverse functions. In recent years, investigators employing classic and state-of-the-art biochemical techniques, mass spectrometry, and genetics have begun to elucidate the detailed structures and functions of these ubiquitous ECM molecules. The results of ongoing studies suggest that sets of sugar-residue sequences containing common modifications, rather than single unique sequences, are responsible for specifying distinct GAG functions. A case in point is a set of five-residue (pentasaccharide) sequences found in a subset of heparin GAGs that controls the activity of antithrombin III (ATIII), an inhibitor of the key blood-clotting protease thrombin. When these pentasaccharide sequences in heparin are sulfated at two specific positions (Figure 20-32), heparin can activate ATIII, thereby inhibiting clot formation. Several other sulfates can be present in the active pentasaccharide in various combinations, but they are not essential for the anticlotting activity of heparin. The rationale for generating sets of similar active sequences rather than a single, unique sequence is not well understood.

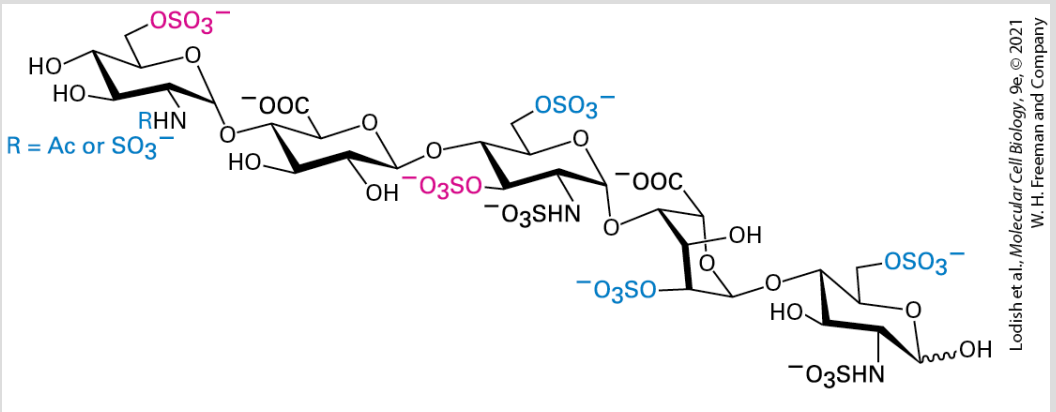


FIGURE 20-32 The pentasaccharide GAG sequence that regulates the activity of antithrombin III (ATIII). Sets of modified five-residue sequences in the much longer GAG called *heparin* with the composition shown here bind to ATIII and activate it, thereby inhibiting blood clotting. The sulfate groups in red type are essential for this heparin function; the modifications in blue type may be present but are not essential. Other sets of modified GAG sequences are thought to regulate the activity of other target proteins.

Diversity of Proteoglycans

The proteoglycans constitute a remarkably diverse group of molecules that are abundant in the ECM of all animal tissues and are also expressed on the cell surface. For example, of the five major classes of heparan sulfate proteoglycans, three are located in the ECM (perlecan, agrin, and type XVIII collagen) and two are cell-surface proteins. The latter include integral membrane proteins (syndecans) and GPI-anchored proteins (glypicans); the GAG chains in both types of cell-surface proteoglycans extend into the extracellular space. The sequences and lengths of proteoglycan core proteins vary considerably, and the number of attached GAG chains ranges from just a few to more than 100. Moreover, a core protein is

often linked to two different types of GAG chains, generating a hybrid proteoglycan. The basal laminal proteoglycan perlecan is primarily a heparan sulfate proteoglycan with three to four GAG chains, although it can sometimes have a bound chondroitin sulfate chain. Additional diversity in proteoglycans occurs because the numbers, compositions, and sequences of the GAG chains attached to otherwise identical core proteins can differ considerably. Laboratory generation and analysis of mutants with defects in proteoglycan production in *Drosophila*, *C. elegans*, and mice have clearly shown that proteoglycans play critical roles in development; for example, as participants in various signaling pathways (see Chapter 16 for examples in the TGF-β and Wnt pathways).

Syndecans are cell-surface proteoglycans expressed by epithelial and nonepithelial cells that bind to collagens and multi-adhesive matrix proteins such as fibronectin, anchoring cells to the ECM. Like that of many integral membrane proteins, the cytosolic domain of syndecan interacts with the actin cytoskeleton and in some cases with intracellular regulatory molecules. In addition, cell-surface proteoglycans such as syndecan bind many protein growth factors and other external signaling molecules, thereby helping to regulate cellular metabolism and function. For instance, syndecans in the hypothalamic region of the brain modulate feeding behavior in response to food deprivation. They do so by participating in the binding to cell-surface receptors of antisatiety peptides that help control feeding behavior. In the fed state, the syndecan extracellular

domain decorated with heparan sulfate GAG chains is released from the cell surface by proteolysis, thus suppressing the activity of the antisatiety peptides and feeding behavior. In mice engineered to overexpress the *syndecan-1* gene in the hypothalamic region of the brain and other tissues, normal control of feeding by antisatiety peptides is disrupted, and the animals overeat and become obese.

Hyaluronan Resists Compression, Facilitates Cell Migration, and Gives Cartilage Its Gel-Like Properties

Hyaluronan, also called *hyaluronic acid (HA)* or *hyaluronate*, is a nonsulfated GAG (see Figure 20-30a) made by a plasma-membrane-bound enzyme called *HA synthase* and is secreted directly into the extracellular space as it is synthesized. (A similar approach is used by plant cells to make the ECM component cellulose.) Hyaluronan is a major component of the ECM that surrounds migrating and proliferating cells, particularly in embryonic tissues. In addition, it forms the backbone of complex proteoglycan aggregates found in many ECMs, particularly cartilage. Because of its remarkable physical properties, hyaluronan imparts stiffness and resilience as well as a lubricating quality to many types of connective tissue such as joints.

Hyaluronan molecules range in length from a few disaccharide repeats to about 25,000. The typical hyaluronan in joints such as the elbow has 10,000 repeats for a total mass of 4×10^6 Da and a length of 10 μm (about the diameter of a small cell). Individual segments of a hyaluronan molecule fold into a rodlike conformation because of the β glycosidic linkages between the sugars and extensive intrachain hydrogen bonding. Mutual repulsion between negatively charged carboxylate groups that protrude outward at regular intervals also contributes to these locally rigid structures. Overall, however, hyaluronan is not a long, rigid rod like fibrillar collagen; rather, it is very flexible in solution, bending and twisting into many conformations, forming a random coil.

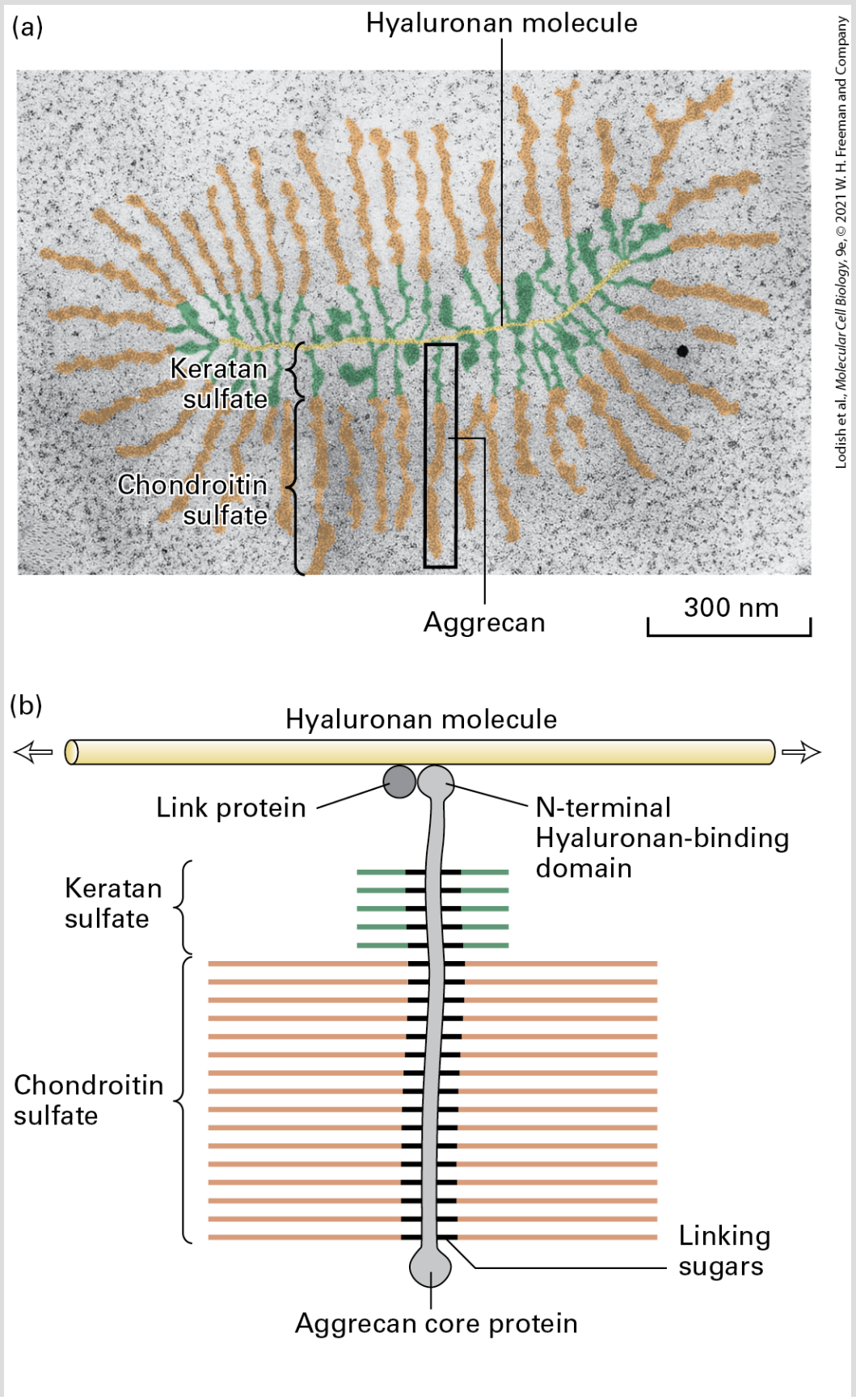
Because of the large number of anionic residues on its surface, the typical hyaluronan molecule binds a large amount of water and behaves as if it were a large hydrated sphere with a diameter of about 500 nm. As the concentration of hyaluronan increases, the long chains begin to entangle, forming a viscous gel. Even at low concentrations, hyaluronan forms a hydrated gel; when placed in a confining space, such as that between two cells, the long hyaluronan molecules tend to push outward. This outward pushing creates a swelling, or *turgor pressure*, within the extracellular space. In addition, the binding of cations by carboxylate (COO^-) groups on the surface of hyaluronan increases the concentration of ions and thus the osmotic pressure in the gel. As a result, large amounts of water are taken up, contributing to the turgor pressure. These

swelling forces give connective tissues their ability to resist compression forces, in contrast to collagen fibers, which are best able to resist stretching forces. Other highly charged GAG chains are similarly hydrated.

Hyaluronan is bound to the surface of many migrating cells by a number of adhesion receptors, such as the receptor called *CD44*, which contains hyaluronan-binding domains, each with a similar three-dimensional conformation. Because of its loose, hydrated, porous nature, the hyaluronan coating bound to cells appears to keep them apart from one another, giving them the freedom to move about and proliferate. The cessation of cell movement and the initiation of cell-cell attachments are frequently correlated with a decrease in hyaluronan, a decrease in hyaluronan receptors, and an increase in the extracellular enzyme hyaluronidase, which degrades hyaluronan in the matrix. These alterations of hyaluronan are particularly important during the many cell migrations that facilitate differentiation and in the release of a mammalian egg cell from its surrounding cells after ovulation.

The predominant proteoglycan in cartilage, called *aggrecan*, assembles with hyaluronan into very large aggregates, illustrative of the complex structures that proteoglycans sometimes form. The backbone of this proteoglycan aggregate is a long molecule of hyaluronan to which multiple aggrecan molecules are bound tightly but noncovalently (Figure 20-33). A single hyaluronan-aggrecan

aggregate, one of the largest macromolecular complexes known, can be more than 4 μm long and have a volume larger than that of a bacterial cell. These aggregates give cartilage its unique gel-like properties and its resistance to deformation, essential for distributing the load in weight-bearing joints.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-33 Structure of proteoglycan aggregate from cartilage. (a) Electron micrograph of an aggrecan aggregate from fetal bovine epiphyseal cartilage (artificially colored). Aggrecan core proteins are bound at ~40-nm intervals to a molecule of hyaluronan. (b) Schematic representation of an aggrecan monomer bound to hyaluronan. In aggrecan, both keratan sulfate and chondroitin sulfate chains are attached to the core protein. The N-terminal domain of the core protein binds noncovalently to a hyaluronan molecule. Binding is facilitated by a link protein, which binds to both the hyaluronan molecule and the aggrecan core protein. Each aggrecan core protein has 127 Ser-Gly sequences at which GAG chains can be added. The molecular weight of an aggrecan monomer averages 2×10^6. The entire aggregate, which may contain upward of 100 aggrecan monomers, has a molecular weight in excess of 2×10^8 and is about as large as the bacterium *Escherichia coli*.

[Part (a) republished with permission from Elsevier, from J. A. Buckwalter and L. Rosenberg, 1983, "Structural Changes During Development in Bovine Fetal Epiphyseal Cartilage," *Collagen Rel. Res.* **3**(6):489–504; permission conveyed through Copyright Clearance Center, Inc.]

The aggrecan core protein (~250,000 MW) has one N-terminal globular domain that binds with high affinity to a specific decasaccharide sequence within hyaluronan. This specific sequence is generated by covalent modification of some of the repeating disaccharides in the hyaluronan chain. The interaction between aggrecan and hyaluronan is facilitated by a link protein that binds to both the aggrecan core protein and hyaluronan (Figure 20-33b). Aggrecan and the link protein have in common a *link* domain, about 100 amino acids long, that is found in numerous ECM and cell-surface hyaluronan-binding proteins in both cartilaginous and noncartilaginous tissues. These proteins almost certainly arose in

the course of evolution from a single ancestral gene that encoded just this domain.

Fibronectins Connect Cells and ECM, Influencing Cell Shape, Differentiation, and Movement

Many different cell types synthesize **fibronectin**, an abundant multi-adhesive matrix protein found in all vertebrates. The discovery that fibronectin functions as an adhesion molecule stemmed from observations that it is present on the surfaces of normal fibroblasts, which adhere tightly to petri dishes in laboratory experiments, but is absent from the surfaces of tumorigenic (i.e., cancerous) cells, which adhere weakly. The 20 or so isoforms of fibronectin are generated by alternative splicing of the RNA transcript produced from a single gene (see Figure 5-28). Fibronectins are essential for the migration and differentiation of many cell types in embryogenesis. These proteins are also important for wound healing because they promote blood clotting and facilitate the migration of macrophages and other immune-system cells into the affected area.

Fibronectins help attach cells to the ECM by binding to other ECM components, particularly fibrillar collagens and heparan sulfate proteoglycans, and to adhesion receptors such as integrins (see

Figure 20-2). Through their interactions with adhesion receptors, fibronectins influence the shape and movement of cells and the organization of the cytoskeleton. Conversely, by regulating their receptor-mediated attachments to fibronectin and other ECM components, cells can sculpt the immediate ECM environment to suit their needs.

Fibronectins are dimers of two similar polypeptides linked at their C-termini by two disulfide bonds; each chain is about 60–70 nm long and 2–3 nm thick. Partial digestion of fibronectin with small amounts of proteases and analysis of the fragments showed that each chain comprises several functional regions with different ligand-binding specificities (Figure 20-34a). Each region, in turn, contains multiple copies of certain domain-encoding sequences that can be classified into one of three types. These domains are designated fibronectin type I, II, and III repeats, on the basis of similarities in amino acid sequence, although the sequences of any two repeats of a given type are not identical. These linked repeats give the molecule the appearance of beads on a string. The combination of the different repeats composing the regions confers on fibronectin its ability to bind multiple ligands.

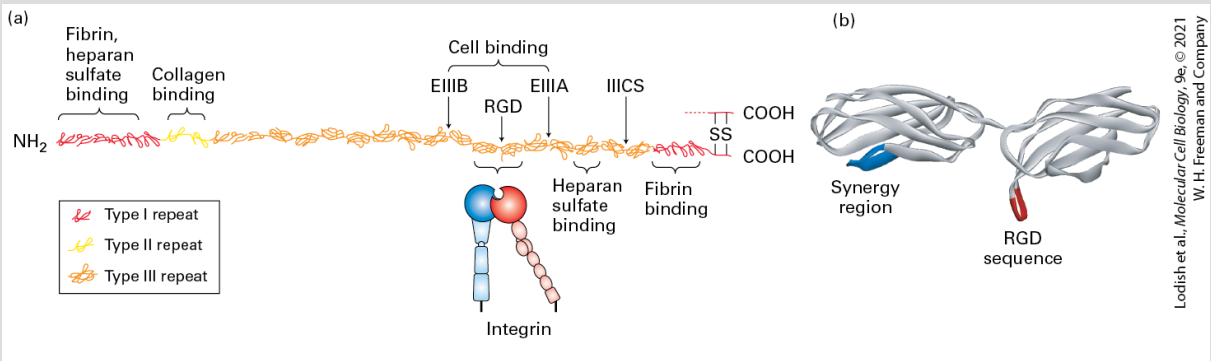
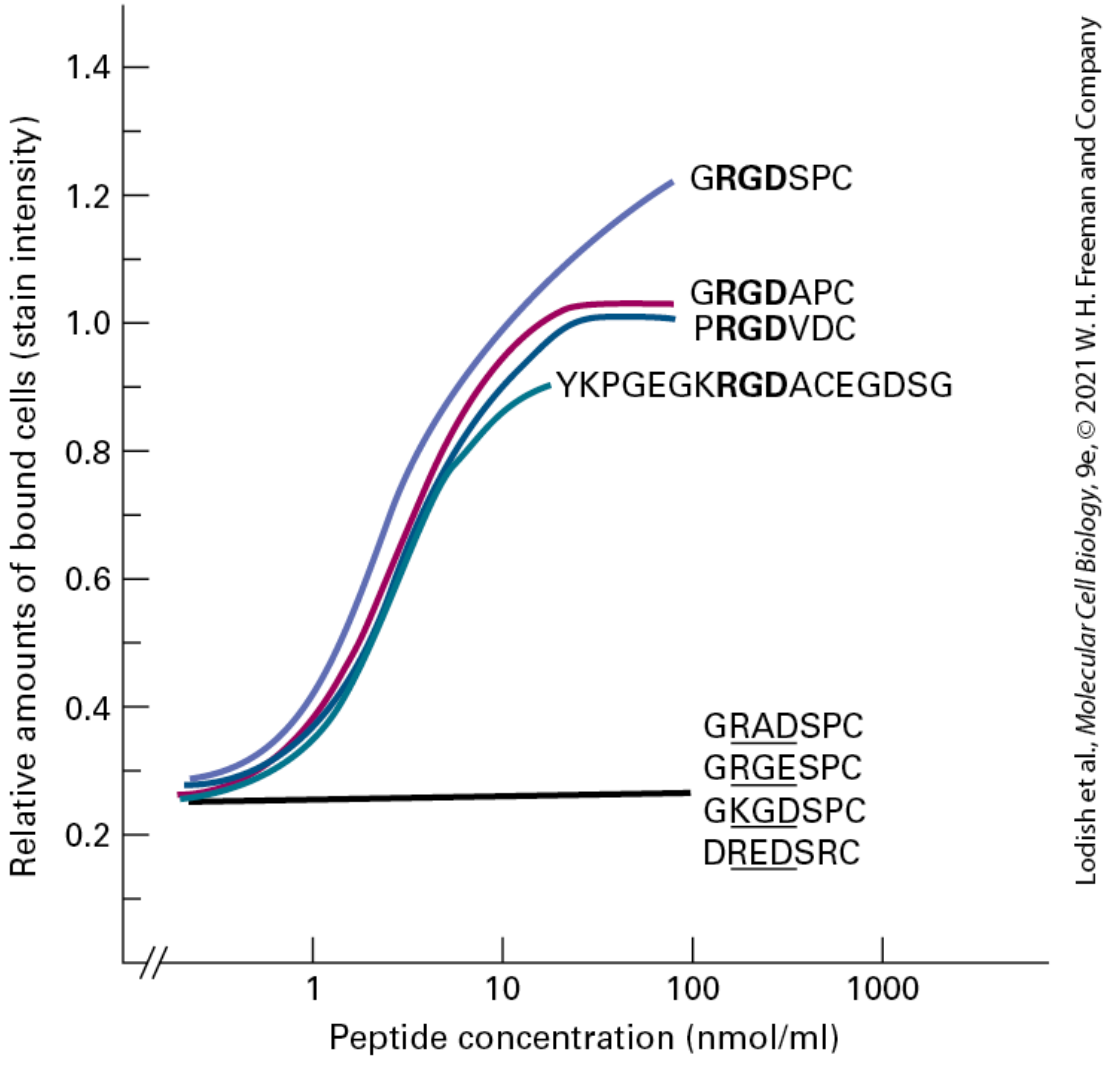


FIGURE 20-34 Organization of fibronectin and its binding to integrin. (a) Scale model of fibronectin is shown docked by two type III repeats to the extracellular domains of integrin. Only one of the two similar chains, which are linked by disulfide bonds near their C-termini, in the dimeric fibronectin molecule is shown. Each chain contains about 2446 amino acids and is composed of three types of repeating amino acid sequences (type I, II, or III repeats) or domains. The EIIIA, EIIIB — both type III repeats — and IIICS domain are variably spliced into the structure at locations indicated by arrows. Circulating fibronectin lacks EIIIA, EIIIB, or both. At least five different sequences may be present in the IIICS region as a result of alternative splicing (see Figure 5-28). Each chain contains several multi-repeat-containing regions, some of which contain specific binding sites for heparan sulfate, fibrin (a major constituent of blood clots), collagen, and cell-surface integrins. The integrin-binding domain is also known as the *cell-binding domain*. Structures of fibronectin's domains were determined from fragments of the molecule. (b) A high-resolution structure shows that the RGD sequence motif extends outward in a loop from its compact type III domain on the same side of fibronectin as the synergy region, which also contributes to high-affinity binding to integrins.

[Data from D. J. Leahy, I. Aukhil, and H. P. Erickson, 1996, *Cell* **84**:155, PDB ID 1fnf.]

One of the type III repeats in the cell-binding region of fibronectin mediates binding to certain integrins. The results of studies with synthetic peptides corresponding to parts of this repeat identified the tripeptide sequence Arg-Gly-Asp, called the *RGD motif*, as the minimal sequence within this repeat required for recognition by those integrins. In one study, heptapeptides with and without the

RGD motif were tested for their ability to mediate the adhesion of rat kidney cells to a culture dish. The results showed that heptapeptides containing the RGD motif mimicked intact fibronectin's ability to stimulate integrin-mediated adhesion, whereas variant heptapeptides lacking this sequence were ineffective (Figure 20-35).



EXPERIMENTAL FIGURE 20-35 A specific tripeptide sequence (RGD) in the cell-binding region of fibronectin is required for cell adhesion. The cell-binding region of fibronectin contains an integrin-binding hexapeptide sequence, GRGDSP in the single-letter amino acid code. Together with an additional C-terminal cysteine (C)

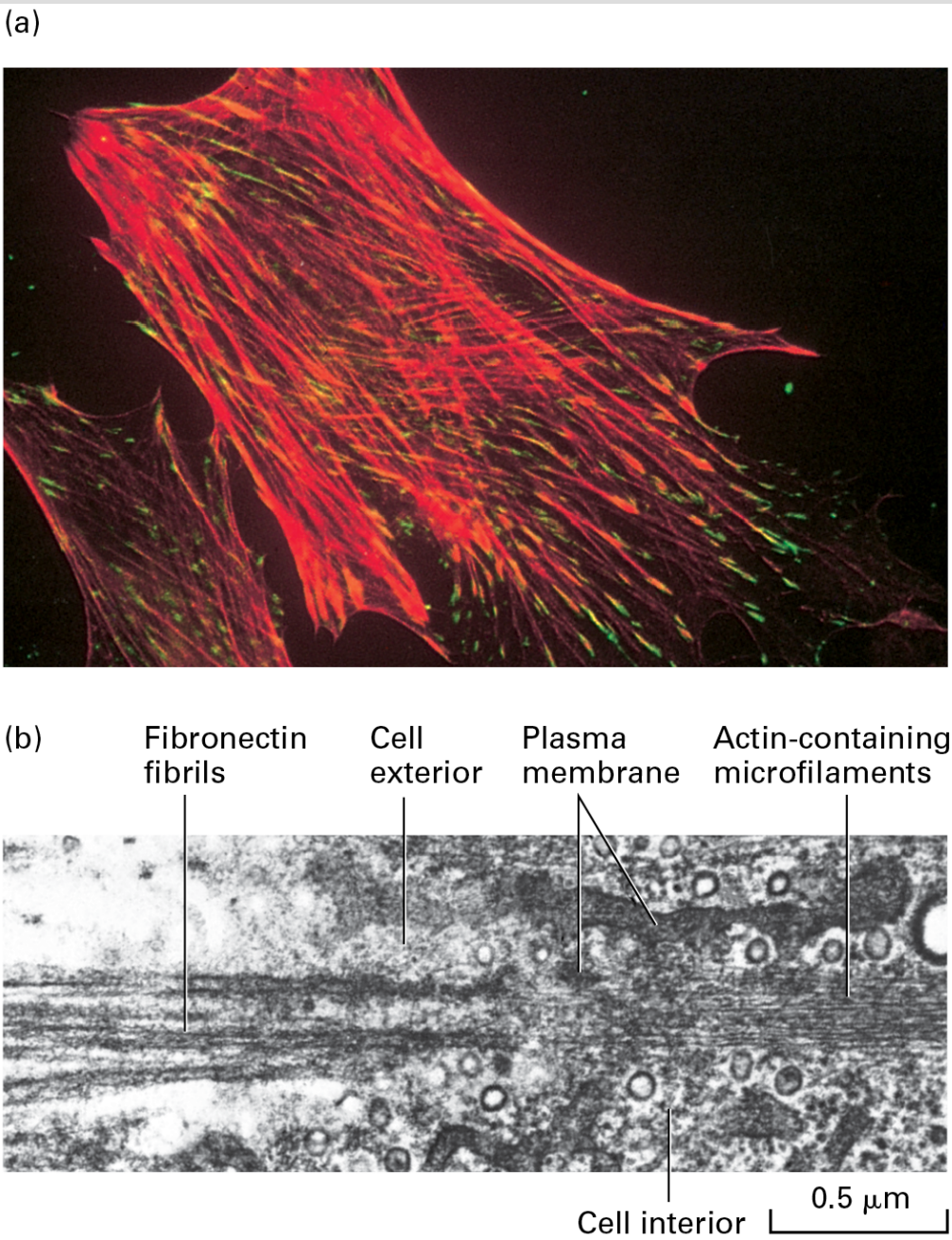
residue, this heptapeptide and several variants were synthesized chemically. Different concentrations of each synthetic peptide were added to polystyrene dishes that had the protein immunoglobulin G (IgG) firmly attached to their surfaces; the peptides were then chemically cross-linked to the IgG. Subsequently, cultured normal rat kidney cells were added to the dishes and incubated for 30 minutes to allow adhesion. After the unbound cells were washed away, the relative amounts of cells that had adhered firmly were determined by staining the bound cells with a dye and measuring the intensity of the staining with a spectrophotometer. The results shown here indicate that cell adhesion increased above the background level with increasing peptide concentration for those peptides containing the RGD motif, but not for the variants lacking this sequence (modification underlined).

[Data from M. D. Pierschbacher and E. Ruoslahti, 1984, *Proc. Nat'l. Acad. Sci. USA* **81**:5985.]

A three-dimensional model of fibronectin binding to integrin, based on partial structures of both fibronectin and integrin, has been assembled. In a high-resolution structure of the integrin-binding fibronectin type III repeat and its neighboring type III domain, the RGD motif is at the apex of a loop that protrudes outward from the molecule, in a position facilitating binding to integrins (Figure 20-34b). Although the RGD motif is required for binding to several different integrins, its affinity for integrins is substantially less than that of intact fibronectin or of the entire cell-binding region in fibronectin. Thus structural features near the RGD motif in fibronectins (e.g., parts of adjacent repeats, such as the synergy region; see Figure 20-34b) and in other RGD-containing proteins must enhance their binding to certain integrins. Moreover, the simple, soluble dimeric forms of fibronectin produced by the liver or by fibroblasts are initially in a nonfunctional conformation that

binds poorly to integrins because the RGD motif is not readily accessible. The adsorption of fibronectin onto a collagen matrix or basal lamina — or, experimentally, to a plastic tissue culture dish — results in a conformational change that enhances the ability of fibronectin to bind to cells. Possibly, this conformational change increases the accessibility of the RGD motif for integrin binding.

Microscopy and other experimental approaches (e.g., biochemical-binding experiments) have demonstrated the role of integrins in cross-linking fibronectin and other ECM components to the cytoskeleton. For example, the colocalization of cytoskeletal actin filaments and integrins within cells can be visualized by fluorescence microscopy (Figure 20-36a). The binding of cell-surface integrins to fibronectin in the ECM induces the actin cytoskeleton–dependent movement of some integrin molecules in the plane of the plasma membrane. The ensuing mechanical tension due to the relative movement of different integrins bound to a single fibronectin dimer stretches the fibronectin (see Figure 20-9), a mechanosensor, and promotes self-association of fibronectins into multimeric fibrils.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-36 Integrins mediate linkage between fibronectin in the ECM and the cytoskeleton. (a) Immunofluorescent micrograph of a fixed cultured fibroblast showing colocalization of the α5β1 integrin (green) and actin-containing stress fibers (red). The cell was incubated with two types of monoclonal antibodies: an integrin-specific antibody linked to a green-fluorescing dye and an actin-specific antibody linked to a red-fluorescing dye. Stress fibers are long bundles of actin microfilaments that radiate inward from

points where the cell contacts a substratum. At the distal ends of these fibers, near the plasma membrane, the coincidence of actin and fibronectin-binding integrin produces a yellow fluorescence. (b) Electron micrograph of the junction of fibronectin and actin fibers in a cultured fibroblast. Individual actin-containing 7 nm microfilaments, components of a stress fiber, end at the obliquely sectioned cell membrane. The microfilaments appear aligned with and in close proximity to the thicker, densely stained fibronectin fibrils on the outside of the cell.

[Part (a) J. L. Duband et al., 1988, *J. Cell Biol.* **107**:1385–1396, cover; https://doi.org/10.1083/jcb.107.4.1385. Part (b) republished with permission from Elsevier, from I. I. Singer, 1979, "The Fibronexus: A Transmembrane Association of Fibronectin-Containing Fibers and Bundles of 5 nm Microfilaments in Hamster and Human Fibroblasts," *Cell* **16**(3):675–685; permission conveyed through Copyright Clearance Center, Inc.]

The force needed to unfold and expose functional self-association sites in fibronectin is much less than that needed to disrupt fibronectin-integrin binding. Thus fibronectin molecules remain bound to integrin while cell-generated mechanical forces induce fibril formation. In effect, the integrins, through adapter proteins, transmit the intracellular forces generated by the actin cytoskeleton to extracellular fibronectin (inside-out signaling via mechanotransduction). Gradually, the initially formed fibronectin fibrils mature into highly stable matrix components by covalent cross-linking. In some electron micrographs, exterior fibronectin fibrils appear to be aligned in a seemingly continuous line with bundles of actin fibers within the cell (Figure 20-36b). These observations and the results from other studies provided the first example of a molecularly well-defined adhesion receptor forming a

bridge between the intracellular cytoskeleton and the ECM components — a phenomenon now known to be widespread.

Elastic Fibers Permit Many Tissues to Undergo Repeated Stretching and Recoiling

Elastic fibers are found in the ECM of a wide variety of tissues that are subject to mechanical strain or deformation, such as the lungs, which expand and contract during breathing (Figure 20-37a); the blood vessels, through which blood pulses due to the heartbeat; and the skin and many other tissues that stretch and contract. Elastic fibers permit the rubberlike reversible elastic stretching and recoiling of these tissues.

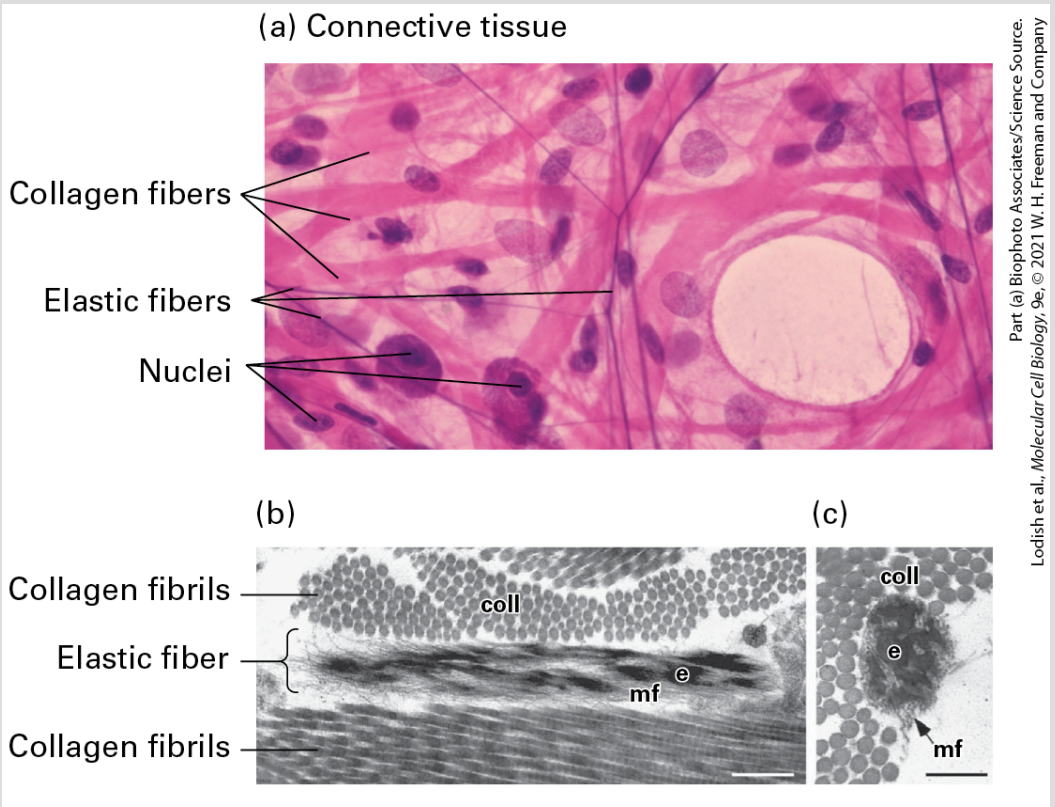


FIGURE 20-37 Elastic and collagen fibers in connective tissue. (a) Light-microscopic image of loose connective tissue from the lung. Collagen fibers (bundles of collagen fibrils) are stained pink; elastic fibers and nuclei of cells are stained purple. (b) Longitudinal and (c) cross-sectional electron microscopic images of elastic fibers and collagen fibrils (coll) in the skin of a mouse. The elastic fibers have a solid core of elastin (e) integrated into and surrounded by a bundle of microfibrils (mf). Scale bars, 0.25 μm.

[Parts (b) and (c) republished with permission from Elsevier, from J. Choi et al., 2009, "Analysis of Dermal Elastic Fibers in the Absence of Fibulin-5 Reveals Potential Roles for Fibulin-5 in Elastic Fiber Assembly," *Matrix Biol*. **28**(4):211–220; permission conveyed through Copyright Clearance Center, Inc.]

The major component of an elastic fiber, which can be several hundred to several thousand nanometers in diameter, is an insoluble, amorphous core composed of the protein **elastin**. Elastin

consists of aggregates of monomeric *tropoelastin* molecules that are covalently cross-linked via a lysyl oxidase–mediated process similar to that seen in collagen. Repetitive proline- and glycine-enriched hydrophobic sequence motifs contribute to the ability of tropoelastins to self-associate, extend under stress, and recoil efficiently after stretching. The elastin core is surrounded by a collection of 10–12-nm-diameter microfibrils (Figure 20-37b) made up of the proteins fibrillin, fibulin, and associated proteins such as LTBPs (*l*atent *T*GF-β *b*inding *p*roteins). The microfibrils serve as scaffolds for the assembly of the elastic fiber's core. Elastin-free microfibrils are found in the eye, where they transmit muscular force to reshape the lens for focusing and may provide structural support to the cornea.

Similar to other components of the ECM, microfibrils participate in cell signaling. In the secretory pathway, LTBPs bind the inactive form of the secreted signaling protein transforming growth factor β (TGF-β; a cytokine, see Figure 16-22) prior to their co-secretion and incorporation into microfibrils. Biomechanical stress mediated by cell-surface integrins binding to and pulling on the LTBP/TGF-β complex or proteolytic cleavage are thought to be the direct cause of active TGF-β release from the ECM and subsequent signaling (see Figure 16-22). Similar integrin-mediated activation of the LTBP/TGF-β complex bound to other ECM proteins provides active TGF-β for regulation of a wide variety of biological processes.

A variety of diseases, many involving skeletal and cardiovascular abnormalities, are consequences of mutations in the genes encoding the structural proteins of elastic fibers or the proteins that contribute to their proper assembly. For example, mutations in the *fibrillin-1* gene cause Marfan syndrome, whose varied symptoms can include bone overgrowth, loose joints, abnormally long extremities and face, and cardiovascular defects due to weakness in the walls of the aorta and other blood vessels. There has been considerable speculation that President Abraham Lincoln's unusually tall, elongated body may have been a consequence of Marfan syndrome. ■

In mammals, most tropoelastin synthesis occurs immediately before and after birth during the late fetal and neonatal periods. Thus most of the body's elastin needs to be very durable, lasting an entire lifetime. The extraordinary stability of elastin has been measured in a variety of ways. Pulse-chase experiments (see Chapter 3) using radioactive amino acid administration can be used to measure the life span of elastin in animals. In humans, two other methods employed to study the longevity of elastin have revealed that the mean lifetime of an elastin molecule in human lungs is about 70 years! The first method takes advantage of a naturally occurring phenomenon: the slow, natural rate of conversion of L-aspartic acid — incorporated into proteins during their synthesis — to D-aspartic acid. Thus the age of a long-lived protein can be estimated using chemical analysis to determine what fraction of its L-aspartic acid

has been converted over time to the D isomer, together with knowledge of the age of the tissue from which it was isolated. The second method is a variation on the classic pulse-chase experiments used in the laboratory. As a consequence of nuclear weapons testing in the 1950s and 1960s, ^{14}C was introduced into the atmosphere and hence the food chain. This environmental ^{14}C has been used as the radioactive *pulse* in what is essentially a pulse-chase experiment to determine the stability of proteins of interest.

Metalloproteases Remodel and Degrade the Extracellular Matrix

Many key physiological processes, including morphogenesis during development, control of cellular proliferation and motility, response to injury, and even survival, require not only the production of ECM, but also its remodeling or degradation. Because of its enormous importance as a key element in the extracellular environment of multicellular organisms, remodeling and degradation of the ECM must be carefully controlled. Degradation of the ECM is often mediated by zinc-dependent ECM metalloproteases. Given the wide array of ECM components, it is not surprising that there are many such metalloproteases with varying substrate specificities and sites of expression. In many cases, their names incorporate the names of their substrates, as for the metalloproteases called *collagenases*, *gelatinases*, *elastases*, and *aggrecanases*. Some are secreted into the

extracellular fluid, and others are closely associated with the plasma membranes of cells, either tightly bound in a noncovalent association with the membrane or as integral membrane proteins. Many are initially synthesized as inactive precursors that must be specifically activated to function.

ECM metalloproteases are divided into three major subgroups based on the enzymes' structures: **matrix metalloproteases (MMPs)** (of which there are 23 in humans), *a d*isintegrin *a*nd *m*etalloproteinases (ADAMs), and *ADAM*s with *t*hrombo*s*pondin motifs (ADAMTSs). These proteases can degrade ECM components as well as non-ECM components such as adhesion receptors. Indeed, a key function of ADAMs is cleaving extracellular domains from integral membrane proteins. In Chapter 16 we learned how ADAMs play key roles in signaling by Notch and EGF receptors. One mechanism used to control the activities of these proteases is the production of protein inhibitors called *TIMPs* (*t*issue *i*nhibitors of *m*etallo*p*roteinases) and *RECK* (*re*version-inducing–*c*ysteine-rich protein with *k*azal motifs). Some of these inhibitors have their own cell-surface receptors and functions independent of their ability to inhibit metalloproteinases. ECM-degrading proteases are associated with a variety of diseases, the best known of which is metastatic (spreading) cancer (see Chapter 25).

KEY CONCEPTS OF SECTION 20.4

The Extracellular Matrix II: Connective Tissue

- Connective tissue, such as tendon and cartilage, differs from other solid tissues in that most of its volume is made up of extracellular matrix (ECM) rather than cells (see Figure 20-4).
- The synthesis of fibrillar collagen (e.g., types I, II, and III) begins inside the cell with the chemical modification of newly made α chains and their assembly into triple-helical procollagen within the endoplasmic reticulum. After secretion, procollagen molecules are cleaved, associate laterally, and are covalently cross-linked into bundles called *fibrils*, which can form larger assemblies called *fibers* (see Figure 20-28).
- The various collagens are distinguished by the ability of their helical and nonhelical regions to associate into fibrils, to form sheets, or to cross-link other collagen types (see Table 20-5).
- Proteoglycans consist of membrane-associated or secreted core proteins covalently linked to one or more glycosaminoglycan (GAG) chains, which are linear polymers of disaccharides that are often modified by sulfation.
- Cell-surface proteoglycans such as the syndecans facilitate cell-matrix interactions and help present certain external signaling molecules to their cell-surface receptors.
- Hyaluronan, a highly hydrated GAG, is a major component of the ECM of migrating and proliferating cells. Certain adhesion receptors bind hyaluronan to cells.
- Large proteoglycan aggregates containing a central hyaluronan molecule noncovalently bound to the core proteins of proteoglycan molecules (e.g., aggrecan) contribute to the ability of the matrix to resist compression forces (see Figure 20-33).
- Fibronectins are abundant multi-adhesive matrix proteins that play a key role in migration and cellular differentiation. They contain binding sites for integrins and ECM components (collagens, proteoglycans) and thus can attach cells to the ECM (see Figure 20-34).
- The tripeptide RGD motif Arg-Gly-Asp, found in fibronectins and some other matrix proteins, is recognized by several integrins.
- Elastic fibers permit repeated stretching and recoiling of tissues because of their highly elastic core of cross-linked, amorphous elastin, which is surrounded by a network of microfibrils that help assemble the fibers and regulate signaling mediated by TGF-β.
- The remodeling or degradation of ECM is mediated by a large number of secreted and cell-membrane-associated zinc metalloproteases that fall into several families (MMPs, ADAMs, ADAMTSs) and whose activities are regulated by protein inhibitors (TIMPs and RECK).

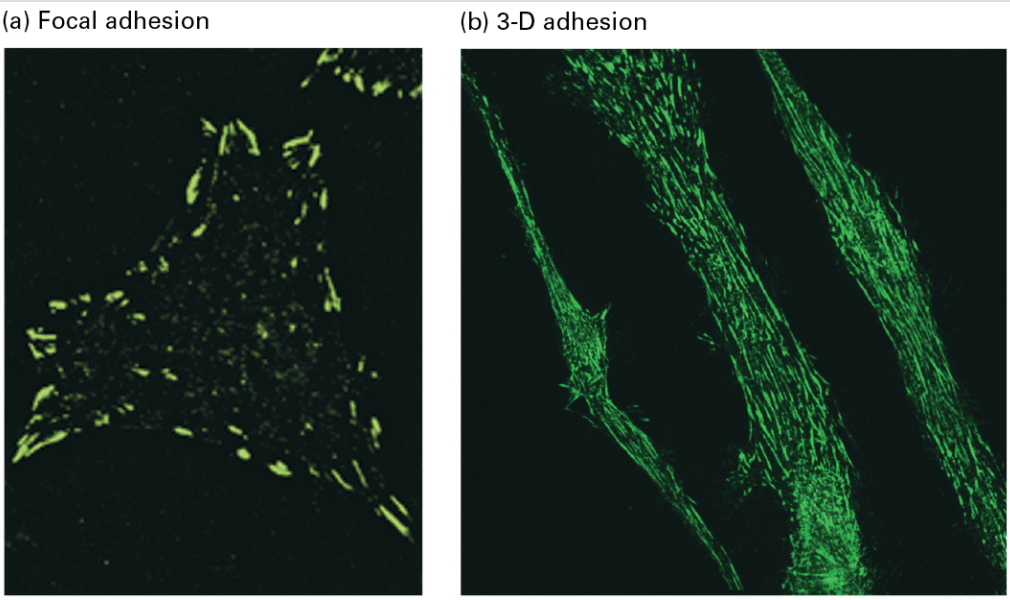
20.5 Adhesive Interactions in Motile and Nonmotile Cells

After adhesive interactions in epithelia form during differentiation, they are often very stable and can last throughout the life span of the cells or until the epithelium undergoes further differentiation. Although such long-lasting nonmotile adhesion also exists in nonepithelial tissues, some nonepithelial cells must be able to crawl across or through a layer of ECM or other cells. Moreover, during development or wound healing and in certain pathological states (e.g., cancer), epithelial cells can transform into more motile cells (the epithelial-to-mesenchymal transition). Changes in expression of adhesion molecules play a key role in this transformation, as they do in other biological processes involving cell movement, such as the crawling of white blood cells into tissue sites of infection. In this section, we describe various cell-surface structures that mediate transient adhesive interactions that are especially adapted for the movement of cells as well as those that mediate long-lasting adhesion. The intracellular mechanisms used to generate the mechanical forces that propel cells and modify their shapes are covered in Chapters 17 and 18.

Integrins Mediate Adhesion and Relay Signals Between Cells and Their Three-

Dimensional Environment

As already discussed, integrins can connect epithelial cells to the basal lamina and, through adapter proteins, to intermediate filaments of the cytoskeleton (see Figure 20-1). In both nonepithelial and epithelial cells, integrins form a bridge between the ECM and the cytoskeleton. In epithelial and nonepithelial cells, integrins in the plasma membrane are also clustered with other molecules in various focal contacts (focal adhesions) and focal contact–like adhesive structures called *focal complexes*, *3-D adhesions*, and *fibrillar adhesions*, as well as in circular adhesions called *podosomes*. These structures are multiprotein complexes that mediate (1) cell adhesion to the ECM — for example, via integrin binding to fibronectin (see Figure 20-36), collagen or laminin, (2) integrin association with the actin cytoskeleton, (3) adhesion-dependent outside-in and inside-out signaling (see Figure 20-8), and (4) mechanosensory coupling between cells and their environments. These complexes are readily observed by fluorescence microscopy with the use of antibodies that recognize integrins or other molecules clustered with them (Figure 20-38).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 20-38 Integrins cluster into adhesive structures with various morphologies in nonepithelial cells. Immunofluorescence methods were used to detect integrin-containing adhesive structures (green) on cultured cells: (a) focal adhesions and (b) 3-D adhesions on the surfaces of human fibroblasts. Cells were grown (a) directly on the flat surface of a culture dish or (b) on a three-dimensional matrix of ECM components. The shape, distribution, and composition of the integrin-based adhesions formed by cells vary depending on the cells' environment.

[Part (a) republished with permission from Nature, from B. Geiger et al., 2001, "Transmembrane Crosstalk Between the Extracellular Matrix and the Cytoskeleton," *Nat. Rev. Mol. Cell Biol*. **2**(11):793–805; permission conveyed through Copyright Clearance Center, Inc. Part (b) Kenneth Yamada and Edna Cukierman.]

Integrin-containing adhesive structures are dynamic due to ongoing import, export, or covalent modification of their components, and each contains dozens of intracellular adapter and associated proteins. These adhesive structures may in some cases assemble due to the formation of phase-separated biomolecular condensates

(see Chapter 3) in which each component has the potential of binding directly to many other components. The hundreds of such proteins identified to date have the potential to engage in many hundreds of distinct protein-protein interactions that may be subject to regulation. For example, binding sites generated by phosphorylation of integrin and its associated proteins, as well as by generation of phosphorylated derivatives of phosphatidylinositol in the adjacent membrane, recruit additional proteins into and can also cause release of some proteins from these multiprotein complexes. A tightly controlled choreography of internal signals, contributions of other signaling pathways such as those involving receptor tyrosine kinases (see Figure 20-8), and external signals (such as the composition and rigidity of the ECM) regulates these complexes. Together, they help define the precise composition and activity of the integrin multiprotein complex and the consequent influence that it has on cellular structure and activity (outside-in effect) as well as the influence of the cellular actin cytoskeleton on the ECM (inside-out effect). The adhesive structures can be stable and substantially contribute to the architecture of a tissue. They can also be transient. For example, the dynamic formation and dissolution of adhesive structures can result in cells grabbing onto the ECM for traction, exerting force to move relative to the adhesion site, and then disassembling to release the cell, permitting net movement as cells crawl through adjacent tissue.

Although found in many nonepithelial cells, integrin-containing adhesive structures have been studied most frequently in fibroblasts

grown in cell culture on flat glass or plastic surfaces called *substrata*. These conditions only poorly approximate the three-dimensional ECM environment that normally surrounds such cells in vivo. When fibroblasts are cultured in three-dimensional ECMs derived from cells or tissues, they form adhesions to the three-dimensional ECM substratum, called *3-D adhesions*. These structures differ somewhat in composition, shape, distribution, and activity from the focal or fibrillar adhesions seen in cells growing on the flat substrata typically used in cell culture experiments (see Figure 20-38). Cultured fibroblasts in three-dimensional ECMs have anchoring junctions that display greater adhesion and mobility, increased rates of cell proliferation, and spindle-shaped morphologies more like those of fibroblasts in tissues than do cells cultured on hard, flat surfaces. These and other observations indicate that the topological, compositional, and mechanical properties of the ECM all play a role in controlling the shape and activity of a cell, due at least in part to integrin-mediated bridging of the ECM to intracellular molecules. Tissue-specific differences in these ECM characteristics probably contribute to the tissue-specific properties of cells.

The importance of the three-dimensional environment of cells has been highlighted by cell-culture studies of the morphogenesis, functioning, and stability of specialized milk-producing mammary epithelial cells and their cancerous transformed counterparts. For example, the three-dimensional ECM-dependent outside-in signaling mediated by integrins influences the epidermal growth

factor–tyrosine kinase receptor signaling system, and vice versa. The three- dimensional ECM also permits the mammary epithelial cells to generate in vivo–like circular epithelial structures, called *acini*, which secrete the major protein constituents of milk. The use of such three-dimensional ECM cell culture systems permits more realistic comparisons of the responses of normal and cancer cells to potential chemotherapeutic agents. Analogous systems employing both natural and synthetic three-dimensional ECMs are being developed to provide more in vivo–like conditions to study other complex tissues and organs, such as the liver. In the presence of the appropriate ECM, many types of 3-D, organ-like tissues called *organoids*, similar to those in intact organisms, can now be generated from self-organizing, pluripotent or adult stem cells (see Chapters 4 and 22).

Regulation of Integrin-Mediated Adhesion and Signaling Controls Cell Function and Movement

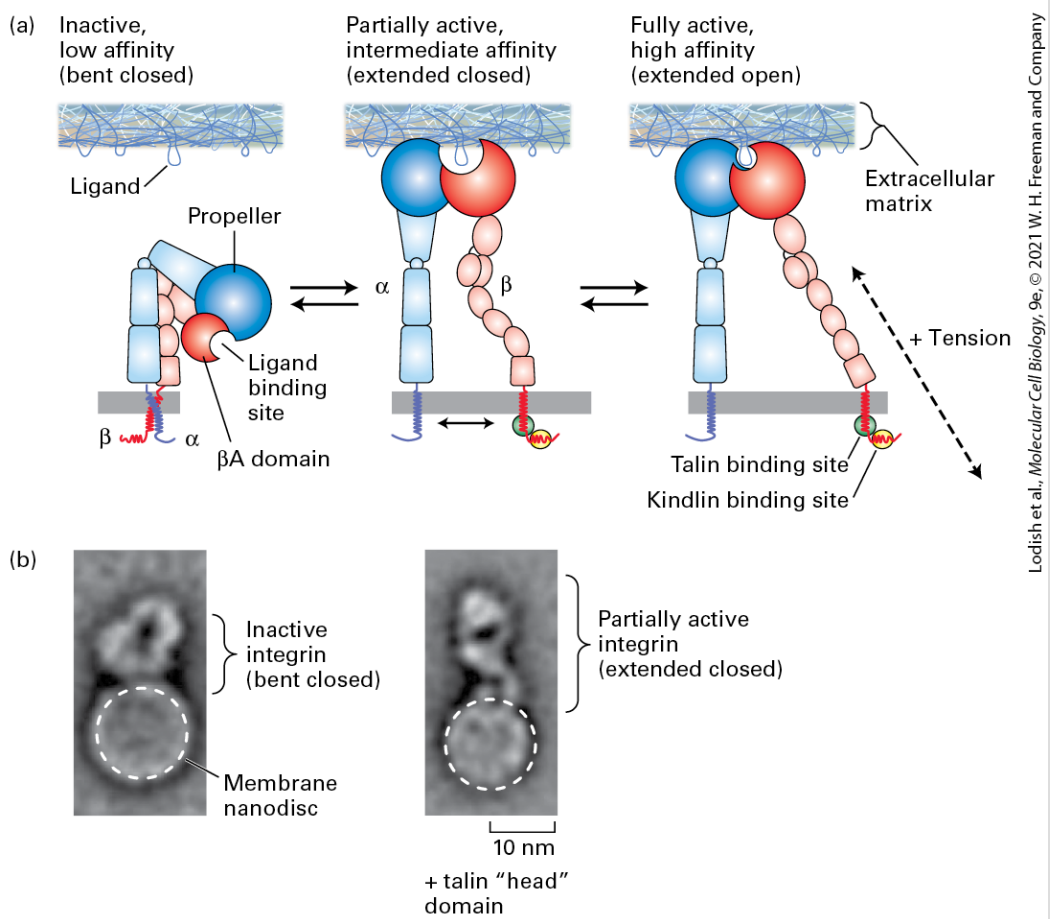
Cells can exquisitely control the strength of integrin-mediated cell-matrix interactions by regulating integrin's expression levels, ligand-binding activities, or both. Such regulation is critical to the role of these interactions in cell migration and other functions involving cell movement.

Integrin Binding

Integrin binding to ECM and neighboring cells can influence the developmental fate of cells, determining the type of cell into which a progenitor cell will differentiate (see Chapter 22). Integrins participate in cells sensing the mechanical characteristics of their environments, such as the rigidity of the ECM, and responding appropriately.

Most integrins can exist in at least three conformations. In the first, inactive state (*bent closed* conformation, Figure 20-39a, *left*), the closely apposed extracellular domains of the αβ heterodimer are bent over, placing the ligand-binding site — located at the interface of the propeller (α chain) and βA (β chain) domains — near the plasma membrane and distant from the ECM. In the bent closed conformation, there is only low-affinity ligand binding and the transmembrane domains and cytoplasmic C-terminal tails of the two subunits are closely bound together. In the second, partially active state, the transmembrane and cytoplasmic domains of the α and β chains at least partially separate and the extracellular domains partially straighten into an *extended closed* conformation in which the ligand-binding site is projected farther away from the cell surface and assumes an intermediate affinity for ligands, such as a variety of ECM proteins (Figure 20-39a, *center*). In the third, fully active, *extended open* conformation, tension on the protein fully extends the extracellular domain of the β chain and confers high affinity ligand binding that can be as much as 4000-fold tighter than

binding to the bent closed conformation (Figure 20-39a, *right*). The changes in the conformations of integrin's cytoplasmic tails can influence the binding of intracellular adaptor proteins. For example, two of integrin's cytoplasmic adaptor proteins, talin and kindlin, can bind to short sequence motifs in the cytoplasmic tail of the β chain when it is separated from the tail of the α chain in the extended closed and open forms, but not when these sites are blocked in the bent closed form (see Figure 20-39a).



EXPERIMENTAL FIGURE 20-39 The three conformations of integrins. (a) Most integrins can be in one of three interconvertible conformations: inactive, low affinity for extracellular ligands (bent closed, *left*); partially active, intermediate

affinity (extended closed, *center*); and fully active, high affinity (extended open, *right*). Tension across the integrin due to pulling by either its extracellular ligand or intracellular binding partners drives the integrin into the extended open conformation. The extracellular ligand binding site is at the interface of the propeller domain on the α chain and the βA domain on the β chain. The cytoplasmic domain of the β chain has binding sites for intracellular adaptor proteins, e.g., talin and kindlin (green and yellow ovals, respectively). These sites are inaccessible (cryptic) in the bent closed form, but accessible in the extended forms due to separation of the transmembrane and cytoplasmic domains of the α and β chains (indicated by double-headed arrow). (b) Single inactive (bent closed) integrin αIIbβ3 molecules (*left*) were incorporated into phospholipid nanodiscs (small bilayers in which the extracellular and cytoplasmic domains of the integrin are exposed to a buffer), and the integrin-binding and activating head domain of the adapter protein talin was added to some of these preparations (*right*). Multiple electron microscopic images of individual nanodiscs were collected and averaged. Phospholipid nanodiscs are indicated by dashed white circles, and the heights of the integrin extracellular regions that extend above the nanodiscs are indicated by brackets. Binding of talin's head domain to the cytoplasmic domain of the β chain traps the integrins in the extended closed conformation.

[Part (a) Information from J. Wang et al., 2019, *Nat. Commun.* **10**:5481. Part (b) F. Ye et al., 2010, *J. Cell Biol.* **188**(1):157–173, Fig. 7; https://doi.org/10:1083/jcb.200908045.]

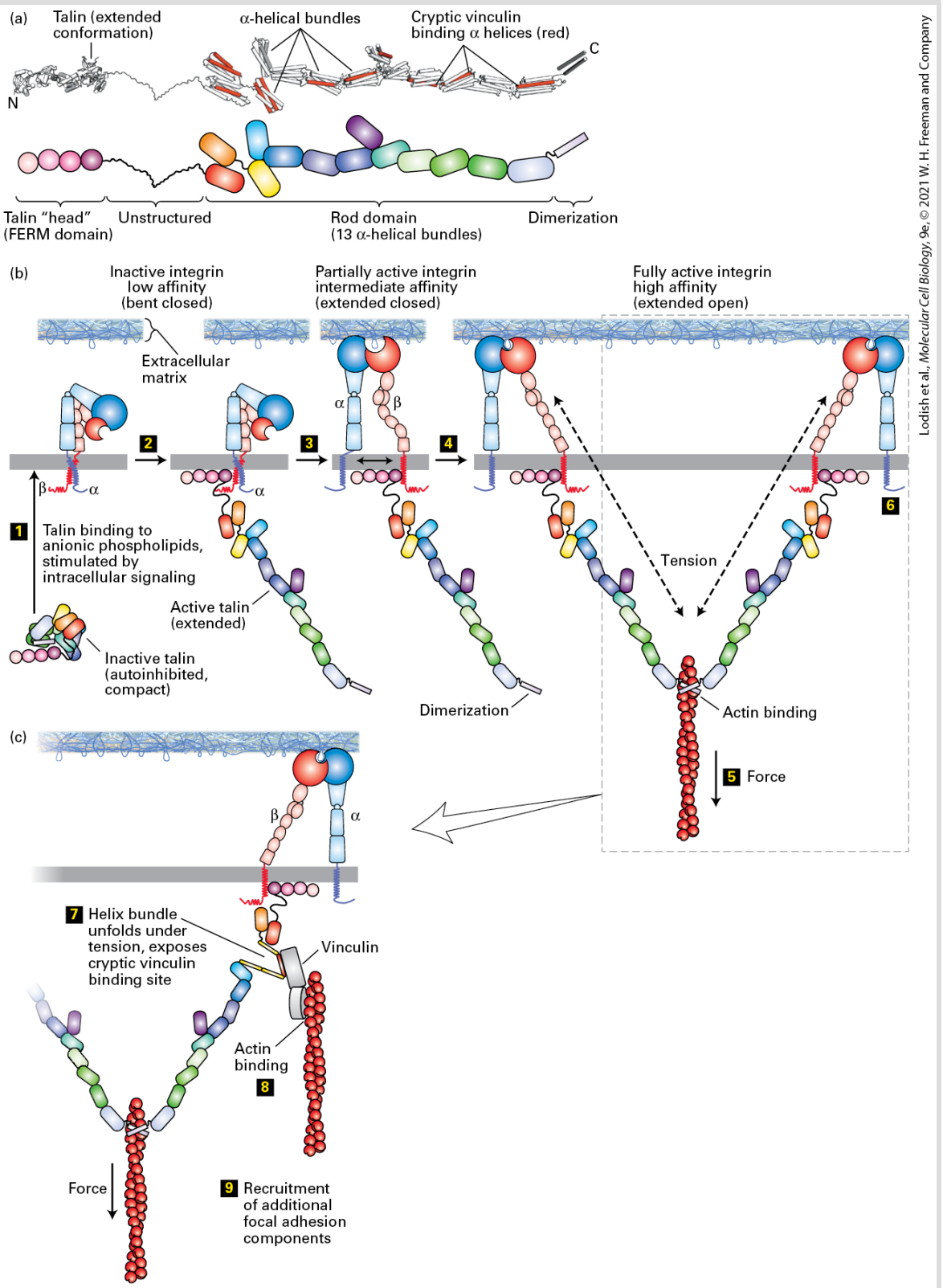
One helpful way to think about these three forms of integrins is that there is an equilibrium distribution of integrin molecules in all three conformations (they can readily interconvert). When the integrin is bound to neither extracellular ligands (e.g., ECM) nor intracellular partners (e.g., talin), the large majority of unbound integrin molecules is in the bent, inactive conformation. Binding of extracellular ligands to the extended conformations traps more of

the integrins in the extended conformations than in the absence of the ligands (shifts the equilibrium distribution) and thus facilitates the binding of intracellular adaptor proteins such as talin or kindlin to the now accessible sites in the β chain's cytoplasmic domain. Similarly, binding of intracellular adaptors to the extended conformations increases the fraction of integrin molecules in extended conformations (again, a shift in the equilibrium distribution) and facilitates extracellular ligand binding to the intermediate and high-affinity ligand-binding sites. Thus binding of ligands to integrins on either side of the membrane can increase binding on the other side.

The structural features of integrin's α and β subunits that permit them to adopt the bent and extended conformations have been preserved through metazoan evolution, indicating their importance for integrin function. Indeed, these three conformations provide an attractive explanation for the ability of integrins to mediate outside-in and inside-out signaling. The binding to the integrin's extracellular ligand-binding site of certain ECM molecules or CAMs on other cells would shift the distribution of conformations and thus result in more integrins in the extended closed or extended open forms with at least partially separated cytoplasmic tails. Intracellular adapter proteins could recognize the separation of the tails and, as a result, bind to the tails. Integrin-binding-dependent changes in these adapters can then alter the cytoskeleton and activate or inhibit intracellular signaling pathways. Conversely, changes in the metabolic or signaling state of the cells could cause

intracellular adapters to bind to or dissociate from the cytoplasmic tails of the integrins and thus alter the distribution of integrin conformations. Such binding can force the tails either to separate or to associate (see Figure 20-39a). As a consequence, a greater fraction of the integrins would be either in the bent closed (inactive), extended closed (partially active), or extended open (fully active) conformations, thereby altering their interactions with the ECM or with other cells. Indeed, in vitro EM studies of purified integrins reconstituted individually into lipid bilayer nanodiscs show that binding of the head domain of the adapter/mechanosensor protein talin (described in detail below; see Figure 20-40) to the cytoplasmic tail of integrin's β chain is sufficient to partially activate integrin, inducing a straightening of the bent closed conformation into the extended closed form (see Figure 20-39b).

We can see an example of integrins mediating intracellular-extracellular communication in the initial formation of focal adhesions in Figure 20-40. Although there are many intracellular proteins involved, this figure focuses on talin, actin, and vinculin. Figure 20-40a shows a model of an active conformation of talin and Figures 20-40b and c show some of the key steps associated with integrin/talin-mediated inside-out signaling. In the absence of stimulation, most of the talin in the cytoplasm is condensed into an inactive, autoinhibited, compact structure in which interdomain binding sterically blocks talin binding to some of its many partners, including integrins and actin (see Figure 20-40b, *lower left*).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-40 Talin- and integrin-mediated intracellular-extracellular communication. (a) *Top:* A hypothetical model of an elongated, active conformation of a talin monomer based on structures of fragments of the intact protein with secondary structures illustrated as tubes (α helices) and ribbons with arrowheads (β sheets). The N-terminal "head" domain (four globular subdomains) is connected by an apparently unstructured domain to a rod domain that comprises 13 α-helical bundles (4 or 5 helices per bundle), 9 of which contain either 1 or 2 helices (red, total of 11) that are cryptic binding sites for the protein vinculin. Some bundles also contain cryptic actin-binding sites. These helices can only bind vinculin when exposed by unfolding (see Figure 20-9b). There is a single α helix at the extreme C-terminus that is a dimerization domain. *Bottom:* A cartoon model. Head subdomains = gray ovals; rod domain bundles = colored capsules; and dimerization domain = narrow rectangle. (b) Intracellular signaling induces the translocation of inactive talin in the cytoplasm to the plasma membrane and its unfolding/activation (steps **1** and **2**). The binding of activated talin to integrin's β chain cytoplasmic domain stabilizes integrin's extended closed conformation, promoting binding to the ECM (step **3**). The unfurled talin homodimerizes and binds several proteins, including actin filaments that bind via talin's C-terminal dimerization domain (shown) and several other sites (not shown) (step **4**). The actin filaments exert force (step **5**), generating tension across the ECM/integrin/talin complex. This tension converts the integrin into the fully active, extended open form (step **6**). (c) Subsequent steps illustrated for the region outlined in the light gray dashed box in part (b) are shown. Additional force unfolds one or more of talin's helical bundles (step **7**), exposing cryptic binding site(s) (red helix) that recruit vinculin and other proteins, including more actin (steps **8** and **9**), resulting in growth of the focal adhesion.

[Part (a) *left* information from B. T. Goult, J. Yan, and M. A. Schwartz, 2018, *J. Cell Biol*. **217**(11):3776–3784; and M. Yao et al., 2014, *Sci. Rep.* **4**:4610. Parts (b) and (c) information from J. Wang et al., 2019, *Nat. Commun*. **10**:5481; D. Dedden et al., 2019, *Cell* **179**(1):120–131; S. Chakraborty et al., 2019, *Biochemistry* **58**(47):4677–4695; Y. A. Kadry and D. A. Calderwood, 2020, *Biochim. Biophys. Acta, Biomembr.* **1862**(5):183206.]

Intracellular signaling can induce the translocation of talin to the inner leaflet of the plasma membrane, for example, by generating polyphosphoinositides to which globular subdomains in talin's head

domain can bind and by small GTP-binding proteins (Figure 20-40b, steps **1** and **2**). In steps **2** and **3**, the compact autoinhibited rod domain unfolds into the active conformation (see Figure 20-40a) and a portion of the head domain binds to the cytoplasmic domain of integrin's β chain. Talin binds to integrin's β chain cytoplasmic domain when the transmembrane and cytoplasmic domains of integrin's α and β chains at least partly separate in the extended closed conformation. In this conformation integrin binds to the ECM with intermediate affinity (see Figure 20-39). The order of these events is not well established. The C-terminal dimerization domain in the unfurled, extended, and active talin mediates homodimerization, and actin filaments can bind to several sites in the rod domains in the dimeric talin, including the C-terminal dimerization domain (step **4**, also see Figure 20-40a). Other binding partners can also bind to the unfurled, active talin. Additional signaling steps mediated by the partially active, extended closed integrin (e.g., activation of kinases, such as focal adhesion kinase and scr-like kinases) induce formin-mediated actin filament formation and myosin engagement with actin filaments to exert force on the talin (step **5**, also see Chapter 17). As a consequence, tension develops in the ECM-bound integrin/talin complex, converting the integrin into the fully active, extended open form (step **6**).

Additional force applied to the integrin/talin complex, exerted either from the cytoplasm (e.g., actin filaments and myosin) or the ECM, can pull apart the α-helical bundles in talin (Figure 20-40c, step **7**),

exposing cryptic vinculin binding sites (red helix). The greater the force, the more α-helical bundles in talin are opened and the more vinculin, actin, and other proteins are recruited to the integrin/talin complex. Vinculin binding to exposed sites on talin can then recruit other proteins, such as actin filaments (steps **8** and **9**). The unfolded bundles can readily refold when the force is reduced. Strikingly, the more rigid the structure of the ECM, the greater the tension across the integrin/talin complex and the larger and stronger the adhesion. During steps **3**–**9**, integrins begin to cluster and the focal adhesion can continue to grow and form a mature focal adhesion that firmly attaches the cell to the ECM. Alternatively, the multiprotein adhesion complex can dissociate with inactivation of the integrins and other components, in part due to recruitment of inhibitory proteins.

Studies suggest that the efficient activation of integrins in intact cells may also require the participation of another class of adapter proteins called *kindlins*, which bind to a distinct site on the cytoplasmic tail of integrin's β chain (see Figure 20-39a). Kindlin plays a key role in the integrin- and microfibril-mediated activation of TGF-β (inside-out signaling involving elastic fibers and their microfibril-associated protein LTBP, described earlier) and other pathways of integrin-mediated signaling, such as recruitment of branched actin networks in lamellipodia (see Chapter 17).

Platelet function, discussed in more detail below, provides a good example of how cell-matrix interactions are modulated by control of integrin-binding activity. Platelets are cell fragments that circulate in the blood and clump together with ECM molecules to form a blood clot. In its basal state, the αIIbβ3 integrin present on the plasma membranes of platelets cannot bind tightly to its protein ligands (including fibrinogen and fibronectin), all of which participate in the formation of a blood clot, because it is in the inactive (bent closed) conformation. During clot formation, platelets are activated by binding to ECM proteins such as collagen and a large protein called *von Willabrand factor* that, through binding to receptors, generate intracellular signals. Platelets may also be activated by ADP or the clotting enzyme thrombin. These signals induce changes in signaling pathways within the platelet that result in an activating conformational change in the platelet's αIIbβ3 integrin. As a consequence, this integrin can bind tightly to extracellular clotting proteins and participate in clot formation. People with genetic defects in the β3 integrin subunit are prone to excessive bleeding, attesting to the role of the αIIbβ3 integrin in the formation of blood clots (see Table 20-4).

Integrin Expression

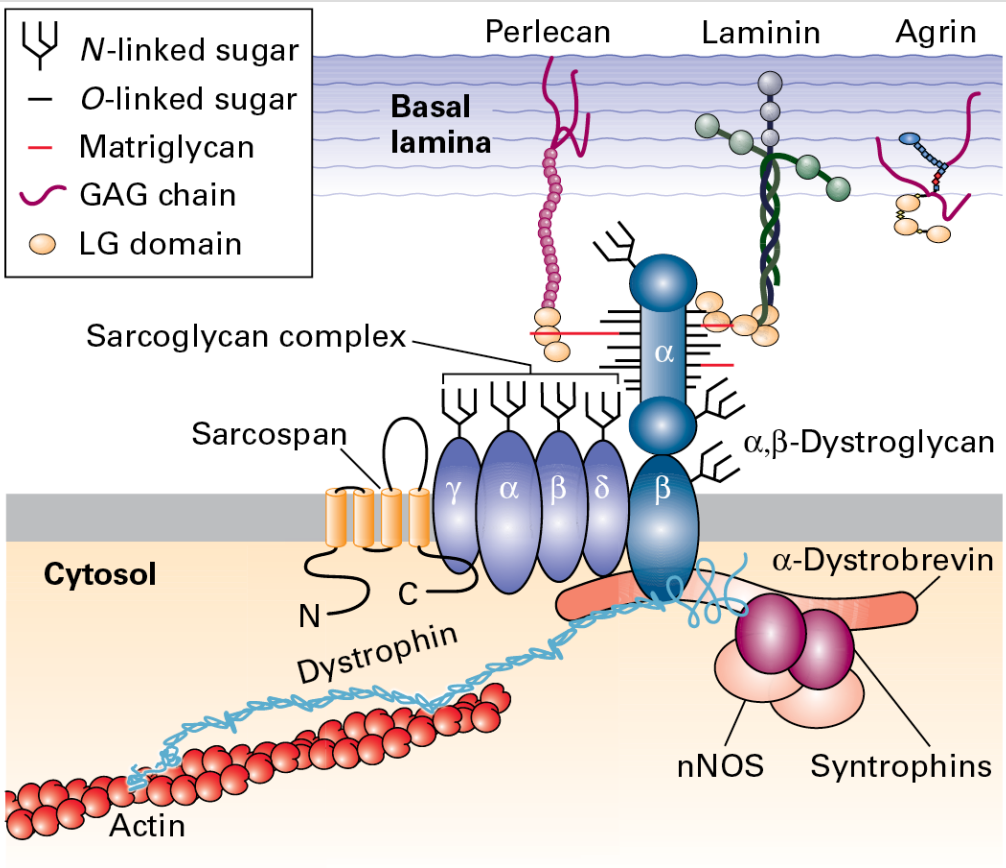
The attachment of cells to ECM components can also be modulated by altering the number of integrin molecules exposed on the cell surface. The α4β1 integrin, which is found on many hematopoietic

cells, offers an example of this regulatory mechanism. For these hematopoietic cells to proliferate and differentiate, they must be attached to fibronectin synthesized by supportive (stromal) cells in the bone marrow. The α4β1 integrin on hematopoietic cells binds to a Glu-Ile-Leu-Asp-Val (EILDV) sequence in fibronectin in the ECM, thereby anchoring the cells to the matrix. This integrin also binds to a sequence in a CAM called *vascular CAM-1 (VCAM-1)*, which is present on stromal cells of the bone marrow. Thus hematopoietic cells directly contact the stromal cells as well as the ECM. Late in their differentiation, hematopoietic cells decrease their expression of this integrin; the resulting reduction in the number of α4β1 integrin molecules on the cell surface is thought to allow mature blood cells to detach from the ECM and stromal cells in the bone marrow and enter the circulation.

Connections Between the ECM and Cytoskeleton Are Defective in Muscular Dystrophy

The importance of the adhesion receptor–mediated linkage between ECM components and the cytoskeleton is highlighted by a set of hereditary muscle-wasting diseases, collectively called *muscular dystrophies*. Duchenne muscular dystrophy (DMD), the most common type, is a sex-linked disorder, affecting 1 in 3300 boys, that results in cardiac or respiratory failure, usually in the late teens

or early twenties. The first clue to understanding the molecular basis of this disease came from the discovery that people with DMD carry mutations in the gene encoding a protein named *dystrophin*. This very large protein was found to be a cytosolic adapter protein that contains a long central rodlike region with 24 repeats of a three-α-helix bundle, flanked by an N-terminal F-actin-binding domain and C-terminal domains that bind to several plasma membrane and cytoplasmic proteins (Figure 20-41). Several of the central repeats also bind to F-actin and dystrophin links the actin filaments to an adhesion receptor called *dystroglycan* (see Figure 20-41). It is noteworthy that when dystrophin is subjected to tension some of its α-helical repeats unfold and the protein substantially lengthens, a feature somewhat reminiscent of the force-induced unfolding of talin (see Figure 20-40). However, a role for dystrophin in mechanotransduction has not been established.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-41 The dystrophin glycoprotein complex (DGC) in skeletal muscle cells. This schematic model shows that the DGC comprises three subcomplexes: the α, β-Dystroglycan subcomplex; the integral membrane protein sarcospan and the sarcoglycan complex; and the cytosolic adapter subcomplex comprising dystrophin, other adapter proteins, and signaling molecules. Through its *O*-linked matriglycan sugars (see Figure 20-31c), α-dystroglycan binds to LG-domain-containing components of the basal lamina, such as laminin and the GAG-containing proteoglycans perlecan and agrin. Dystrophin — the protein that is defective in DMD — links β-dystroglycan to the actin cytoskeleton, and α-dystrobrevin links dystrophin to the sarcoglycan complex and sarcospan. Neuronal nitric oxide synthase (nNOS), which produces the gaseous signaling molecule nitric oxide, binds to dystrophin via the adaptors called *syntrophins*. See S. J. Winder, 2001, *Trends Biochem. Sci.* **26**:118–124; D. E. Michele and K. P. Campbell, 2003, *J. Biol. Chem.* **278**(18):15457–15460; T. Yoshida-Moriguchi and K. P. Campbell, 2015,

Glycobiology **25**(7):702–713; and Y. M. Kobayashi and K. P. Campbell, 2012, Chapter 66, "Skeletal Muscle Dystrophin-Glycoprotein Complex and Muscular Dystrophy," in J. A. Hill and E. N. Olson, eds., *Muscle*, vol. 2, Elsevier/Academic Press.

Dystroglycan is a heterodimeric (α and β subunits) cell-surface glycoprotein (see Figure 20-41). As noted in Section 20.4, about three of the more than 20 *O*-linked oligosaccharide chains in dystroglycan's α subunit are directly linked to the hydroxyl group via a mannose sugar that is phosphorylated (see Figure 20-31c). A phosphorylated hexasaccharide connects the phosphorylated mannose to a GAG-like polymer of xylose–glucuronic acid disaccharides called **matriglycan**. One key enzyme in the synthesis of the matriglycan oligosaccharide that is located in the Golgi complex is called *LARGE*.

The *O*-linked matriglycan binds to various components of the basal lamina, including the multi-adhesive matrix protein laminin (see Figures 20-25 and 20-41) and the proteoglycans perlecan and agrin. The neurexins, a family of cell-surface adhesion molecules expressed by neurons, also bind matriglycan. All of these bind to matriglycan via LG domains (see Figure 20-41).

Dystroglycan's transmembrane β subunit associates with a group of integral membrane proteins called the *sarcoglycan complex* and *sarcospan*, while its cytosolic domain binds directly to dystrophin and directly or indirectly to other adapter and signaling proteins

(see Figure 20-41). Sarcospan is a tetraspanin, as are claudin and occludin in tight junctions (see Figure 20-18b). The resulting large, heteromeric assemblage, the *dystrophin glycoprotein complex (DGC)*, links the ECM to the actin cytoskeleton. Binding to laminin in the basal lamina leads to a cluster of DGCs on the cell surface. The DGC is associated with signaling pathways within muscle and other types of cells. For instance, the signaling enzyme neuronal nitric oxide synthase (nNOS) is associated through syntrophin with the DGC in skeletal muscle. The rise in intracellular Ca^{2+} during muscle contraction activates nNOS to produce nitric oxide (NO), a signaling molecule that diffuses into smooth muscle cells surrounding nearby blood vessels. NO promotes smooth muscle relaxation, leading to a local rise in the flow of blood supplying nutrients and oxygen to the active skeletal muscle. Heart (cardiac) muscle contraction may be influenced by similar NOS-syntrophin interactions.

Mutations in the genes encoding dystrophin, other DGC components, laminin, or the multiple enzymes that mediate the addition of matriglycan to dystroglycan can all disrupt the DGC-mediated link between the exterior and the interior of muscle cells and cause muscular dystrophies. In addition, dystroglycan mutations have been shown to greatly reduce the clustering of acetylcholine receptors on muscle cells at the neuromuscular junctions, which is also dependent on the basal lamina proteins laminin and agrin. These and possibly other effects of DGC defects apparently lead to a cumulative loss of the mechanical stability of

muscle cells as they undergo contraction and relaxation, resulting in deterioration of the cells and muscular dystrophy.

Dystroglycan is an elegant and medically relevant example of the intricate networks of connectivity in cell biology. While originally discovered in the context of studying muscular dystrophy, it was later shown to be expressed in nonmuscle cells and, through its binding to laminin, to play a key role in the assembly and stability of at least some basement membranes. The extent of glycosylation of dystroglycan, such as the length of the matriglycan repeats, varies in different tissues, suggesting that alterations in its glycosylation regulate function. Dystroglycan is essential for normal development, including the proper functioning of some neuronal synapses (see Chapter 23). Additional studies led to its identification as a cell-surface receptor for the virus that causes the frequently fatal human disease Lassa fever and other related viruses, all of which bind via matriglycan. Furthermore, dystroglycan is the receptor on specialized cells in the nervous system — Schwann cells — to which binds the pathogenic bacterium *Mycobacterium leprae*, the organism that causes leprosy. ■

IgCAMs Mediate Cell-Cell Adhesion in Neural and Other Tissues

Numerous transmembrane proteins characterized by the presence of multiple immunoglobulin domains in their extracellular regions constitute the immunoglobulin (Ig) superfamily of CAMs, or *IgCAMs* (e.g., see NCAM in Figure 20-2). The Ig domain is a common protein domain, containing 70–110 residues. It was first identified in antibodies, the antigen-binding immunoglobulins (see Chapter 24), but has a much older evolutionary origin in CAMs. The human, *Drosophila*, and *C. elegans* genomes include about 765, 150, and 64 genes, respectively, that encode proteins containing Ig domains. Immunoglobulin domains are found in a wide variety of cell-surface proteins, including the T-cell receptors produced by lymphocytes and many proteins that take part in adhesive interactions. Among the IgCAMs are neural CAMs; intercellular CAMs (ICAMs), which function in the movement of leukocytes into tissues; and junction adhesion molecules (JAMs), which are present in tight junctions (see Figure 20-18b).

As their name implies, neural CAMs are of particular importance in neural tissues. One type, the NCAMs, primarily mediate homophilic interactions. First expressed during morphogenesis, NCAMs play an important role in the differentiation of muscle cells, glial cells, and neurons. Their role in cell adhesion has been directly demonstrated by the inhibition of adhesion with anti-NCAM antibodies. Numerous NCAM isoforms, encoded by a single gene, are generated by alternative mRNA splicing and by differences in glycosylation. Other neural CAMs (e.g., L1-CAM) are encoded by different genes. In

humans, mutations in different parts of the L1-CAM gene cause various neuropathologies (e.g., mental retardation, congenital hydrocephalus, and spasticity).

An NCAM comprises an extracellular region with five Ig domains and two fibronectin type III domains, a single membrane-spanning segment, and a cytosolic segment that interacts with the cytoskeleton (see Figure 20-2). In contrast, the extracellular region of L1-CAM has six Ig domains and four fibronectin type III domains. As with cadherins, cis (on the same cell) interactions and trans (intercellular) interactions probably play key roles in IgCAM-mediated adhesion (see Figure 20-3); however, adhesion mediated by IgCAMs is Ca^{2+} independent.

Leukocyte Movement into Tissues Is Orchestrated by a Precisely Timed Sequence of Adhesive Interactions

In adult organisms, several types of white blood cells (leukocytes) participate in defense against infection caused by bacteria and viruses and respond to tissue damage due to trauma or inflammation. To fight infection and clear away damaged tissue, these cells must move rapidly from the blood, where they circulate as unattached, relatively quiescent cells, into the underlying tissue at sites of infection, inflammation, or damage. We know a great deal

about the movement into tissue, termed *extravasation*, of four types of leukocytes: neutrophils, which release several antibacterial proteins; monocytes, the precursors of macrophages, which can engulf and destroy foreign particles; and T and B lymphocytes, the antigen-recognizing cells of the immune system (see Chapter 24).

Extravasation requires the successive formation and breakage of cell-cell contacts between leukocytes in the blood and endothelial cells lining the vessels. Some of these contacts are mediated by **selectins**, a family of CAMs that mediate leukocyte–vascular endothelium interactions. Endothelial cells express *P-* and *E-selectins* on their blood-facing surfaces, activated platelets express P-selectin, and leukocytes express *L-selectin*. All selectins contain a Ca^{2+}-dependent *lectin domain*, which is located at the distal end of the extracellular region of the molecule and recognizes particular sugars in glycoproteins or glycolipids (see Figure 20-2). For example, the primary ligand for P- and E-selectins is an oligosaccharide called the *sialyl Lewis-x antigen*, a part of longer oligosaccharides present in abundance on leukocyte glycoproteins and glycolipids.

Figure 20-42 illustrates the basic sequence of cell-cell interactions leading to the extravasation of leukocytes. Various inflammatory signals released in areas of infection or inflammation first cause activation of the vascular endothelium. P-selectin exposed on the surfaces of activated endothelial cells mediates the weak adhesion of passing leukocytes. Because of the force of the blood flow and the

rapid “on” and “off” rates of P-selectin binding to its ligands, these bound leukocytes are slowed, but not stopped, and literally roll along the surface of the endothelium. Among the signals that promote activation of the endothelium are chemokines, a group of small, secreted proteins (8–12 kDa) produced by a wide variety of cells, including endothelial cells and leukocytes.

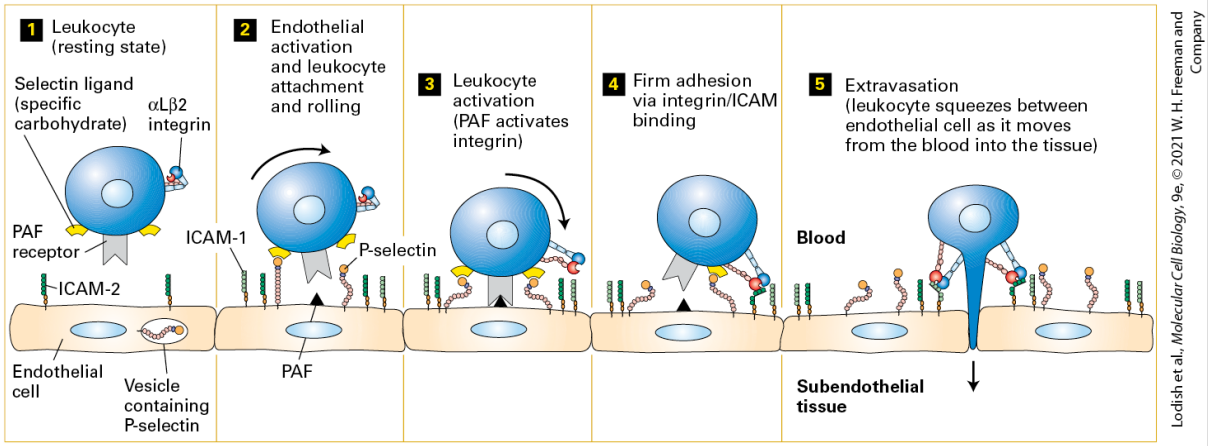


FIGURE 20-42 Endothelium-leukocyte interactions: activation, binding, rolling, and extravasation. Step 1: In the absence of inflammation or infection, leukocytes and endothelial cells lining blood vessels are in a resting state and not interacting. Step 2: Inflammatory signals released only in areas of inflammation, infection, or both activate resting endothelial cells, resulting in the movement of vesicle-sequestered selectins to the cell surface. The exposed selectins mediate weak binding of leukocytes by interacting with carbohydrate ligands on leukocytes. Blood flow forces the loosely bound leukocytes to roll along the endothelial surface of the blood vessel (curved arrow). Activation of the endothelium also causes synthesis of platelet-activating factor (PAF) and ICAM-1, both expressed on the endothelial cell surface. PAF and other, usually secreted, activators, including chemokines, then induce changes in the shapes of the leukocytes and activation of leukocyte integrins such as αLβ2, which is expressed by T lymphocytes (step 3). The subsequent tight binding between activated integrins on leukocytes and CAMs on the endothelium (e.g., ICAM-2 and ICAM-1) results in firm adhesion (step 4) and subsequent movement (extravasation) into the underlying tissue (step 5). See R. O. Hynes and A. D. Lander, 1992, *Cell* **68**:303.

For tight adhesion to occur between activated endothelial cells and leukocytes, β2-containing integrins on the surfaces of the leukocytes must be activated indirectly by chemokines or by other local activation signals such as *platelet-activating factor (PAF)*. Platelet-activating factor is unusual in that it is a phospholipid rather than a protein; it is exposed on the surfaces of activated endothelial cells at the same time that P-selectin is exposed. The binding of PAF or other activators to their G protein–coupled receptors on leukocytes leads to activation of the leukocyte integrins (see Figure 20-39). These activated integrins then bind to distinct IgCAMs on the surfaces of endothelial cells. These IgCAMs include ICAM-2, which is expressed constitutively, and ICAM-1, whose synthesis is induced by activation. ICAM-1 does not usually contribute substantially to leukocyte adhesion to endothelial cells immediately after activation, but rather participates at later times in cases of chronic inflammation. The tight adhesion mediated by these Ca^{2+}-independent integrin-ICAM interactions leads to the cessation of rolling and to the spreading of leukocytes on the surface of the endothelium; soon the adhered cells move between adjacent endothelial cells and into the underlying tissue. The extravasation step itself (also called *transmigration* or *diapedesis*; step **5** in Figure 20-42) requires the dissociation of otherwise stable adhesive interactions between endothelial cells that are primarily mediated by the CAM VE-cadherin. There is general agreement that the leukocyte interactions with endothelial cells mediated by CAMs

initiate outside-in signaling in the endothelial cells that involves phosphorylation, activation of small GTPases, and an increase in cytosolic calcium concentration. These signals weaken or disrupt VE-cadherin-mediated inter-endothelial-cell adherens junctions and increase actin-myosin contraction, which pulls the endothelial cells apart, thus permitting the paracellular, amoeboid movement of the leukocyte between adjacent endothelial cells that is responsible for most extravasation.

The selective adhesion of leukocytes to the endothelium near sites of infection or inflammation thus depends on the sequential appearance and activation of several different CAMs on the surfaces of the interacting cells. Different types of leukocytes express different integrins, though all contain the β2 subunit. Nonetheless, all leukocytes move into tissues by the general mechanism depicted in Figure 20-42.

Many of the CAMs used to direct leukocyte adhesion are shared among different types of leukocytes and target tissues, yet often only a particular type of leukocyte is directed to a particular tissue. How is this specificity achieved? A three-step model has been proposed to account for the cell-type specificity of such leukocyte-endothelium interactions. First, endothelial activation promotes initial relatively weak, transient, and reversible binding (e.g., the interaction of selectins and their carbohydrate ligands). Without additional local activation signals, the leukocyte will quickly move on. Second, cells in the immediate vicinity of the site of infection or inflammation

release or express chemical signals such as chemokines and PAFs that activate only special subsets of the transiently attached leukocytes, depending on the types of chemokine receptors those leukocytes express. Third, additional activation-dependent CAMs (e.g., integrins) engage their binding partners, leading to strong, sustained adhesion. Only if the proper combination of CAMs, binding partners, and activation signals are engaged together with the appropriate timing at a specific site will a given leukocyte adhere strongly. Such combinatorial diversity and cross talk allows a small set of CAMs to serve diverse functions throughout the body — a good example of biological efficiency.

Leukocyte-adhesion deficiency is caused by a genetic defect in the synthesis of the integrin β2 subunit. People with this disorder are susceptible to repeated bacterial infections because their leukocytes cannot extravasate properly and thus cannot effectively fight infection within a tissue.

Some pathogenic viruses have evolved mechanisms to exploit cell-surface proteins that participate in the normal response to inflammation. For example, many of the RNA viruses that cause the common cold (rhinoviruses) bind to and enter cells through ICAM-1, and chemokine receptors can be important entry sites for human immunodeficiency virus (HIV), the cause of AIDS. Integrins appear to participate in the binding or internalization of a wide variety of viruses, including reoviruses (which cause fever and gastroenteritis, especially in infants), adenoviruses (which cause conjunctivitis and acute respiratory disease), and foot-and-mouth disease virus (which causes fever in cattle and pigs). ■

KEY CONCEPTS OF SECTION 20.5

Adhesive Interactions in Motile and Nonmotile Cells

- Many cells have integrin-containing clusters of proteins (e.g., focal adhesions, 3-D adhesions, podosomes) that physically and functionally connect cells to the ECM and facilitate inside-out and outside-in signaling.
- Via interaction with integrins, the three-dimensional structure of the ECM surrounding a cell can profoundly influence the behavior of the cell.
- Integrins exist in at least three conformations (bent closed, extended closed, extended open) that differ in their affinity for ligands and in their interactions with cytosolic adapter proteins and extracellular ligands (see Figure 20-39). Switching between these conformations depends, in part, on the application of mechanical force and allows regulation of integrin activity, which is important for control of cell adhesion and movements. Talin, one of many cytoplasmic adaptors that interact with integrins, can function as a mechanosensor. Talin can mediate integrin's interactions with the cytoskeleton and thus its switching between conformational and functional states.
- Dystroglycan, an adhesion receptor, forms a large complex with dystrophin, other adapter proteins, and signaling molecules (see Figure 20-41). This complex links the actin cytoskeleton to the surrounding ECM, providing mechanical stability to muscle.

Mutations in various components of this complex cause different types of muscular dystrophy.

- Neural cell-adhesion molecules, which belong to the immunoglobulin (Ig) family of CAMs, mediate Ca^{2+}- independent cell-cell adhesion in neural and other tissues.
- The combinatorial and sequential interaction of several types of CAMs (e.g., selectins, integrins, and ICAMs) is critical for the specific adhesion of different types of leukocytes to endothelial cells in response to local signals induced by infection or inflammation (see Figure 20-42).

20.6 Plant Tissues

We turn now to the assembly of plant cells into tissues. The overall structural organization of plants is generally simpler than that of animals. For instance, plants have only four broad types of cells, which in mature plants form four basic classes of tissue: *dermal tissue* interacts with the environment, *vascular tissue* transports water and dissolved substances such as sugars and ions, space-filling *ground tissue* constitutes the major sites of metabolism, and *sporogenous tissue* forms the reproductive organs. Plant tissues are organized into just four main organ systems: *stems* have support and transport functions, *roots* provide anchorage and absorb and store nutrients, *leaves* are the sites of photosynthesis, and *flowers* enclose the reproductive structures. Thus at the cell, tissue, and organ levels, plants are generally less complex than most animals.

Moreover, unlike animals, plants do not replace or repair old or damaged cells or tissues; they simply grow new organs. Indeed, the developmental fate of any given plant cell is primarily based on its position in the organism rather than on its lineage, whereas both are important in animals (see Chapter 22). In both plants and animals, a cell's direct communication with its neighbors is important. Most importantly for this chapter, and in contrast with animals, few cells in plants contact one another directly through molecules incorporated into their plasma membranes. Instead, plant cells are typically surrounded by a rigid cell wall that contacts the cell walls

of adjacent cells (Figure 20-43a). Also in contrast with animal cells, a plant cell rarely changes its position in the organism relative to other cells. These features of plants and their organization have determined the distinctive molecular mechanisms by which plant cells are incorporated into tissues and communicate with one another.

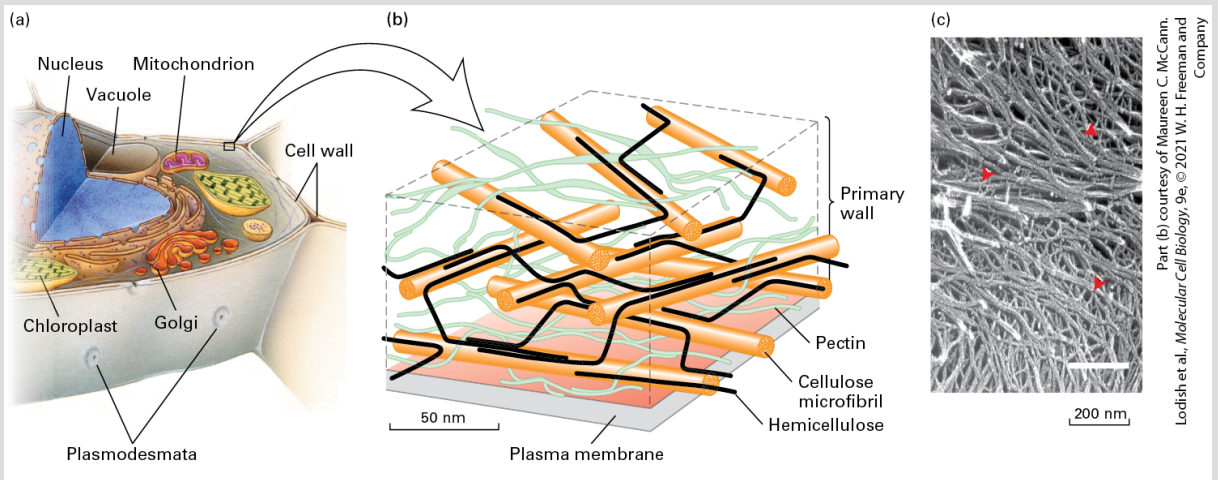


FIGURE 20-43 Structure of the plant cell wall. (a) Overview of the organization of a typical plant cell, in which the organelle-filled cell with its plasma membrane is surrounded by a well-defined extracellular matrix called the *cell wall*. (b) Schematic representation of the cell wall of an onion. Cellulose (highly ordered glucose polymers) and hemicellulose (amorphous polysaccharide) are arranged into at least three layers in a matrix of the polysaccharide pectin and glycoproteins. The sizes of the polymers and their separations are drawn to scale. To simplify the diagram, most of the hemicellulose cross-links and other matrix constituents (e.g., glycoproteins such as extensin, complex organic polymers called *lignins*) are not shown. (c) Quick-freeze deep-etch electron micrograph of the cell wall of a garden pea in which some of the pectin molecules were removed by chemical treatment. The abundant thicker fibers are cellulose microfibrils, and the thinner fibers are hemicellulose cross-links (red arrowheads).

[Part (c) republished with permission from Oxford University Press, from T. Fujino et al., 2000, “Characterization of Cross-Links Between Cellulose Microfibrils, and Their Occurrence

During Elongation Growth in Pea Epicotyl," *Plant Cell Physiol*. 2000, **41**(4):486–494; permission conveyed through Copyright Clearance Center, Inc.]

The Plant Cell Wall, a Plant's ECM, Is a Laminate of Cellulose Fibrils in a Matrix of Polysaccharides and Glycoproteins

The plant **cell wall**, an extracellular matrix that is mainly composed of polysaccharides and is about 0.2 μm thick, completely coats the outside of the plant cell's plasma membrane. This structure serves some of the same functions as the ECM produced by animal cells, even though the two structures are composed of entirely different macromolecules and have a different organization. About 1000 genes in the plant *Arabidopsis*, a small flowering plant (also called "thale cress"; see Chapters 1 and 7), are devoted to the synthesis and functioning of its cell wall, including approximately 414 glycosyltransferase genes, which encode enzymes that transfer specific sugar residues to proteins or polysaccharides, and more than 316 glycosyl hydrolase genes, which encode enzymes that degrade sugar containing polymers. Similar to animal ECMs, the plant cell wall organizes cells into tissues, signals a plant cell to grow and divide, and controls the shapes of plant organs (morphogenesis). It is a dynamic structure that plays important roles in controlling the differentiation of plant cells during

embryogenesis and growth, and it provides a barrier to protect against pathogen infection. Just as the ECM helps define the shapes of animal cells, the cell wall defines the shapes of plant cells. When the cell wall is digested away from plant cells by hydrolytic enzymes, spherical cells enclosed by a plasma membrane are left.

Because a major function of the plant cell wall is to withstand the turgor pressure of the cell (between 14.5 and 435 pounds per square inch; as much as three orders of magnitude greater than in animal cells and about 13 times greater than in an automobile tire; see Chapter 11), the cell wall is built for lateral strength. It is arranged into layers of **cellulose** microfibrils: bundles of 18–36 parallel chains of extensively hydrogen-bonded, long (as much as 4 μm or greater), linear polymers of glucose in β glycosidic linkages. The cellulose microfibrils are embedded in a matrix composed of *pectin*, a negatively charged polymer of D-galacturonic acid and other monosaccharides, and *hemicellulose*, a short, highly branched polymer of several five- and six-carbon monosaccharides. The mechanical strength of the cell wall depends on cross-linking of the microfibrils by hemicellulose chains (Figure 20-43b, c). The layers of microfibrils prevent the cell wall from stretching laterally. Cellulose microfibrils are synthesized on the exoplasmic face of the plasma membrane from UDP-glucose and ADP-glucose formed in the cytosol. The polymerizing enzyme, called *cellulose synthase*, moves within the plane of the plasma membrane along tracks of intracellular microtubules as cellulose is formed, providing a

distinctive mechanism for intracellular-extracellular communication and ensuring that the cellulose microfibrils are oriented properly to permit cell wall, and thus whole-cell, growth.

Unlike cellulose, pectin and hemicellulose are synthesized in the Golgi complex and transported to the cell surface, where they form an interlinked network that helps bind the walls of adjacent cells to one another and cushions them. When purified, pectin binds water and forms a gel in the presence of Ca^{2+} and borate ions — hence the use of pectins in many processed foods. As much as 15 percent of the cell wall may be composed of *extensin*, a glycoprotein that contains abundant hydroxyproline and serine. Most of the hydroxyproline residues are linked to short chains of arabinose (a five-carbon monosaccharide), and the serine residues are linked to galactose. Carbohydrate accounts for about 65 percent of extensin by weight, and its protein backbone forms an extended rodlike helix with the hydroxyl or *O*-linked carbohydrates protruding outward. *Lignin* — a complex, insoluble polymer of phenolic residues — associates with cellulose and is a strengthening material. Like cartilage proteoglycans, lignin resists compression forces.

The cell wall is a selective filter whose permeability is controlled largely by pectins. Whereas water and ions diffuse freely across and through cell walls, the diffusion of large molecules, including proteins larger than 20 kDa, is limited. This limitation may explain why many plant hormones are small, water-soluble molecules,

which can diffuse across the cell wall and interact with receptors in the plasma membrane of plant cells.

Loosening of the Cell Wall Permits Plant Cell Growth

Because the cell wall surrounding a plant cell prevents it from expanding, the wall's structure must be loosened when the cell grows. The amount, type, and direction of plant-cell growth are regulated by small-molecule hormones called *auxins*. The auxin-induced weakening of the cell wall permits the expansion of the intracellular vacuole (see Figure 20-42a) by uptake of water, leading to elongation of the cell. We can grasp the magnitude of this phenomenon by considering that, if all cells in a redwood tree were reduced to the size of a typical liver cell, the tree would have a maximum height of only 1 meter, about a hundredfold less than normal.

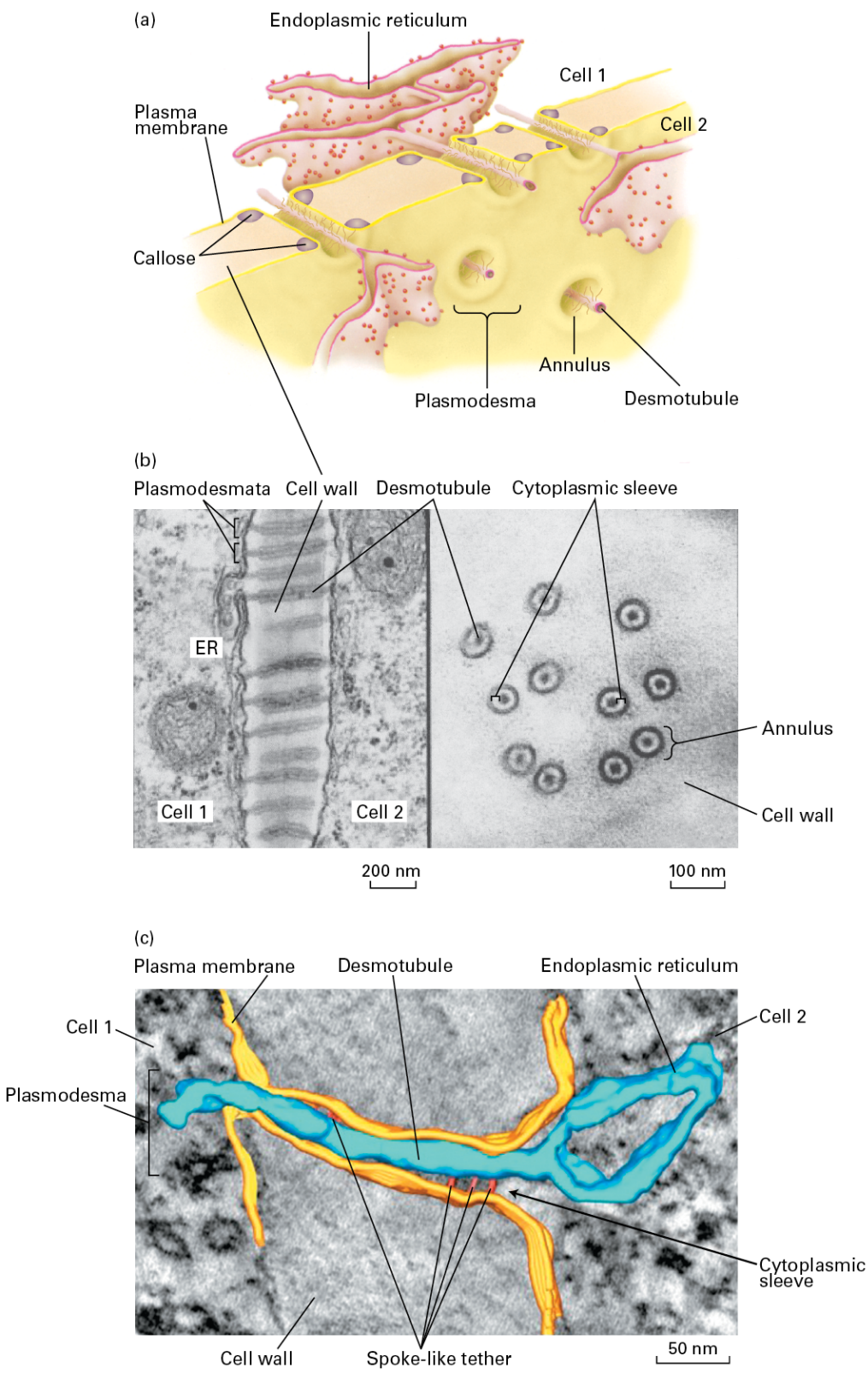
The cell wall undergoes its greatest changes at the *meristem* in a root or shoot tip. Meristems are where cells divide and grow, as described in Chapter 22. Young meristematic cells are connected by thin, flexible *primary cell walls*, which can be loosened and stretched to allow subsequent cell elongation. Their major components are cellulose, pectin, and polysaccharides called *xyloglucans*. After cell elongation ceases, the cell wall is generally thickened, either by the

secretion of additional macromolecules into the primary wall or, more usually, by the formation of a *secondary cell wall* composed of several layers. Cellulose, lignin, and polysaccharides called *xylans* and *glucomannans* are the major components in secondary cell walls. In mature tissues such as the xylem — the tubes that conduct salts and water from the roots through the stems to the leaves — most of the cell eventually degenerates, leaving only the cell wall. The unique properties of wood and of plant fibers such as cotton are due to the molecular properties of the cell walls in the tissues of origin.

Plasmodesmata Directly Connect the Cytosols of Adjacent Cells

The presence of a cell wall separating cells in plants imposes barriers to cell-cell communication not faced by animals. One distinctive mechanism used by plant cells to communicate directly is specialized cell junctions called **plasmodesmata**, which are large tubes that extend through the cell wall (Figure 20-44) and were discovered in 1885. Like gap junctions, plasmodesmata can function as channels that directly connect the cytosol of a cell to that of an adjacent cell without a membrane barrier. Plasmodesmata appear to play an especially important role in protection from pathogens and in regulating the development of plant cells and tissues, as is

suggested by their ability to mediate intracellular movement of transcription factors and ribonuclear protein complexes.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 20-44 Plasmodesmata. (a) Schematic model of plasmodesmata, showing the desmotubule, an extension of the endoplasmic reticulum (ER), and the annulus, a plasma-membrane-lined channel filled with cytosol that interconnects the cytosols of adjacent cells. The regulated deposition of a glucose polymer called *callose* in the extracellular spaces in the cell wall adjacent to the entrances of the channels has the potential to block intercellular transport through the plasmodesmata, apparently by forcing the closing of the channels by narrowing the annulus. (b) Electron micrographs of thin sections of a sugarcane leaf (brackets indicate individual plasmodesmata). *Left:* Longitudinal view, showing the ER and desmotubule running through each annulus. *Right:* Perpendicular cross-sectional views of plasmodesmata, each with its annulus, desmotubule, and cytoplasmic sleeve. (c) Cells from the plant *Arabidopsis* were grown in culture, frozen, and, while cold, fixed with glutaraldehyde and stained with osmium tetroxide and uranyl acetate. Multiple views of one plasmodesma were visualized by tilting the sample in an electron microscope and a three-dimensional tomographic model of the structure was computed. Shown is the electron micrograph of one view (black and white) with the colored model of its plasmodesma; plasmamembrane = yellow; ER and desmotubule membrane = blue; and spoke-like tethers connecting these membranes = red. The gap between the two membranes is the cytoplasmic sleeve..

[Part (b) republished with permission from Springer, from K. Robinson-Beers and R. F. Evert, 1991, "Fine Structure of Plasmodesmata in Mature Leaves of Sugarcane," *Planta* **184**(3):307–318; permission conveyed through Copyright Clearance Center, Inc. Part (c) republished with permission from Springer Nature, from W. J. Nicolas et al., 2017, "Architecture and Permeability of Post-Cytokinesis Plasmodesmata Lacking Cytoplasmic Sleeves." *Nature Plants* **3**:17082; permission conveyed through Copyright Clearance Center, Inc.]

Although plasmodesmata and gap junctions appear to resemble each other functionally with respect to forming channels for direct diffusion of molecules from one cell's cytosol to that of another,

their structures differ dramatically in two significant ways. First, in animal-cell gap junctions, the membranes of the connected cells are distinct, and the cytosols are connected by a protein-lined tube of connexon hemichannels (see Figure 20-21). In plant plasmodesmata (see Figure 20-44), the plasma membranes of the adjacent plant cells merge to form a continuous single membrane that lines the wall of the channel, called the *annulus*, that connects the two cells. The diameter of the annulus varies from about 30–60 nm; its length can vary and may be greater than 1 μm. The lipid and protein composition of the specialized plasma membrane lining the annulus differs from that of the rest of the plasma membrane that surrounds the cell. Second, passing though the center of the plasmodesmata's annulus is a narrow tube, called a *desmotubule*, that is an extension of endoplasmic reticulum (ER) from each of the connected cells. Between the desmotubule and the plasmodesmata's specialized plasma membrane is a gap whose size is regulated and varies between 2–10 nm. This gap is called the *cytoplasmic sleeve*, through which molecules diffuse from the cytosol of one cell to that of the other. Filamentous actin runs throughout the length of the cytoplasmic sleeve. The close apposition of the plasma membrane and ER/desmotubule's membrane in plasmodesmata is an example of a specialized membrane contact site (MCS), such as that described for mitochondria and ER in Chapter 12. A set of transmembrane proteins in the desmotubule called *MCTPs* can span the gap of the cytoplasmic sleeve and bind to anionic phospholipids on the plasma membrane, thus tethering these two membranes. The

MCTP proteins appear to compose spoke-like projections observed by electron microscopy that tether the two membranes (see Figure 20-44c), help determine the size and properties of the cytoplasmic sleeve, and thus control the movement of molecules through plasmodesmata. In animal cells, tunneling nanotubes resemble plasmodesmata in that they are membrane-lined channels that permit molecules to flow between cells (see Section 20.2). However, unlike plasmodesmata, there is no evidence for ER passing through tunneling nanotubes.

There are simple plasmodesmata (with a single pore, like those in Figure 20-44) and complex plasmodesmata that branch into multiple channels. The density of plasmodesmata varies depending on the plant and cell type, and even the smallest meristematic cells have more than a thousand connections with their neighbors. Proteomic analysis suggests that there are more than 115 different proteins that are concentrated in plasmodesmata and play important roles in determining their structures and functions. These include the MCTPs, actin-binding proteins (e.g., formin and the adaptor protein NET1A), a variety of receptor-like proteins (e.g., kinases), and enzymes.

Many types of molecules travel from cell to cell through plasmodesmata. Molecules smaller than about 1000 Da, including a variety of metabolic and signaling compounds (ions, sugars, amino acids), can generally diffuse freely through the cytoplasmic sleeve, whose size and properties are highly regulated. In some

circumstances, the channel is clamped shut; in others, it is dilated sufficiently to permit the passage of molecules larger than 10,000 Da, including some transcription factors, nucleic acid/protein complexes, metabolic products, and even plant viruses. Some of these require special chaperones to facilitate transport. Soluble molecules pass through the cytosolic sleeve, whereas membrane-bound molecules or certain proteins within the ER lumen may pass from cell to cell via the desmotubule. The deposition and breakdown of a glucose polymer called *callose* in the extracellular spaces adjacent to the entrances of the channels (see Figure 20-44a) is thought to regulate the closing and opening of the channels, respectively. Specialized kinases may also phosphorylate plasmodesmal components to regulate their activities (e.g., opening of the channels). Among the factors that affect the permeability of plasmodesmata is the cytosolic Ca^{2+} concentration: an increase in cytosolic Ca^{2+} reversibly inhibits movement of molecules through these structures.

The Molecules That Plants Depend on for Adhesion and Mechanotransduction Differ from Those in Animals

As noted earlier in this section, there are dramatic differences in the organization of cells into tissues in plants and animals. These

differences have determined the distinctive molecular mechanisms by which plant cells are incorporated into tissues and communicate with one another, including their cell-cell and cell-matrix interactions.

Plant Adhesion Molecules

Systematic analyses of the *Arabidopsis* genome and biochemical analyses of other plant species have provided no evidence for the existence of plant homologs of most animal CAMs, adhesion receptors, and ECM components; plants use other adhesive molecules for their cell-cell and cell-matrix interactions. Currently, our understanding of cell-matrix adhesion in plants is less advanced than our understanding of cell-matrix adhesion in animals.

Because plant cells are surrounded by thick cell walls, their assembly into tissues is primarily determined by cell-matrix interactions. Most of the polysaccharides in the cell wall (cellulose, hemicellulose, pectins, glycoproteins) can cross-link to one another. The adhesive properties of pectins are thought to be especially important for determining both the physical characteristics of cell walls and the interactions of the walls with the plasma membranes of the underlying cells. Disruption of the gene encoding glucuronyltransferase 1, a key enzyme in pectin biosynthesis, dramatically interferes with the synthesis of specialized pectins that help hold the cells in meristems tightly together. As a consequence,

the mutation prevents normal cell adhesion and differentiation into photosynthetic cells. In vitro binding assays, combined with in vivo studies and analyses of plant mutants, have identified other macromolecules in the ECM that are important for specialized types of adhesion. For example, a cysteine-rich protein called *stigma/stylar cysteine-rich adhesin (SCA)* and a specialized pectin that binds to SCA mediate normal adhesion of pollen, which contains sperm cells, to the stigma or style in the female reproductive organ of the Easter lily.

A number of plasma membrane proteins exhibit the potential to bind to cell wall polysaccharides, particularly pectins, and thus mediate cell-matrix adhesion. These include receptor-like proteins, such as Wall-Associated Kinases (WAKs), GPI-anchored proteins, Hydroxyproline-Rich Glycoproteins (HRGPs), and Glycosyl Inositol Phosphoceramides (GIPCs). For example, five WAKs and WAK-like proteins are expressed in the plasma membrane of *Arabidopsis* cells. These transmembrane proteins have a cytoplasmic serine/threonine kinase domain, and their extracellular regions contain multiple epidermal growth factor (EGF) repeats, frequently found in animal cell-surface receptors. Some WAKs have an extracellular pectin-binding domain that can recognize and bind full-length pectin and pectin degradation fragments. Such binding has been proposed to help cells monitor and respond to the status of the cell wall during normal growth and in the context of cell-wall damage (wounding) or infection by pathogens. Thus some WAKS in plant cells appear to be

analogous to adhesion receptors in animal cells, binding and sensing the ECM and mediating outside-in signaling.

Mechanosensation in Plants

As with animal cells, plant cells are subject to a variety of mechanical forces (stresses and strains), both internal (e.g., turgor pressure, cytoskeletal forces) and external (e.g., wind), which they sense (mechanosensing) and to which they respond (e.g., alterations in cytoskeletal organization, calcium flux, gene expression, wood formation, cell wall stiffness, and extent and direction of growth). Possible mechanosensors in the cell wall include components whose structures respond to changes in mechanical forces and thus change the properties of the wall and its interactions with the plasma membrane. Potential mechanosensors in the cell's plasma membrane are those components that bind to the cell wall (e.g., receptor-like kinases such as pectin-binding WAKs) and mechanosensitive ion channels. There is evidence that suggests inside-out and outside-in communication via adhesive molecules contribute to plant responses to mechanical forces.

One well-known example of macroscopic plant mechanosensation is the Venus flytrap (*Dionaea muscipula*), first described by the colonial governor of North Carolina, Arthur Dobbs, in 1759. Charles Darwin studied the Venus flytrap extensively, writing a book in 1875 in which he called the plant "one of the most

wonderful in the world" and once noting that he cared more for one of them "than the origin of all the species in the world." The Venus flytrap detects the presence of an insect on its leaf when an insect hits one of its sensory hairs, which are around 200 μm in diameter at the base and 2 mm long. The sensory hair is a multicellular organ that, when bent by prey, puts shear stress on a cluster of cells at its base, which opens mechanosensitive calcium channels. Calcium entry triggers an action potential in the leaf (see Chapter 23) that leads to closing the trap and hormonally induced secretion of digestive enzymes that disassemble the prey into basic nutrients over a one-to-two-week period. ■

We are all familiar with one aspect of cell-wall–mediated adhesion in plants: the softening of fruits, such as tomatoes and strawberries, during ripening. The ripening of fruits is a complex process that influences color, aroma, flavor, nutrient content, and texture (including softening and water content). Ripening of seed-bearing fruit is advantageous because it encourages fruit-eating animals to eat and disperse the seeds or breaks down the uneaten fruit's structure (softening and decay) to release the seeds in situ. However, excessive softening can interfere with successful storage and transport and can reduce shelf life. Expression of a group of cell-wall–degrading enzymes, including pectate lyase (PL), and their remodeling of cell walls by degrading pectin plays a key role in determining the rate and extent of softening. Experiments using both strawberries and tomatoes have suggested that inhibiting the

expression of pectate lyase using either CRISPR gene editing or RNA interference (see Chapter 6) can substantially delay strawberry and tomato softening and may not block other desirable consequences of ripening (changes in color, flavor, etc.). In the future, targeted genetic modifications of cell-wall metabolism may improve the quality of available fruits while reducing the costs of production and distribution. ■

KEY CONCEPTS OF SECTION 20.6

Plant Tissues

- The organization of cells into tissues in plants is fundamentally different from the assembly of animal tissues, primarily because each plant cell is surrounded by a relatively rigid cell wall.
- The plant cell wall comprises layers of cellulose microfibrils embedded within a matrix of hemicellulose, pectin, extensin, and other less abundant molecules.
- Cellulose, a large, linear glucose polymer, assembles spontaneously into microfibrils stabilized by hydrogen bonding.
- The cell wall defines the shapes of plant cells and restricts their elongation. Auxin-induced loosening of the cell wall permits cell elongation.
- Adjacent plant cells can communicate through plasmodesmata, junctions formed by tubes of specialized plasma membrane and ER that connect the cytosols of adjacent cells, and by allowing molecules to pass between cells (see Figure 20-44).
- Plants do not produce homologs of the common adhesion molecules found in animals.
- The molecules mediating cell-matrix adhesion in plants are not as well-defined as those in animals, although both cell wall and plasma membrane molecules implicated in plant cell adhesion have been identified and are under investigation.
- As in the case of animal cells, mechanotransduction in plants plays an important role in maintaining tissue integrity and directing differentiation and development.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

adapter proteins
adherens junction
adhesion receptor
anchoring junction
apical
basal
basal lamina
basolateral
cadherin
cell-adhesion molecule (CAM) or cell-cell adhesion receptor
cell junction
cell-matrix adhesion receptor

cell wall
cellulose
cis (lateral) interactions
classical cadherin
clustered protocadherin
collagen
connective tissue
connexin
desmosomal cadherin
desmosome
elastin
epithelium
epithelial-to-mesenchymal transition
extracellular matrix (ECM)
fibrillar collagen
fibroblast
fibronectin
focal adhesion/focal contact
gap junction
glycosaminoglycan (GAG)
heterophilic binding
heterotypic adhesion
homophilic binding
homotypic adhesion
hyaluronan
inside-out signaling
integrin
isoform

laminin
lateral
matriglycan
matrix metalloproteases (MMPs)
mechanosensor
mechanotransduction
multi-adhesive matrix protein
N-linked oligosaccharide
O-linked oligosaccharide
outside-in signaling
paracellular pathway
PDZ domain
perlecan
plasmodesmata
proteoglycan
RGD motif
selectin
tight junction
talin
trans (intercellular) interactions
transcellular pathway

Review the Concepts

1. Describe the two phenomena that give rise to the diversity of adhesion molecules, such as cadherins. What additional phenomenon gives rise to the diversity of integrins?

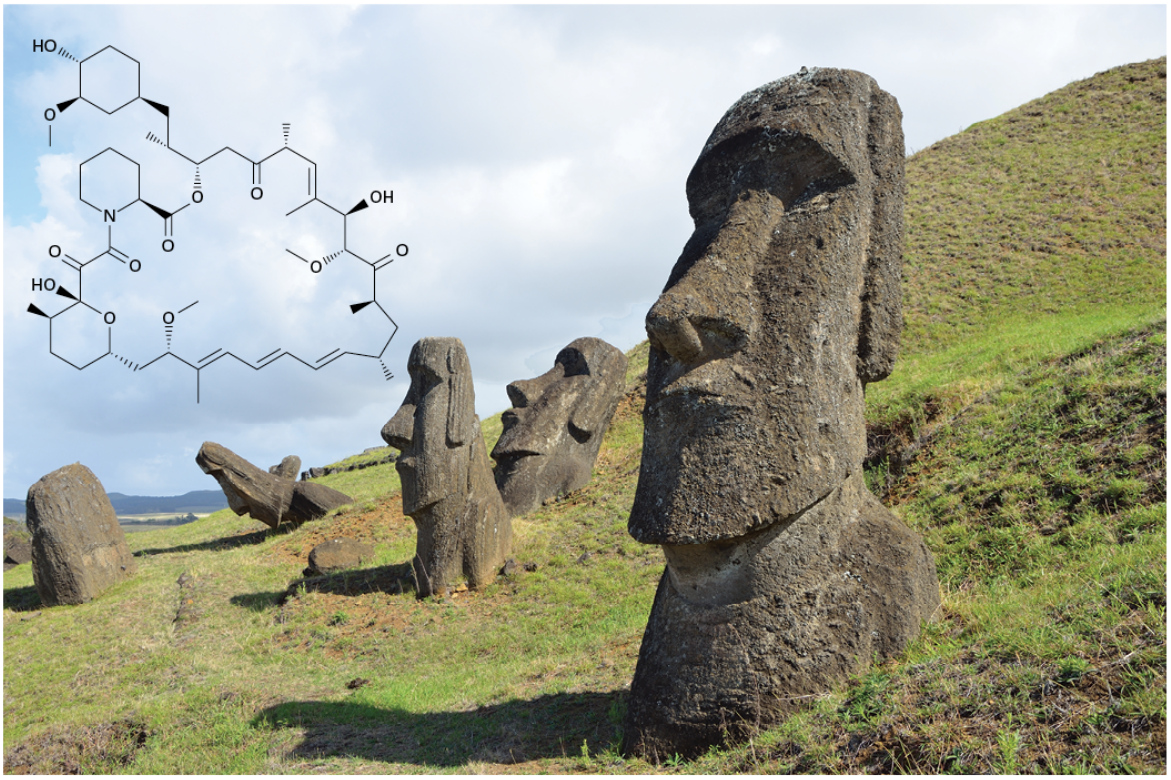
2. Cadherins are known to mediate homophilic interactions between cells. What is a homophilic interaction, and how can it be demonstrated experimentally for E-cadherins? What component of the extracellular environment is required for the homophilic interactions mediated by cadherins, and how can this requirement be demonstrated?
3. Together with their role in connecting the lateral membranes of adjacent epithelial cells, adherens junctions play a role in controlling cell shape. What associated intracellular structure and proteins are involved in this role?
4. What is the normal function of tight junctions? What can happen to tissues when tight junctions do not function properly?
5. Gap junctions between cardiac muscle cells and gap junctions between uterine smooth-muscle cells form connections that provide for rapid communication. What is this phenomenon called? How is communication among uterine smooth-muscle cells up-regulated for parturition (childbirth)?
6. What is collagen, and how is it synthesized? How do we know that collagen is required for tissue integrity?
7. Explain how changes in integrin structure mediate outside-in and inside-out signaling.
8. Compare the functions and properties of each of three types of macromolecules that are abundant in the ECM of all tissues.

9. Many proteoglycans have signaling roles. Regulation of feeding behavior by syndecans in the hypothalamic region of the brain is one example. How is this regulation accomplished?
10. You have synthesized an oligopeptide containing an RGD motif surrounded by other amino acids. What is the effect of this peptide when added to a fibroblast cell culture grown on a layer of fibronectin adsorbed to the tissue culture dish? Why does this happen?
11. Describe the major activity and possible localization of the three major subgroups of proteins that remodel or degrade the ECM in physiological or pathological tissue remodeling. Identify a pathological condition in which these proteins play a key role.
12. Blood clotting is a crucial function for mammalian survival. How do the multi-adhesive properties of fibronectin lead to the recruitment of platelets to blood clots?
13. How do changes in molecular connections between the ECM and the cytoskeleton give rise to Duchenne muscular dystrophy?
14. To fight infection, leukocytes move rapidly from the blood into sites of infection in the tissues. What is this process called? How are adhesion molecules involved in this process?
15. The structure of a plant cell wall needs to loosen to accommodate cell growth. What signaling molecule controls this process?

16. Compare plasmodesmata in plant cells with gap junctions and tunneling nanotubes in animal cells.

CHAPTER 21

Responding to the Cellular Environment



Yulia Malega/Shutterstock.

Rapamycin (in black), a clinically important antifungal drug and inhibitor of the mTORC kinase complex, was isolated from the bacterium *Streptomyces hygroscopicus* found on Easter Island. It was named after the native name of the island, Rapa Nui.

OUTLINE

21.1 Regulating Blood Glucose Level
21.2 Integrating Cell Growth Signals with Nutrient and Energy Levels

21.3 Responding to Changes in the Levels of Cholesterol and Unsaturated Fatty Acids
21.4 Responding to Low Oxygen
21.5 Responding to Elevated Temperatures
21.6 Sensing Day and Night: Circadian Rhythms
21.7 Sensing and Responding to the Physical Environment

All cells sense and respond to changes in their environment, often by changing their pattern of gene expression. Bacteria constantly monitor the levels of sugars and amino acids in their environment and respond by swimming toward sources of these nutrients and inducing or repressing genes that enable them to most efficiently metabolize these molecules. Yeasts and other unicellular, eukaryotic organisms also sense the presence of sugars, amino acids, and many other metabolites in their environment and adjust their biosynthetic pathways, membrane transporters, and metabolism accordingly. In Chapter 1, we saw how the green unicellular alga *Chlamydomonas reinhardtii* senses light and swims toward or away from the light source depending on its intensity.

This chapter focuses on the multiple ways cells in metazoans, and particularly vertebrates, sense the levels of glucose, amino acids, lipids such as cholesterol, oxygen, and other chemicals in their environment, as well as day and night and contact with other cells, and respond appropriately. These extracellular signals can have short-term effects that usually occur within minutes and are induced by modification of existing enzymes or other proteins, as we learned in Chapters 3 and 15. All of the extracellular signals we

discuss also affect gene expression and thus induce long-term changes in cell function.

The pathways we describe have much in common with the signaling pathways described in Chapters 8, 15, and 16. **Sensors**, like receptors for extracellular signaling molecules, generally bind a specific target molecule and then undergo a conformational change; **mechanoreceptors** sense changes in the physical environment (e.g., pressure or tension, see Chapter 20). These sensors activate **signal transduction pathways** that can have one or many steps. Ultimately one or more **effector proteins** become activated that induces the cell's response, which can be a short-term change in a metabolic pathway, or induction or repression of specific genes, or both. **Signal amplification** and **feedback repression** (see Chapter 15) characterize all of the signal transduction pathways we describe in this chapter, enabling an appropriate but not overreactive adjustment of the cell's metabolism and pattern of gene expression. These pathways enable a cell to respond to a variety of stresses, for instance, lack of a nutrient or an elevated temperature, and enable **homeostasis**, the restoration of the cell's normal chemical and physical conditions and even the survival of the cell itself.

The availability of blood glucose is regulated during periods of abundance (following a meal) or scarcity (fasting) principally by two hormones, insulin and glucagon. As we learn in Section 21.1, when blood glucose concentrations rise too high, β islet cells in the pancreas, cells that synthesize and store insulin, secrete insulin into

the bloodstream. (The 1923 Nobel Prize in Physiology or Medicine was awarded for the discovery of insulin.) Insulin signals fat and muscle cells to increase their uptake of glucose from the blood; it also signals liver cells to store glucose as glycogen, lowering the glucose concentration in the blood. When blood glucose concentration falls too low, other cells in the islets, α cells, release glucagon from secretory vesicles; glucagon instructs liver cells to release glucose into the blood (see Chapter 15). Working together, insulin and glucagon maintain blood glucose homeostasis, restoring the normal glucose levels of about 5 mM in humans. Defects in these pathways lead to major diseases, including diabetes and cardiovascular disease, with dire consequences for the individual and increasingly for public health.

Cell proliferation, the doubling of cell numbers, requires the presence of one or more growth factors, as discussed in Chapter 16. Cell division also requires the presence of sufficient ATP, amino acids, and other nutrients to produce all of the proteins, nucleic acids, membranes, and other components required to double the cell mass in preparation for division. In Section 21.2 we discuss the **mTOR protein kinase**, a component of a large multiprotein complex called **mTORC1**. mTORC1 contains multiple sensor proteins that detect the levels of specific nutrients, energy availability (ATP levels), and the activity of intracellular signal transduction proteins activated downstream of growth factor receptors. mTOR kinase becomes activated and consequently stimulates cell growth, metabolism, and proliferation only when all of these signals are

present simultaneously; we will see how multiple sensors in the mTORC1 complex interact to activate the mTOR kinase.

Among the many nutrients essential for cell growth and function are membrane lipids and their biosynthetic precursors, including cholesterol and unsaturated fatty acids (see Chapters 2 and 10). Cells obtain the necessary lipids either by biosynthesis or import from the extracellular environment. If there are not enough of these lipids, the cells cannot grow, divide, or function properly. If there is too much, the excess can interfere with normal cell function. In Section 21.3, we learn how cells sense levels of intracellular cholesterol and unsaturated fatty acids and how they respond by adjusting the rates of fatty acid and cholesterol biosynthesis and import so that the cholesterol:phospholipid ratio is kept within the narrow, desirable range required for normal membrane functions. Accumulation of excess cholesterol in our arteries is a major cause of cardiovascular diseases and strokes, and in Chapter 10 we learned how the anti-atherosclerosis medications called *statins* inhibit cholesterol biosynthesis and reduce the concentration of low-density lipoproteins (often called "bad cholesterol"; see Chapter 14) and reduce the formation of atherosclerotic plaques. The 1985 Nobel Prize in Physiology or Medicine was awarded for discoveries of how cholesterol metabolism is regulated.

The Nobel Prize in Physiology or Medicine was awarded in 2019 to three physician scientists for elucidating how metazoan cells sense

and adapt to oxygen availability. In Section 21.4, we learn how a reduction in oxygen levels in blood is sensed by many cells in the body and how the transcription factor Hif-1α coordinates many body responses to low oxygen. At ambient oxygen levels, Hif-1α undergoes covalent modifications by oxygen-requiring enzymes that trigger its polyubiquitination and destruction by proteasomes. Hif-1α becomes progressively more stable as oxygen levels decrease. In certain kidney cells, Hif-1α induces the expression of the erythropoietin gene; erythropoietin in turn increases the production of red blood cells that transport oxygen from the lungs to body tissues (see Chapter 16). In other cells, Hif-1α induces expression of genes to promote survival in low-oxygen conditions. Hif-1α is found only in metazoan animals; we also discuss a different family of oxygen-sensing transcription factors that are expressed both in plants and animals and that arose very early in evolution.

All organisms, be they unicellular, plants, invertebrates, or cold-blooded vertebrates, are subject to changes in the temperatures of their internal and external environments. Even vertebrates feel the effects of elevated temperature during fevers or during the hottest periods of summer days. One of the effects of elevated temperatures is partial unfolding or complete denaturation of many cellular proteins that, if uncorrected, could have lethal consequences for the cell. We learn in Section 21.5 how all studied eukaryotes sense the presence of unfolded proteins by the binding of chaperones to unfolded protein domains (see Chapter 3) and how this leads to the **heat-shock response**, during which a particular transcription factor

becomes activated and then induces the synthesis of many chaperones that enable the cell to survive an elevated temperature until the temperature is reduced to normal. Similar chaperones play key roles in facilitating folding of proteins within the endoplasmic reticulum lumen and mitochondria, as discussed in Chapter 13.

Most living things on Earth are subject to a day-night cycle of almost exactly 24 hours. Their cells contain a **circadian clock**, or **circadian oscillator** — a group of proteins whose signaling oscillates with an endogenous period of 24 hours and that receives daily signals from the environment, primarily daylight and darkness, that keep it on time. These synchronized cell clocks, which we discuss in Section 21.6, make it possible for the organism to anticipate the environmental changes that derive from the day-night cycle and adjust its biology and behavior accordingly. In 2017, the Nobel Prize in Physiology or Medicine was awarded to three scientists for their discoveries of molecular mechanisms controlling the circadian rhythm in fruit flies; similar clocks are found in plant, fungal, and vertebrate cells that adjust many aspects of their function to a day-night cycle. Circadian rhythms in certain brain neurons control aspects of organismal behavior, such that mice and lions are active at night while humans are more active during the day.

During development, cells proliferate and differentiate to generate organs of specific sizes. Indeed, certain mutations in the fly *Drosophila melanogaster* (see Chapter 1) result in excessive organ

growth; one was called *Hippo*, as mutations caused massive overgrowth of larvae that looked like a hippopotamus. When part of an adult liver is surgically removed, the liver rapidly regenerates to its former size; clearly gene products must exist that negatively regulate growth when the appropriate organ size is achieved. In Section 21.7, we discuss the **Hippo pathway**, a protein kinase signaling cascade conserved in all metazoans that senses and integrates multiple environmental cues and modulates cell growth in response to the physical environment. This pathway senses physical and mechanical cues from the extracellular matrix in which the cell is imbedded, senses tension (mechanotransduction), and receives signals from the adherens junctions (Chapter 20) that connect them with other cells. The Hippo pathway regulates the expression of multiple genes involved in growth control and differentiation, and we will see that abnormalities in components of this pathway contribute to developmental disorders, cell transformation, and metastases by many types of cancers.

21.1 Regulating Blood Glucose Level

In this section, we consider how multiple hormones and signal transduction pathways interact, focusing on one of the most important physiological control systems: regulation of the body's need for glucose. Defects in these pathways lead to major diseases, particularly type 2 or *adult-onset* diabetes, a principal cause of blindness, kidney failure, and limb amputations in adults. Cellular responses to changes in amino acids and other nutrients, which are largely reflected in alterations in gene expression, are covered in Section 21.2.

Insulin and Glucagon Work Together to Maintain a Stable Blood Glucose Level

Our focus in this section is mainly on vertebrates, whose normal blood glucose concentration is about 5 mM. Because most body cells, including brain neurons, utilize glucose as their primary energy source, a prolonged fall in blood glucose, called *hypoglycemia*, can cause seizures, loss of consciousness, and death. In contrast, prolonged elevations in blood glucose above 9–10 mM can cause diabetes. During normal daily living, the maintenance of normal blood glucose concentrations depends on the balance between two peptide hormones, **insulin** and **glucagon**, which are

made in distinct and adjacent endocrine cells in island-like domains of the pancreas termed *pancreatic islet cells* (Figure 21-1a) and are stored in secretory vesicles prior to regulated secretion (see Figure 14-2). Insulin, which lowers blood glucose, contains two polypeptide chains linked by disulfide bonds and is synthesized by the β cells in the islets (see Figure 14-24). Glucagon is produced by the α cells in the islets and acts to raise blood glucose by inducing cleavage of glycogen into glucose in the liver and secretion of the glucose into the blood (see Chapter 15). The availability of blood glucose is regulated during periods of abundance (following a meal) or scarcity (following fasting) by the amounts of insulin and glucagon secreted from pancreatic islet β and α cells.

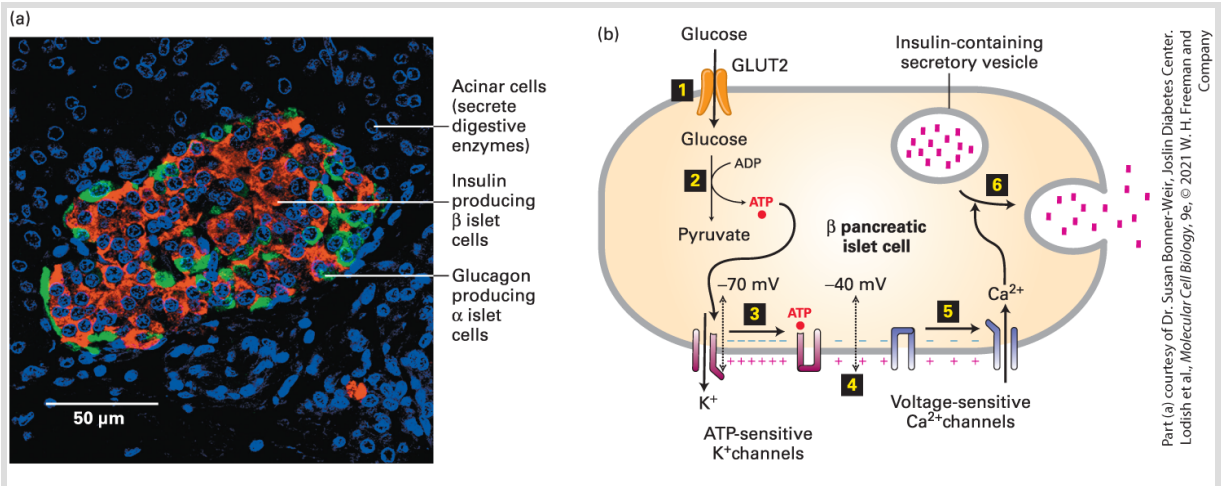


FIGURE 21-1 Secretion of insulin by pancreatic islet β cells in response to a rise in blood glucose above 5 mM. (a) A section of an islet in a human pancreas. Islets comprise about 2 percent of the mass of a pancreas. A human pancreas contains about one million islets, each of which contains about 1000 β cells. β cells, which comprise ~70 percent of islet cells, are revealed by a red histochemical stain for insulin. Glucagon, synthesized in the α cells largely on the periphery of the islet, is visualized by a green stain. The DAPI dye, blue, stains all nuclei. (b) Regulation of insulin secretion. Entry of glucose into pancreatic β cells is mediated by the GLUT2 glucose transporter (step 1). Because the K_m of GLUT2 for glucose

is 20 mM, a rise in extracellular glucose from 5 mM, characteristic of the fasting state, causes a proportional increase in the rate of glucose entry (see Figure 11-5). The conversion of glucose into pyruvate (Figure 12-3) is thus accelerated, resulting in an increase in the concentration of ATP in the cytosol (step 2). The binding of ATP to ATP-sensitive K^+ channels in the β cells closes those channels (step 3), thus reducing the efflux of K^+ ions from the cell. The resulting small depolarization of the plasma membrane from ~ −70 mV to ~ −40 mV (step 4) triggers the opening of voltage-sensitive Ca^{2+} channels (step 5). The influx of Ca^{2+} ions raises the cytosolic Ca^{2+} concentration, triggering the fusion of insulin-containing secretory vesicles with the plasma membrane and the secretion of insulin (step 6). See J. C. Henquin, 2000, *Diabetes* **49**:1751.

Secretion of glucagon from pancreatic α cells is inhibited by insulin when blood glucose is high; secretion is enhanced by signals from the central nervous system when blood glucose should increase. Like the epinephrine receptor, the glucagon receptor, found primarily on liver cells, is coupled to the $G_\alpha s$ protein, whose effector protein is adenylyl cyclase. The binding of glucagon to its receptor induces a rise in cAMP, leading to activation of protein kinase A, which inhibits glycogen synthesis and promotes glycogenolysis, yielding glucose 1-phosphate (see Figure 15-22). Liver cells convert glucose 1-phosphate into glucose, which is released into the blood, thus raising blood glucose back toward its normal fasting level.

Our focus here will be on the hormone insulin, which acts in several short-term and long-term ways to reduce the level of blood glucose:

- Within seconds of its secretion, insulin induces an increase in the uptake of glucose from the blood into muscle and fat cells,

primarily by increasing the number of glucose transporters in the plasma membrane.

- Within seconds to minutes, insulin acts on the liver to both stimulate glycogen synthesis from glucose and accelerate glucose flux through the glycolytic pathway; both processes provide a *sink* that helps lower glucose levels in the blood.
- Insulin released from β cells acts on the nearby α cells in the pancreatic islets to inhibit glucagon synthesis.
- Over a longer time frame of hours, insulin acts on the liver to inhibit synthesis of enzymes that catalyze the process termed **gluconeogenesis** — the synthesis of glucose from smaller metabolites.

In the following sections, we see how insulin secretion is stimulated by a rise in blood glucose and then the multiple ways insulin acts to lower blood glucose.

A Rise in Blood Glucose Triggers Insulin Secretion from the β Islet Cells

After a meal, when blood glucose rises above its normal level of 5 mM, the pancreatic β cells respond to the rise in glucose (and a concurrent rise in amino acids in the blood) by releasing insulin into the blood (Figure 21-1b). We saw in Chapter 14 that these cells store insulin in a dehydrated, almost crystalline form in secretory vesicles; as with all regulated secretory pathways, fusion of these

vesicles with the plasma membrane and secretion of their contents is triggered by a rise in cytosolic Ca^{2+}. Insulin secretion is triggered by a rise in extracellular glucose, which, via GLUT2 glucose transporters (Figure 11-5), causes a proportionate increase in the rate of glucose entry into the cells and a corresponding increase in the rate of glycolysis. The resulting rise in the ratio of cytosolic ATP:ADP concentrations causes closing of an ion channel unique to the β cells, an ATP-gated K^+ channel, reducing the efflux of K^+ ions from the cell. As occurs at the axon termini of nerve cells (see Chapter 23), the resulting depolarization of the plasma membrane from ~ −70 mV to ~ −40 mV triggers the opening of voltage-sensitive Ca^{2+} channels, an increase in cytosolic Ca^{2+}, and insulin secretion.

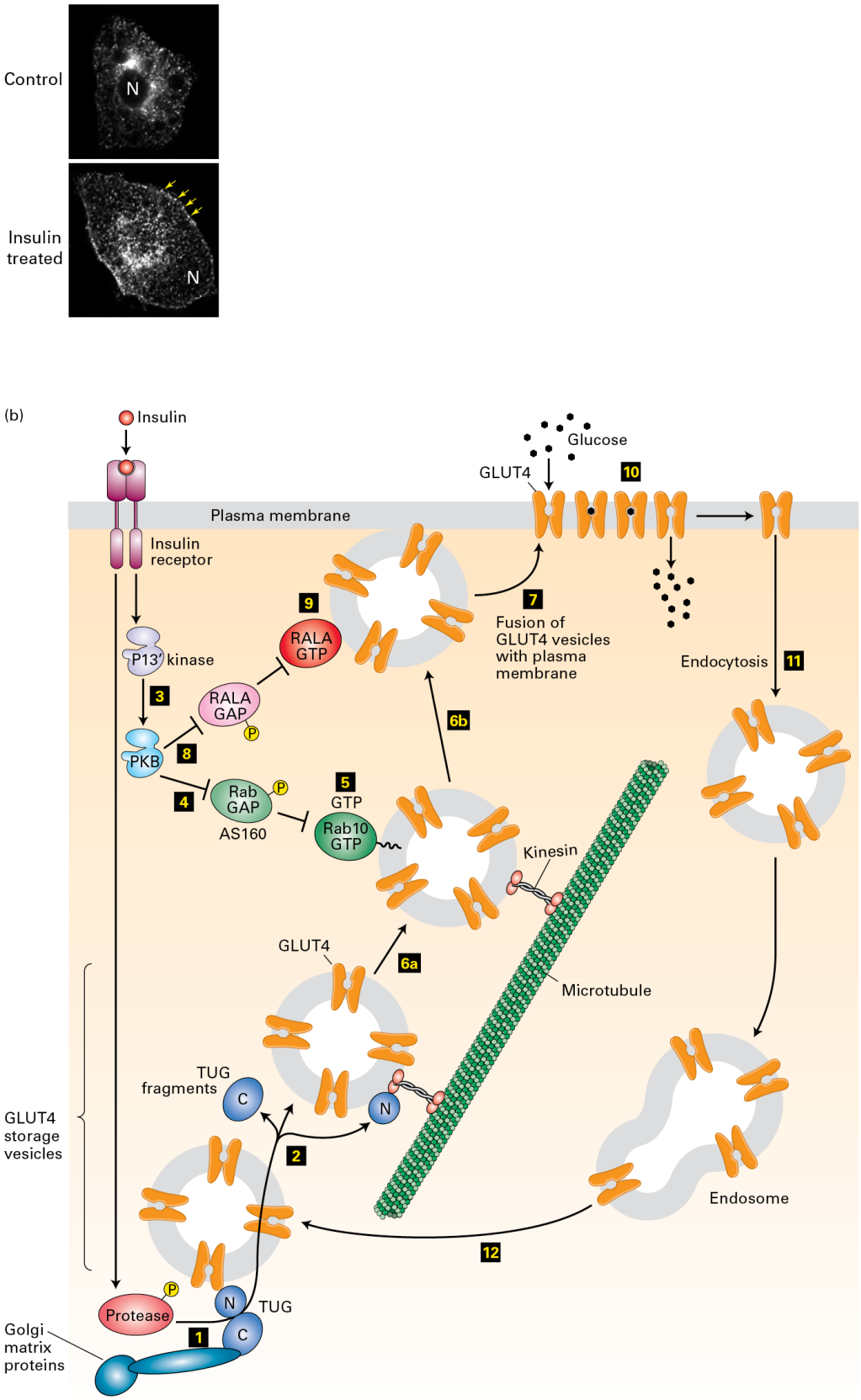
In Fat and Muscle Cells, Insulin Triggers Fusion of Intracellular Vesicles Containing the GLUT4 Glucose Transporter with the Plasma Membrane, Thus Increasing the Rate of Glucose Uptake

Once secreted from pancreatic β cells, insulin circulates in the blood and binds to insulin receptors, which are present on many different kinds of cells, including liver, muscle and adipocyte (fat) cells. In fat and muscle cells the insulin receptor, a receptor tyrosine kinase (see

Chapter 16), activates several signal transduction pathways leading, within minutes, to a ~10-fold increase in the rate of uptake of glucose from the blood. Once imported, glucose is rapidly phosphorylated by hexokinase to form glucose 6-phosphate (Figure 12-3), which cannot be transported out of the cell; the increase in glucose import therefore results in rapid lowering of the level of blood glucose.

Like the plasma membranes of most body cells, the membranes of fat and muscle cells contain the GLUT1 glucose transporter, which allows the cell to import sufficient glucose for its basal metabolic needs. Fat and muscle cells also express large amounts of a glucose transporter, the insulin-responsive glucose transporter GLUT4, made only by these cells. In resting (unstimulated) cells, virtually all of the GLUT4 is located in small vesicles in the cytosol (Figure 21-2a). While some GLUT4 is in endosomes, most is in a small organelle unique to fat and muscle cells termed the **GLUT4 storage vesicle** (**GSV**; Figure 21-2b). These vesicles, which are separate from but, as we will see, derived from endosomes, are tethered to the Golgi matrix, a network of coiled-coil proteins surrounding the Golgi complex, by a protein termed *TUG* (*T*ether containing a *U*BX domain for *G*LUT4). The N-terminal domain of TUG binds GLUT4 and other vesicle proteins while its C-terminal domain binds Golgi matrix proteins, preventing the movement of GSVs to the plasma membrane in the absence of insulin signaling.

(a) Total GLUT4



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Co

EXPERIMENTAL FIGURE 21-2 Insulin stimulation of fat cells induces translocation of GLUT4 from intracellular vesicles to the plasma membrane. (a) Cultured adipose cells were engineered to express a chimeric protein comprising GLUT4 with a green fluorescent protein (GFP) fused to its C-terminus; live cells were visualized with a confocal fluorescence microscope. In the absence of insulin, virtually all of the GLUT4 is in intracellular membranes that are not connected to the plasma membrane; there is little surface staining. Treatment with insulin triggers fusion of the GLUT4-containing membranes with the plasma membrane, thereby moving GLUT4 to the cell surface and enabling it to transport glucose from the blood into the cell. Arrows highlight GLUT4 present at the plasma membrane; *N* indicates the position of the nucleus. (b) In fat and muscle cells, insulin signaling acts at multiple steps to increase the level of GLUT4 at the plasma membrane. In resting cells, the majority of the GLUT4 protein is localized to specialized GLUT4 storage vesicles (GSVs); much of the GLUT4 is tethered to proteins surrounding the Golgi complex, termed *Golgi matrix proteins*, by the TUG protein and is thus unable to move to the cell surface. Binding of insulin to the insulin receptor leads to activation of a protease (step **1**) that cleaves the TUG protein, releasing GLUT4-containing vesicles from the Golgi matrix (step **2**). The N-terminal fragment of TUG remains bound to the GSVs and binds to the microtubule motor protein kinesin (step **6a**; see Chapter 18), enabling the GSVs to move along microtubules to the cell surface (step **6b**; *below*). Insulin also activates PKB (step **3**; see Figure 16-17). PKB then phosphorylates the Rab GAP protein AS160 (step **4**), inhibiting its ability to accelerate GTP hydrolysis by the GTP-binding protein Rab10 that is attached to a GSV by a covalently attached lipid. Rab10 then accumulates in its active GTP-bound state (step **5**) and directs the GLUT4 storage vesicles to move along microtubules and interact with target proteins on the cell surface (steps **6a** and **6b**). Finally, these GSVs fuse with the plasma membrane (step **7**). This step is catalyzed by another monomeric GTP-binding protein, RalA, when in its active GTP-bound state. PKB stimulates this membrane fusion event by phosphorylating and thus inactivating the RalA GAP protein (step **8**), allowing RalA to bind GTP (step **9**). The resultant increase in plasma membrane GLUT4 allows the cell to incorporate glucose from the extracellular fluids at a rate of about 10 times that of unstimulated cells (step **10**). Following removal of insulin, the plasma membrane GLUT4 is internalized by endocytosis (step **11**) and eventually transported to GSVs (step **12**). Many other proteins, not shown here, participate in these signaling and vesicle budding and fusion events; some of these proteins are also activated downstream of the insulin receptor in pathways other than those requiring PKB. See J. S. Bogan, 2012, *Annu. Rev. Biochem.* **81**:507; A. Klip, T. E. McGraw, and D. E. James, 2019, *J. Biol. Chem.* **294**:11369.

[Part (a) republished with permission from American Society for Biochemistry and Molecular Biology, from C. Yu et al., 2007, "The Glucose Transporter 4-Regulating Protein TUG Is Essential for Highly Insulin-Responsive Glucose Uptake in 3T3-L1 Adipocytes," 2007, *J. Biol. Chem.* **282**:7710; permission conveyed through Copyright Clearance Center, Inc.]

Binding of insulin to its receptor triggers activation of a protease that catalyzes a site-specific endoproteolytic cleavage of TUG, separating the N-terminal GLUT4-binding segment from the C-terminus that is anchored to the Golgi matrix. (The signaling pathway by which this takes place is only now being completely identified.) The cleavage of TUG releases GLUT4 storage vesicles from the Golgi matrix and allows the released N-terminal TUG fragment to bind the microtubule motor protein kinesin (see Chapter 18). This initiates the microtubule-based movement of the GSVs to the plasma membrane.

Movement of the GSVs to the plasma membrane is also regulated by two monomeric GTP-binding proteins, RalA and Rab10. Recall that several monomeric GTP-binding proteins are essential for the budding of intracellular transport vesicles (e.g., Sar proteins; see Figures 14-7 and 14-9); others, the Rabs, are essential for vesicle fusion (see Figure 14-11). As with other monomeric GTP-binding proteins, RalA and Rab10 (see Figure 21-2b) switch between the inactive GDP-bound and active GTP-bound state. Rab10 is bound to a GSV by a covalently attached lipid and regulates the movement of GLUT4 vesicles to the plasma membrane; RalA regulates vesicle fusion with the plasma membrane.

Like other receptor tyrosine kinases, the insulin receptor activates the PI-3 kinase/PKB pathway (Figure 16-17). Active PKB, in turn, phosphorylates and by so doing inhibits the GTP → GDP hydrolysis activity of two GAP proteins termed *AS160* and *RGC*; phosphorylation also causes the binding of AS160 to the 14-3-3 protein (Figure 16-13) that sequesters it in the cytosol away from its target Rab proteins. In the basal (unstimulated) state, these GAPs inhibit RalA and Rab10 function by enhancing their rates of GTP hydrolysis, thus keeping the GLUT4 storage vesicles from moving along microtubules to and then fusing with the plasma membrane. Phosphorylation of both GAPs by PKB allows RalA and the Rab10 to accumulate in their active GTP-bound states, enabling them to facilitate multiple steps in the GLUT4 pathway, including transport of the GLUT4 storage vesicles along microtubules to the cell surface and fusion of these vesicles with the plasma membrane (see Figure 21-2b). The mechanisms by which RalA and Rab10 regulate GSV targeting and fusion are the subject of much current research.

As blood glucose level drops, insulin secretion and insulin blood levels drop, and insulin receptors are no longer activated as strongly. In fat and muscle cells, plasma-membrane GLUT4 becomes internalized by endocytosis. When insulin levels are low, GLUT4 diverges from other endocytosed cargo, and moves from endosomes to a segment of the *trans*-Golgi network, and from there to the GSVs where it is stored, thus lowering the level of cell-surface GLUT4 and thus of glucose import.

In the Liver, Insulin Inhibits Glucose Synthesis, Accelerates the Rate of Glycolysis, and Enhances Storage of Glucose as Glycogen

In Chapter 15, we learned that hepatocytes, the major cell type in the liver, are important regulators of glucose metabolism; these cells store excess glucose as glycogen. Insulin acts on the liver in several ways: by inhibiting gluconeogenesis (the metabolic pathway by which metabolites including pyruvate, oxaloacetate, and acetyl CoA are converted to glucose), by accelerating the rate of glycolysis, and by enhancing storage of glucose as glycogen. All of these actions have a common end point — reducing the generation by and release of glucose from hepatocytes and thus reducing the level of glucose in the blood.

Insulin Promotes Glycogen Synthesis from Glucose

Within minutes, insulin stimulation of hepatocytes enhances the conversion of glucose to glycogen; PKB, which is activated downstream of the insulin receptor, plays a crucial role in this process as well. Active PKB phosphorylates GSK3 (the same enzyme that functions in the Wnt and Hh pathways; see Chapter 16). In resting (non-insulin-stimulated) cells, GSK3 is catalytically active

and phosphorylates glycogen synthase, inhibiting its activity and thus blocking glycogen synthesis from glucose (see Figure 15-22). In contrast, in insulin-treated liver, when GSK3 becomes phosphorylated by PKB its catalytic activity is inhibited, so it cannot phosphorylate and thus inhibit glycogen synthase. Insulin-stimulated activation of PKB therefore results in rapid activation of glycogen synthase and glycogen synthesis, helping remove glucose from the circulation and storing it as glycogen in muscle and liver.

Insulin Accelerates the Rate of Glycolysis

As we learned in Chapter 12, phosphofructokinase-1, the enzyme that catalyzes the reaction of fructose 6-phosphate and ATP to form fructose 1,6-bisphosphate and ADP, is a key rate-limiting enzyme in glycolysis. Insulin stimulation results in an enhancement of the activity of phosphofructokinase-1, thus accelerating the catabolism of glucose. It does so in part by enhancing the concentration of fructose 2,6-bisphosphate, as fructose 2,6-bisphosphate allosterically activates phosphofructokinase-1 in liver cells (see Figure 12-4). Through a signal transduction pathway activated downstream of the insulin receptor, the activity of phosphofructokinase-2, the enzyme that forms fructose 2,6-bisphosphate from fructose 6-phosphate and ATP, is enhanced, while the activity of fructose bisphosphate phosphatase-2, the enzyme that hydrolyzes fructose 2,6-bisphosphate back to fructose 6-phosphate and P_i, is inhibited. The end result is enhancement of

phosphofructokinase-1 activity by fructose 2,6-bisphosphate, acceleration of glucose catabolism through the glycolytic pathway, and a reduction in blood glucose levels.

Insulin Inhibits Synthesis of Key Enzymes Essential for Gluconeogenesis

Insulin also acts on hepatocytes to inhibit glucose synthesis from smaller molecules (gluconeogenesis), such as lactate, pyruvate, and acetate (see Chapter 12) and to enhance glycogen synthesis from glucose. Many of these effects are manifest at the level of gene transcription because insulin signaling reduces the expression of genes that encode enzymes that catalyze key steps in gluconeogenesis. The net effect of all these actions is to lower blood glucose to the fasting concentration of about 5 mM while storing the excess glucose intracellularly as glycogen for future use.

If the blood glucose level falls below about 5 mM, for example, due to sudden muscular activity, the reduced insulin secretion from pancreatic β cells results in increased glucagon secretion from the adjacent pancreatic α cells. The secretion of glucagon into the blood quickly triggers the breakdown of glycogen (Figure 15-22) to increase blood glucose levels and thus to restore blood glucose to the normal range.

Unfortunately, these intricate and powerful control systems sometimes fail, causing serious, even life-threatening, diseases, mainly *diabetes mellitus.* In diabetes, the regulation of blood glucose is impaired, leading to persistent elevation of the blood glucose concentration (hyperglycemia) that, if left untreated, leads to the major complications noted above. Type 1 diabetes mellitus, common in children and young adults, is caused by an autoimmune process that destroys the insulin-producing β cells in the pancreas. Sometimes called *insulin-dependent diabetes,* this form of the disease is generally responsive to regulated, lifelong insulin injections and constant external monitoring of blood glucose levels.

Most adults in developed countries with diabetes mellitus have type 2, sometimes called *non-insulin-dependent diabetes*; this condition results from a decrease in the ability of muscle, fat, and liver cells to respond to insulin and from a loss of functional insulin-producing β cells as the body tries to compensate for an elevated glucose level by overproducing insulin. While the underlying causes of this form of the disease are not well understood, obesity is correlated with a huge increase in the incidence of type 2 diabetes. Obesity also contributes to the malfunction of adipocytes, the cells that store fatty acids as triglycerides. The resulting accumulation of lipids (particularly diacylglycerols and sphingolipids) in muscle and liver impairs insulin action in these tissues. Further identification of the signaling pathways that control energy metabolism is expected to

provide insight into the pathophysiology of diabetes, hopefully leading to new methods for its prevention and treatment. ■

KEY CONCEPTS OF SECTION 21.1

Regulating Blood Glucose Level

- The maintenance of normal blood glucose concentrations depends on the balance between two hormones, insulin and glucagon, which are made in adjacent pancreatic islet cells.
- Insulin acts to lower blood glucose when the level rises above the normal 5 mM, while glucagon raises blood glucose levels.
- Insulin secretion by β islet cells is triggered by an increase in glucose influx, which leads to an increase in glycolysis, a rise in ATP levels, a closing of K^+ channels, and an opening of Ca^{2+} channels.
- Insulin acts on fat and muscle cells to increase the rate of glucose import. Activation of insulin receptors triggers several signal transduction pathways that lead to movement of GLUT4 glucose transporters from internal GSVs to and fusion with the plasma membrane.
- Insulin acts on hepatocytes to inhibit glucose synthesis, accelerate the rate of glycolysis, and enhance storage of glucose as glycogen.

21.2 Integrating Cell Growth Signals with Nutrient and Energy Levels

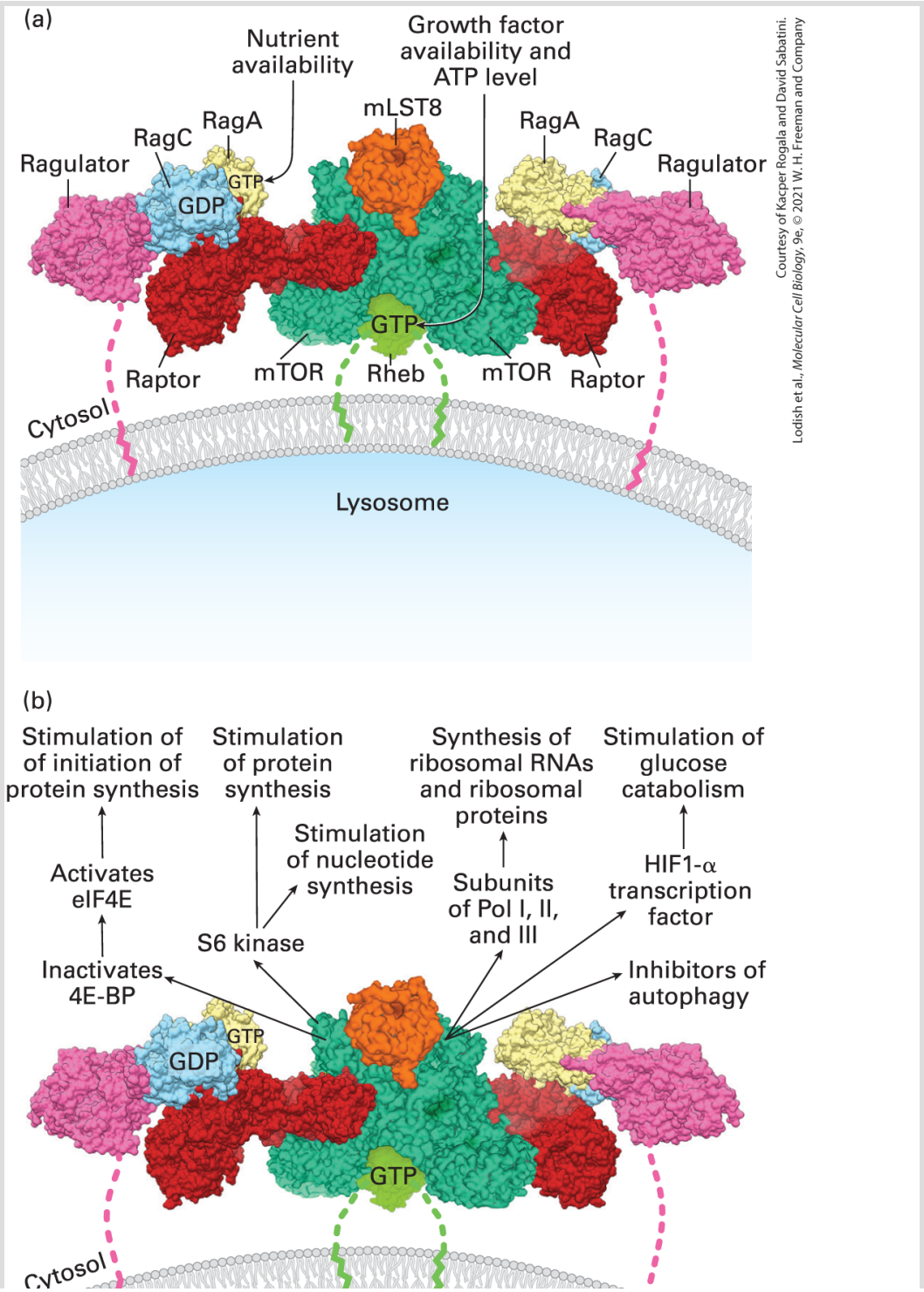
In the previous section, we saw how sensing and maintaining the level of a single molecule, glucose, is essential to the metabolism of vertebrate cells and organisms. However, for some processes cells must be able to sense multiple diverse molecules simultaneously and integrate these signaling inputs into a signal transduction pathway that allows activation of multiple downstream pathways. An important example is the decision of a cell to enter the cell division cycle (see Chapter 19). Initiation of cell division requires not only stimulation by specific growth factors that are on the surface of adjacent cells or in the surrounding medium (see Chapter 16) but also the presence of sufficient amino acids, nucleotides, and other small molecules in the cell's environment. Additionally, the rate of metabolism of glucose, fatty acids, and other energy sources must be sufficient to produce a level of ATP that can power synthesis of all of the proteins, nucleic acids, membranes, and other components necessary for the cell mass to double in preparation for its division.

In this section, we discuss the TOR pathway and its key component, the mTORC1 protein kinase, a large multisubunit complex that regulates cell growth by integrating inputs from multiple cytosolic sensor proteins. We will see in detail how the activity of this kinase

is regulated by levels of certain amino acids in the cytosol and within the lysosome, by the level of ATP, and by the levels of several intracellular signal transduction proteins, mainly kinases, activated downstream of growth-factor receptors. We also learn how the multiple substrates of active mTORC1 kinase in turn activate many signaling pathways necessary for cell growth and proliferation and repress autophagy, which otherwise would recycle these macromolecules back to their small molecule constituents. We also will see how improper activation of the mTORC1 signaling network contributes to human disease and learn that inhibitors of this kinase are used in the clinic.

The TOR pathway was discovered through research into the mechanism of action of rapamycin, an antibiotic produced by a strain of *Streptomyces* bacteria (see the chapter-opening figure). Rapamycin also proved to be useful for suppressing the immune response in organ transplant patients, initially, without knowledge of how it functioned. The *t*arget *o*f *r*apamycin *(TOR)*, a protein that is inhibited by rapamycin, was identified by isolating yeast mutants resistant to rapamycin inhibition of cell growth. Yeast TOR is a large (~2400 amino acid) protein kinase that regulates several cellular processes in yeast cells in response to nutritional status. Both yeast TOR and *m*etazoan *TOR (mTOR)* are incorporated into two multisubunit complexes, mTORC1 and mTORC2. Here we will focus on the better understood mTORC1 complex (Figure 21-3). The large mTORC1 complex is assembled on the cytosolic surface of the

lysosome. We first discuss the multiple signal transduction pathways mTORC1 activates that in turn stimulate cell growth and division, and then the mechanism of mTORC1 activation.



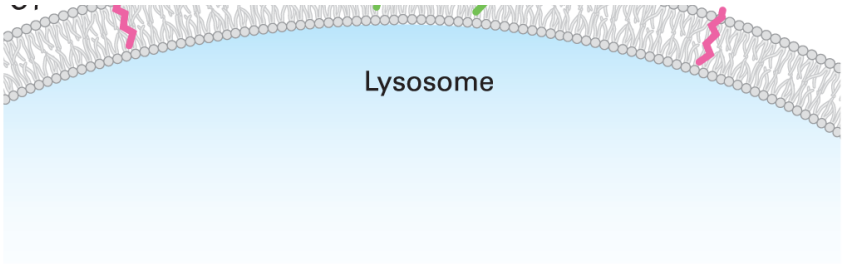


FIGURE 21-3 The mTORC1 complex. (a) Model of the active mTORC1 complex on the surface of the lysosome. The mTORC1 kinase consists of the kinase mTOR, Raptor (regulatory protein associated with mTOR), and many other proteins that regulate mTOR activity. Activation of the mTORC kinase requires its localization to the lysosome surface by the simultaneous conversion of two small GTP-binding proteins, Rheb and RagA, to their active GTP-bound state. Conversion of RagA to its active GTP-bound state occurs when there are sufficient levels of amino acids and other nutrients in the cytosol (Figure 21-4). Conversion of Rheb to its active GTP-bound state occurs downstream of activated kinases such as ERK and PKB that in turn are part of the signal transduction pathways activated downstream of many growth-factor receptors (Figure 21-5); conversion of Rheb to its active GTP-bound state also requires a high level of ATP. Thus mTORC1 integrates signals from many cellular states and in turn regulates many aspects of cell growth and metabolism. Note that both Rheb and the Ragulator protein that binds RagA (and its associated RagC) have lipid-modified terminal residues that tether these components of the supercomplex to the lysosomal membrane. (b) Multiple anabolic signal transduction pathways activated by the mTORC1 complex. Shown are some of the target proteins directly phosphorylated by active mTORC1 that stimulate RNA and protein synthesis and accelerate glycolysis, processes essential for cell growth. See text for details.

[Part (a) data from K. B. Rogala et al., 2019, *Science* **366**(6464):468–475; https://doi.org/10.1126/science.aay0166.]

The Active mTORC1 Complex Activates Many Anabolic Signal Transduction Pathways

In this section, we describe some of the target proteins directly phosphorylated by active mTORC1 that in turn stimulate ribosome, RNA, and protein synthesis, and accelerate the rate of glucose metabolism, all processes essential for cell growth (Figure 21-3b). In the next section, we detail how mTORC1 activity is regulated (Figure 21-3a).

Enhancing the Rate of Messenger RNA Translation and Protein Synthesis

mTORC1 stimulates the rate of messenger RNA translation, and thus overall protein synthesis, through two principal mechanisms: stimulating the production of ribosomes and other molecules necessary for protein synthesis and stimulating the rate of polypeptide chain initiation by ribosomes on mRNAs.

Recall that the first step in translation of a eukaryotic mRNA is binding of the eIF4 initiation complex to the 5′ cap on mRNAs via its eIF4E cap-binding subunit (see Figure 5-36). The concentration of active eIF4E, and thus the rate of polypeptide chain initiation, is down-regulated by a small family of homologous *eIF4E-binding proteins (4E-BPs)* that bind to and in so doing inhibit the interaction

of eIF4E with mRNA 5′ caps. 4E-BPs are directly phosphorylated by active mTORC1. When phosphorylated by mTORC1, 4E-BPs release eIF4E, stimulating the assembly of a translation initiation complex on the mRNA and thus initiation of protein translation. mTORC1 also phosphorylates and activates another protein kinase, S6K, that in turn phosphorylates and thus activates several translation initiation factors, leading to a further increase in the rate of protein synthesis.

Translation of a specific subset of mRNAs that have a string of pyrimidines in their 5′ untranslated regions (called *TOP* mRNAs for *t*ract of *o*ligo*p*yrimidine) is stimulated particularly strongly by phosphorylation of 4E-BP and S6K phosphorylation of another substrate of mTORC1, the RNA binding protein LARP1. TOP mRNAs encode ribosomal proteins and translation initiation and elongation factors; thus enhanced production of these proteins contributes to the production of additional ribosomes in the cell and further enhances the rate of mRNA translation. S6K also phosphorylates its namesake substrate, the S6 subunit of the small ribosomal subunit, but the effects of this phosphorylation on protein synthesis are not understood.

Enhancing the Synthesis of Ribosomal RNAs and tRNAs

Ribosomes constitute about 15 percent of the mass of an average metazoan cell. Active mTORC1 stimulates ribosome production by enhancing the synthesis not only of ribosomal proteins, as noted above, but also of ribosomal RNAs. In ways that are not fully understood, mTORC1 activates the RNA polymerase I transcription factor TIF-1A, stimulating transcription of the large rRNA precursor (see Figure 8-51). mTOR activates transcription by RNA polymerase III as well, by phosphorylating and thereby activating protein kinases that phosphorylate MAF1, a protein inhibitor of RNA polymerase III transcription. MAF1 phosphorylation causes it to be exported from the nucleus, relieving repression of RNA polymerase III transcription and enabling it to synthesize 5S ribosomal RNA as well as many tRNAs. When mTOR activity falls, MAF1 in the cytoplasm is rapidly de-phosphorylated and imported into the nucleus, where it represses transcription by RNA polymerase III.

In addition, mTOR activates two RNA polymerase II activators that stimulate transcription of ribosomal protein and translation factor genes. Finally, mTOR stimulates processing of the rRNA precursor (see Section 9.5) into mature ribosomal RNAs. As a consequence of phosphorylation of these several mTOR substrates, the synthesis and assembly of ribosomes as well as the synthesis of translation factors and tRNAs are greatly increased. Additionally, mTORC1 stimulates the synthesis of pyrimidines, a principal building block of RNA, from smaller molecules, again contributing to the enhanced production of ribosomal and other RNAs. Alternatively, when mTOR kinase activity is inhibited, these substrates become de-

phosphorylated, greatly decreasing the rate of protein synthesis and the production of ribosomes, translation factors, and tRNAs, thus slowing or halting cell growth.

In summary, active mTORC1 stimulates the overall rate of protein synthesis and thus of cell growth by activating transcription factors that enhance expression of ribosomal components, tRNAs, and translation factors. mTORC1 also activates two critical proteins that directly enhance mRNA translation.

Stimulating Glycolysis

Cell growth requires a ready supply of ATP, and activated mTORC1 stimulates glycolysis, the conversion of glucose to pyruvate with the concomitant production of ATP, by increasing expression of the glucose transporter GLUT1, which enhances the rate of glucose import into the cell (see Chapter 11). Activated mTORC1 also increases the production of many enzymes in the glycolytic pathway (see Chapter 12) through increasing the level of the transcription factor hypoxia-inducible factor 1α (HIF-1α). As we will learn in Section 21.4, the HIF-1α protein is rapidly degraded under normal oxygen concentrations but is stabilized during oxygen depletion, a regulation independent of mTORC1 activity. However, because of the selective increase in the rate of translation of HIF-1α mRNA induced by mTORC1, mTORC1 activation leads to an increase in HIF-1α levels even under normal levels of oxygen. HIF-1α in turn increases the rate of transcription of many genes encoding enzymes in the

glycolytic pathway, enhancing ATP levels even in the presence of oxygen.

Inhibiting Autophagy

In addition to enhancing the global rate of cellular protein synthesis and the production of ribosomes, tRNAs, and translation factors, mTORC1 regulates other processes involved in the response to low levels of nutrients, particularly macroautophagy (or simply **autophagy**). During starvation of one or more essential nutrients, when mTOR activity falls, cells degrade cytoplasmic constituents, including whole organelles, to supply energy and precursors for essential cellular processes (see Chapter 14). During this process, a large, double-membrane structure engulfs a region of cytoplasm to form an autophagosome, which then fuses with a lysosome where the entrapped proteins, lipids, and other macromolecules are degraded and their nutrient constituents recycled. Activated mTORC1 inhibits macroautophagy in growing cells when nutrients are plentiful, in part by phosphorylation and inhibition of the autophagy-initiating kinase Unc-51-like kinase-1 (abbreviated *ULK1*). Many of the details of this and other actions of mTORC1 are only now being discovered.

mTORC1 Kinase Activation Requires Amino Acids, a High ATP:AMP Ratio,

and Activation of Signal Transduction Pathways Downstream of Growth-Factor Receptors

Having learned how activated mTORC1 stimulates many aspects of macromolecular synthesis and cell growth, we turn now to its mechanism of activation: how mTORC1 kinase activation simultaneously requires high levels of cytosolic amino acids, a high ATP:AMP ratio, and signals from growth-factor receptors. This process involves the assembly on the cytosolic surface of lysosomes of the huge multisubunit mTORC1 complex depicted in Figure 21-3a.

Activation by Cytosolic Amino Acids

A high level of all 20 amino acids is obviously necessary for protein synthesis to proceed, but somehow vertebrate evolution has selected the levels of only three amino acids, leucine, arginine, and the methionine metabolite *S*-adenosylmethionine (SAM), as the ones that regulate mTORC1 kinase activity (Figure 21-4). (Flies and yeasts utilize different sets of amino acids and nutrients.) The levels of these amino acids in the cytosol regulate a pair of small GTP-binding proteins, RagA and RagC, which are always anchored to a lysosome surface by a lipid attached to Ragulator, a Rag-binding protein.

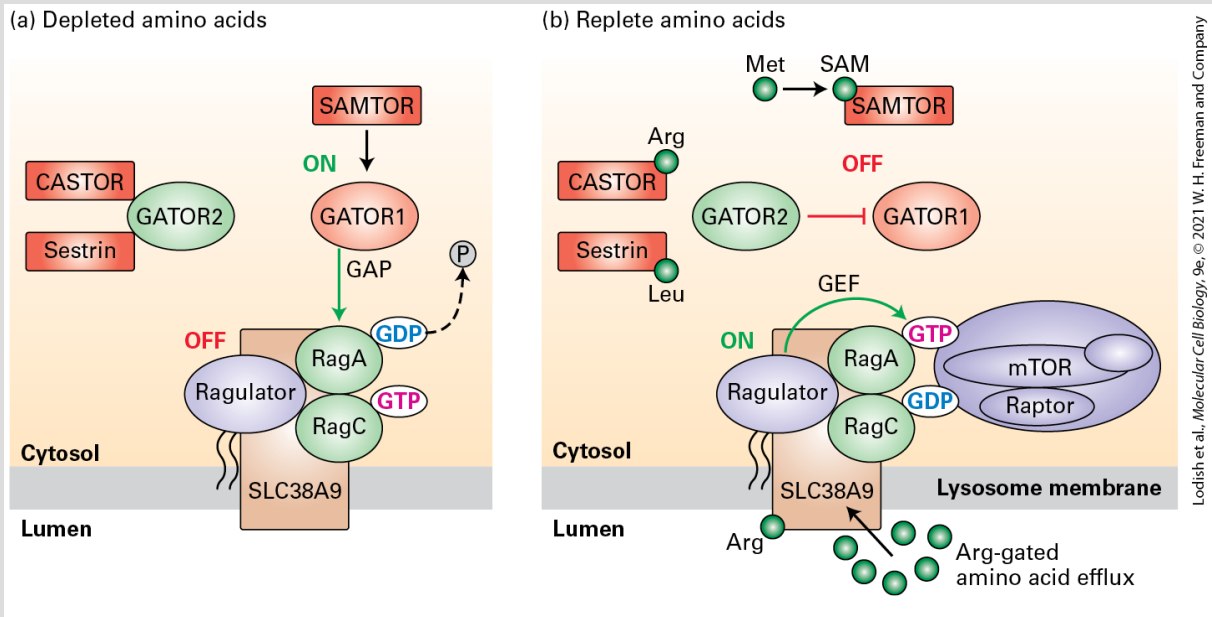


FIGURE 21-4 Model of the activation of the RagA regulator of the mTORC1 complex by three cytosolic amino acids. RagA and RagC are bound to the Ragulator protein that is tethered to the cytosolic surface of a lysosome membrane by a long covalently attached lipid. Like all small GTP-binding proteins, RagA cycles between an active state with bound GTP and an inactive state with bound GDP. RagA must be in its GTP-bound state (and oddly, RagC in its GDP state) to recruit the core mTORC1 complex to the lysosome surface where it can become activated by other signals downstream of growth-factor receptors. GATOR1 is the GAP that converts active RagA·GTP to inactive RagA·GDP and its activity is regulated by the level of three amino acids, leucine, arginine, and a metabolite of methionine, *S*-adenosylmethionine (SAM). (a) Under conditions of amino acid depletion, GATOR1 is active and promotes the conversion of RagA·GTP to RagA·GDP, thus generating RagA/C heterodimers that cannot bind mTORC1. GATOR2, a protein that binds to and inhibits GATOR1 GAP activity, is constrained from doing so by binding to the unoccupied cytosolic leucine receptor Sestrin and the unoccupied arginine receptor CASTOR. SAMTOR, the cytosolic receptor for SAM, when not bound to SAM binds to and activates GATOR1 activity, further favoring the conversion of RagA·GTP to RagA·GDP. (b) In the presence of sufficient arginine, leucine, and SAM, the GAP activity of GATOR1 becomes inactivated and RagA accumulates in the active GTP-bound state. Under these amino acid–replete conditions, leucine, arginine, and methionine-derived SAM directly bind to their sensors, Sestrin, CASTOR, and SAMTOR, respectively. This triggers inhibition of GATOR1 GAP activity: release of Sestrin and CASTOR from GATOR2 allows GATOR2 to bind to and inhibit GATOR1 through an unknown mechanism, and SAMTOR is unable to activate GATOR1 GAP activity. Efflux of

arginine from the lumen of the lysosome into the cytosol, catalyzed by the transport protein SLC38A9, also contributes to accumulation of RagA in its GTP bound active state. RagA·GTP in turn binds mTOR kinase to the lysosome surface, but activation of mTORC1 kinase activity requires additional signals downstream of growth-factor receptors (Figure 21-5). See A. J. Valvezan and B. D. Manning, 2019, *Nat. Metab.* **1**:321.

The small GTP-binding protein RagA is a key regulator of mTORC1 activity; it enables mTORC1 kinase activity to become activated only when the levels of leucine, arginine, and SAM in the cytosol are sufficiently high. (The associated GTP-binding protein RagC has a minor role in mTORC regulation, and oddly its active state is one with a bound GDP, not GTP.) Like all small GTP-binding proteins, RagA shuttles between its active GTP-bound and inactive GDP bound state; RagA must be in its GTP-bound state to recruit the core mTOR kinase complex to the lysosome surface where it can become activated by other signals downstream of growth-factor receptors. The RagA GEF, the lysosome membrane protein SLC38A9 that also functions as an arginine regulated transport protein for several amino acids, continuously triggers the dissociation of the GDP from RagA·GDP, promoting the binding of GTP to RagA to form RagA·GTP. Rag regulation focuses on GATOR1, the GAP that converts active RagA·GTP to the inactive RagA·GDP.

GATOR1 is active when any of the three crucial amino acids, leucine, arginine, and SAM, are depleted in the cytosol, putting RagA in the GDP state and blocking recruitment of mTOR to the lysosome (Figure 21-4a). Binding of another protein complex, termed *GATOR2*,

to GATOR1 inhibits the GATOR1 GAP activity. However, if the levels of arginine or leucine are low, the binding of GATOR2 to GATOR1 is blocked, probably because GATOR2 becomes bound instead to the unoccupied arginine sensor *CASTOR (c*ellular *a*rginine *s*ensor for m*TOR*C1) or the unoccupied leucine sensor Sestrin2. When present at sufficient levels, arginine and leucine bind to defined sites on CASTOR and Sestrin2, respectively. This binding triggers conformational changes that cause CASTOR and Sestrin2 to dissociate from their inhibitory interactions with GATOR2, allowing GATOR2 to bind to GATOR1 and inhibit GATOR1 GAP activity; this promotes accumulation of RagA in its active GTP-bound state (Figure 21-4b), capable of recruiting mTOR to the complex on the surface of the lysosome.

SAMTOR, the SAM-binding protein, works differently. In the absence of SAM, SAMTOR binds directly to GATOR1 and enhances its GAP activity; binding of SAM to SAMTOR prevents its binding to GATOR1 and reduces GATOR1 GAP activity. Efflux of arginine from the lumen of the lysosome into the cytosol, catalyzed by the transport protein SLC38A9, also contributes to accumulation of RagA in its GTP bound active state.

Thus RagA accumulates in the active GTP-bound state only when there are sufficient levels of leucine, arginine, and SAM in the cytosol. RagA·GTP in turn recruits the inactive mTOR from the cytosol to the lysosome surface, but activation of mTORC1 kinase

activity requires additional signals downstream of growth factor receptors, as we see in the next section.

The mTORC1 Complex Is Activated by a High ATP:AMP Ratio and by ERK and PKB Kinases Activated Downstream of Growth Factor Receptors

Having sufficient levels of amino acids, nucleotides, and other small molecules is necessary for a cell to undergo division, but it is not sufficient. The key to understanding how activation of growth factor receptors also promotes the activation of mTORC1 is another small GTP-binding protein, Rheb (Figure 21-5). Like RagA, Rheb switches between its active GTP-bound and inactive GDP-bound state. Rheb is tethered to the surface of the lysosome by a covalently attached lipid, and it must be in its GTP-bound state to bind to and activate the mTOR complex that has been brought to the lysosome surface by RagA·GTP.

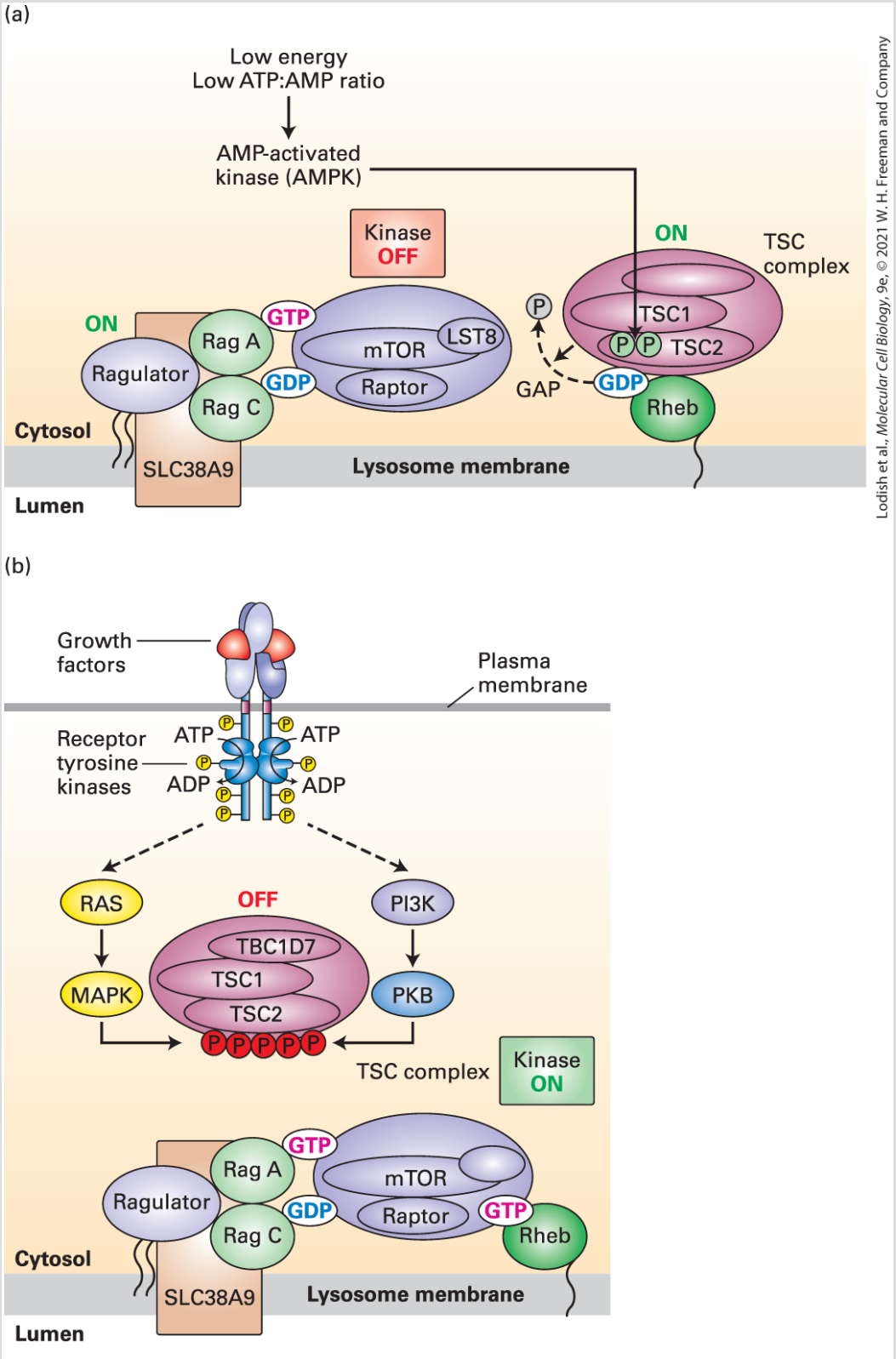


FIGURE 21-5 Model of the activation of the mTORC1 complex by a high ATP:AMP ratio and by kinases activated downstream of growth factor

receptors. The presence of sufficient levels of three amino acids in the cytosol is required to bind the mTOR complex to the lysosome surface (see Figure 21-4). Full activation of mTORC1 kinase activity requires the conversion of the inactive Rheb·GDP protein to its active Rheb·GTP state by kinases activated downstream of growth factor receptors and by a high ATP:AMP ratio. The TSC GAP complex is the only well-established regulator of Rheb identified to date. (a) In absence of growth factors or in a state of low ATP, the mTORC1 kinase is anchored to the lysosome surface (Figure 12-4) but is catalytically inactive because Rheb accumulates in the GDP-bound state and is unable to activate mTORC1 kinase activity. In the absence of growth factor signaling the TSC protein complex, comprising TSC1 and TSC2, localizes to the lysosome surface. TSC possesses highly specific GTPase-activating protein (GAP) activity toward Rheb and thus converts Rheb to the inactive GDP-bound state. A low-energy state, such as during glucose deprivation, results in an increase in the level of cytosolic AMP, which activates the AMP-activated protein kinase AMPK. AMPK phosphorylates TSC2 at a site that further activates the GAP activity of the TSC complex toward Rheb·GTP. Thus a low ATP:AMP ratio also causes Rheb to accumulate in the inactive GDP-bound state and prevents binding and activation of the mTORC1 complex. (b) Addition of growth factors causes Rheb to accumulate in the GTP-bound state, activating mTORC1 kinase activity. TSC2 is a heavily phosphorylated protein, and several protein kinases activated downstream of different growth factor receptors phosphorylate sites on the TSC2 subunit of the TSC complex, inactivating TSC GAP activity and causing dissociation of the TSC complex from lysosomal-bound Rheb. Two such kinases are depicted here, protein kinase B (PKB), activated by the PI-3 kinase pathway (Figure 16-17) and MAP kinase (MAPK), activated by the Ras/MAP kinase pathway (Figure 16-13). Inhibition of TSC GAP activity by these phosphorylation events causes Rheb to accumulate in its active GTP-bound state, able to bind a subunit of the mTORC1 complex already linked to the lysosome membrane, and by so doing activating mTORC1 kinase activity and triggering multiple downstream signal transduction pathways. See A. J. Valvezan and B. D. Manning, 2019, *Nat. Metab.* **1**:321.

As with RagA, GEF proteins, whose identity is unknown, continuously trigger dissociation of the GDP from Rheb·GDP and allow formation of Rheb·GTP and thus activation of mTORC1. Regulation is focused on the GAP that converts the active Rheb·GTP to inactive Rheb·GDP, a heterodimer composed of subunits TSC1 and TSC2, named for their involvement in the medical syndrome *t*uberous *s*clerosis *c*omplex, as discussed below. This TSC complex converts Rheb to its GDP-bound conformation, which does not bind mTOR or activate mTORC1 kinase activity.

TSC2 is phosphorylated by multiple protein kinases that regulate its activity and thus the ability of mTORC1 to accumulate in its active state on the lysosomal surface (Figures 21-5a and b). As we learned in Chapter 16, signaling from many cell-surface growth factor receptor tyrosine kinases leads to activation of the Ras MAPK pathway and MAP kinase (MAPK) in particular. Protein kinase B (PKB) is also activated as part of the PI-3 kinase signaling pathway downstream of many receptor tyrosine kinases. Both MAPK and PKB phosphorylate TSC1/TSC2 at sites that inhibit its GAP activity, causing an increase in accumulation of Rheb·GTP and activation of mTORC1 kinase activity (Figure 21-5b). This type of regulation through cell-surface receptors links the control of cell growth to developmental processes controlled by cell-cell interactions.

Cells also require a high concentration of ATP to undergo proliferation. A reduction in ATP levels cause an increase in ADP

and AMP levels; an accumulation of AMP triggers activation of AMP-activated protein kinase (AMPK). Active AMPK phosphorylates distinct sites on TSC2 that enhance its GAP activity and thus turn off Rheb·GTP–mTORC1 signaling under conditions of cellular energy depletion (Figure 21-5a). AMPK is inactive under conditions of sufficient ATP, but hypoxia and other cellular stresses also activate the TSC1/TSC2 Rheb-GAP.

In summary, activation of MAPK, PKB, and other kinases downstream of growth factor receptors, together with a high cytosolic level of ATP, cause Rheb to accumulate in its active GTP state and to bind to and activate the mTORC1 kinase already at the surface of the lysosome, triggering the multiple downstream anabolic processes discussed above.

Genes encoding components of the mTOR pathway are mutated in many human cancers, resulting in cell growth in the absence of normal growth signals. TSC1 and TSC2 (see Figure 21-5) were initially identified because one or the other of these proteins is mutated in a rare human genetic syndrome: *t*uberous *s*clerosis *c*omplex. Patients with this disorder develop benign tumors in multiple tissues. The disease results because inactivation of either TSC1 or TSC2 eliminates the Rheb-GAP activity of the TSC1/TSC2 heterodimer, resulting in an abnormally high and unregulated level of Rheb·GTP and in the resulting high, unregulated activity of mTORC1. Mutations in components of cell-surface receptor signal

transduction pathways that lead to inhibition of TSC1/TSC2 Rheb-GAP activity are also common in human tumors and contribute to cell growth and replication in the absence of normal signals for growth and proliferation (see Chapter 25).

High mTORC1 protein kinase activity in tumors correlates with a poor clinical prognosis. Consequently, mTOR inhibitors are currently in clinical trials to test their effectiveness for treating cancers in conjunction with other modes of therapy. Rapamycin and other structurally related mTOR inhibitors are potent suppressors of the immune response because they inhibit activation and replication of T lymphocytes in response to foreign antigens (see Chapter 24). Several viruses encode proteins that activate mTOR early after viral infection. The resulting stimulation of translation has an obvious selective advantage for these cellular parasites. ■

KEY CONCEPTS OF SECTION 21.2

Integrating Cell Growth Signals with Nutrient and Energy Levels

- The mTORC1 kinase complex is assembled on the cytosolic surface of the lysosome and consists of the kinase mTOR, the small GTP-binding proteins RagA and Rheb, and many other proteins that regulate mTOR kinase activity (see Figure 21-3).
- The active mTORC1 kinase activates many anabolic signal transduction pathways, including mRNA translation, synthesis of ribosomal RNAs and proteins and tRNAs, and glycolysis, and inhibits autophagy (see Figure 21-3b).
- mTORC1 kinase activation requires the presence in the cytosol of the amino acids arginine, leucine, and *S*-adenosylmethionine. In the absence of any these amino acids GATOR1, the GAP for RagA, causes RagA to accumulate in the GDP-bound state; the

RagA·GDP complex is unable to bring the mTOR kinase to the lysosome surface where it can become activated.

- Binding of these three amino acids to their respective cytosolic sensor proteins triggers an inhibition of the GATOR GAP activity, allowing RagA to accumulate in the active GTP-bound state where it tethers the mTOR kinase to the surface of the lysosome (see Figure 21-4).
- Activation of the mTOR kinase also requires a high ATP:AMP ratio and active ERK and PKB, kinases that are activated downstream of receptor tyrosine kinases. These signaling pathways act on the TSC complex, a GAP for another small GTP-binding protein called *Rheb*, to allow Rheb to accumulate in the active GTP-bound state.
- Accumulation of Rheb·GTP triggers activation of the mTOR kinase that has been tethered to the lysosome surface by RagA·GTP (see Figure 21-5).

21.3 Responding to Changes in the Levels of Cholesterol and Unsaturated Fatty Acids

All cells need to control the amount and composition of membrane lipids. We learned in Chapter 10 that cells must maintain an appropriate fluidity of their membranes and that this is determined by both the amount of cholesterol or other steroids in each cellular membrane and the percentage of unsaturated fatty acids in membrane phospholipids. Indeed, a cell would soon face a crisis if it did not have enough phospholipids to make adequate amounts of membranes or had so much cholesterol that large crystals formed and damaged cellular structures (see Chapter 10). Here we describe how cells maintain cholesterol and fatty acid homeostasis by regulating both the synthesis of enzymes that catalyze key steps in their biosynthesis and the import of exogenous cholesterol from the extracellular fluid. Critical features of this regulation are the sensing of these metabolites by proteins in the endoplasmic reticulum.

Fatty Acid and Cholesterol Biosynthesis as Well as Cholesterol Import Are Regulated at the Level of Gene Transcription

The transcription of genes encoding many enzymes in the cholesterol biosynthetic pathway (see Figure 10-26) is subject to feedback regulation. When cellular cholesterol levels are adequate, cholesterol biosynthesis is suppressed so that excess unesterified cholesterol, which can be toxic, does not accumulate. When cellular cholesterol levels fall too low, threatening the integrity of many cell membranes, biosynthesis is increased. Synthesis of one of the enzymes in the cholesterol biosynthetic pathway, HMG-CoA reductase, is tightly regulated as is the rate-controlling enzyme in cholesterol biosynthesis.

Feedback regulation also controls the import of extracellular cholesterol. As we learned in Chapter 14, lipoproteins are lipid-filled particles that transport hydrophobic molecules through the circulatory system. In humans, the most abundant transporter of cholesterol, in the form of cholesteryl esters, is low-density lipoprotein (LDL; see Figure 14-27). LDL delivers its cholesterol to cells by binding to cell-surface LDL receptors, which mediate the endocytic uptake of LDL particles and their subsequent disassembly in lysosomes, from where the cholesterol enters the cellular pool of cholesterol (see Figure 14-29). The transcription of LDL receptors is transcriptionally regulated to maintain cellular cholesterol homeostasis, and similar systems are used to control unsaturated fatty acid levels.

Sterol Regulatory Elements (SREs)

The levels of sterols such as cholesterol, as well as unsaturated fatty acids, regulate the levels of expression of a number of genes in a cell. The promoters of most cholesterol-sensitive and unsaturated fatty acid–sensitive genes, including HMG-CoA and the LDL receptor, contain one or more 10-base-pair *sterol regulatory elements (SREs)*, or SRE half-sites, in their promoters. (These SREs differ from the *s*erum *r*esponse *e*lements that control many early response genes, discussed in Section 16.2.) The interaction of lipid-regulated transcription factors called **SRE-binding proteins** (**SREBPs**) with these response elements modulates the expression of the target genes.

Mammals express three isoforms of SREBP: SREBP-1a and SREBP-1c, which are generated from alternatively spliced RNAs produced from the same gene; and SREBP-2, which is encoded by a different gene. Together, these transcription factors control expression of proteins that regulate availability not only of cholesterol, but also of fatty acids and the triglycerides and phospholipids made from fatty acids. In mammalian cells, SREBP-1a and SREBP-1c exert a greater influence on fatty acid metabolism than on cholesterol metabolism, whereas the reverse is the case for SREBP-2.

How do cells sense how much cholesterol and unsaturated fatty acids they have, and how are these signals used to control the level of SREBPs in the nucleus and thus gene expression? The SREBP-mediated pathway begins in the membranes of the endoplasmic

reticulum (ER) and includes several other proteins besides SREBP. In Chapter 16, we saw how *regulated intramembrane proteolysis* plays a key role in the Notch and EGF signaling pathways; here we see that *regulated intramembrane proteolysis* also plays an important role in this cellular response to altered cholesterol and fatty acid levels.

The Endoplasmic Reticulum SCAP Protein Senses the Level of Cellular Cholesterol

When cells have adequate levels of cholesterol, SREBP is found in the ER membrane complexed with SCAP (SREBP cleavage-activating protein), insig-1 (or its close homolog insig-2), and perhaps other proteins (Figure 21-6a). SREBP has three distinct domains: an N-terminal cytosolic domain, containing a basic helix-loop-helix (bHLH) DNA-binding motif (see Figure 8-25d), that functions as a transcription factor when cleaved from the rest of SREBP; a central membrane-anchoring domain containing two transmembrane α helices; and a C-terminal cytosolic-facing regulatory domain.

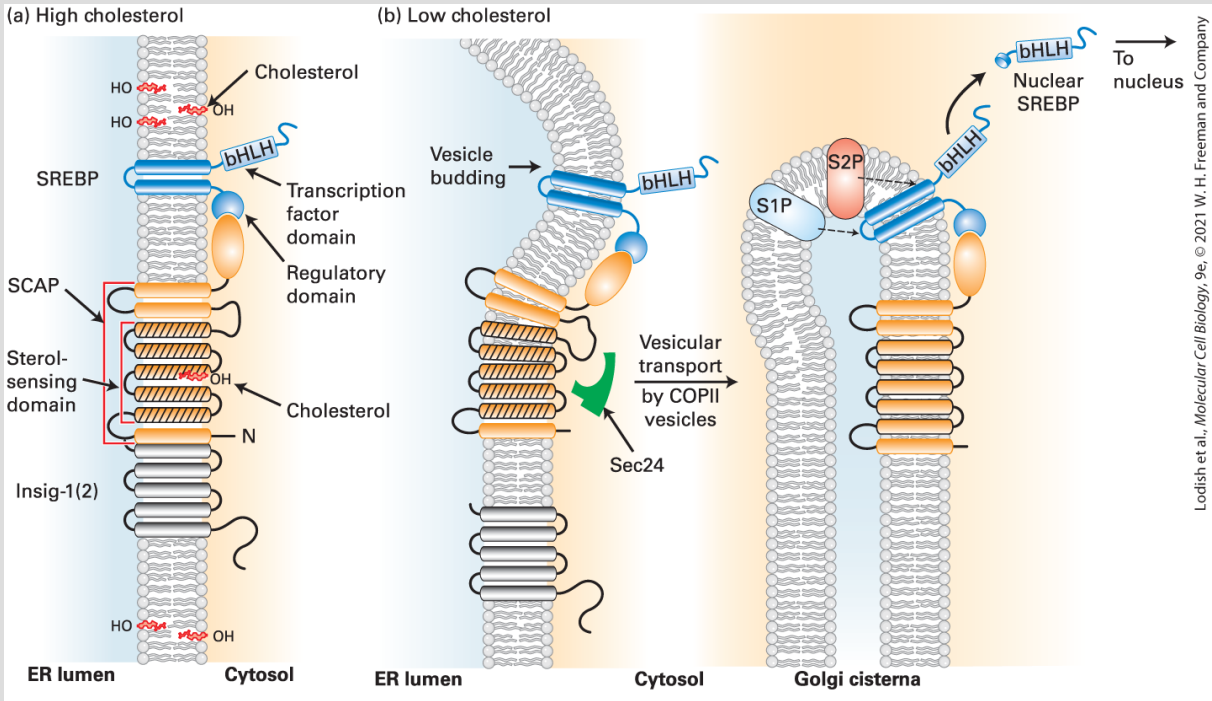


FIGURE 21-6 Cholesterol-sensitive control of SREBP activation. The cellular pool of cholesterol is monitored by the cholesterol-sensor SCAP, located in the ER membrane and stably bound via its C-terminus to the C-terminus of the membrane protein SREBP. Membrane-spanning helices 2–6 of SCAP (orange with black lines) form the sterol-sensing domain. (a) When cholesterol levels are high enough that ER cholesterol exceeds 5 percent of total ER membrane lipid molecules, cholesterol binds to the sterol-sensing domain in SCAP, triggering a conformational change that enables the SREBP/SCAP-cholesterol complex to bind to insig-1(2), anchoring the entire SCAP-SREBP complex in the ER membrane. (b) At low-cholesterol levels, cholesterol dissociates from the SCAP sterol-sensing domain, triggering a reverse conformational change that both dissociates SCAP from insig-1(2) and enables SCAP to bind to Sec24 (green), a subunit of the COPII complex (see Figure 14-9). As a consequence, the SCAP-SREBP complex enters COPII-coated vesicles that transport it to the Golgi. In the Golgi, the sequential cleavage of SREBP by the site 1 and site 2 proteases (S1P, S2P) releases the N-terminal bHLH transcription factor domain of SREBP into the cytosol, which then translocates to the nucleus. There the released bHLH domain of SREBP, called *nuclear SREBP (nSREBP)*, controls the transcription of genes containing sterol regulatory elements (SREs) in their promoters The SCAP protein recycles from the Golgi back to the ER (not shown). See M. S. Brown, A. Radhakrishnan, and J. L. Goldstein, 2018, *Annu. Rev. Biochem.* **87**:783–807; and R. A. DeBose-Boyd and J. Ye, 2018, *Trends Biochem. Sci.* **43**:358–368.

SCAP has eight transmembrane α helices and a large C-terminal cytosolic domain that binds to the regulatory domain of SREBP. Five of the transmembrane α helices in SCAP form a *sterol-sensing domain* similar to that in HMG-CoA reductase (see Figure 21-6a; also see Section 10.3). When cholesterol binds to the sterol-sensing domain in SCAP, SCAP assumes a conformation such that it binds to insig-1(2). When the SCAP-cholesterol complex is tightly bound to insig-1(2), it is in a conformation that blocks its binding to the Sec24 subunit of the COPII vesicle coat complex (see Figure 14-9). This prevents incorporation of the SCAP-SREBP complex into COPII ER-to-Golgi transport vesicles (see Chapter 14). As a consequence the SCAP-SREBP complex is retained in the ER, and thus the transcription factor in SREBP's N-terminal domain cannot enter the nucleus and regulate gene expression. This occurs when cholesterol levels in the ER membrane exceed 5 percent of total ER membrane lipids.

Regulated Intramembrane Proteolysis of SREBP in the Golgi Releases a bHLH Transcription Factor That Acts to Maintain Appropriate Phospholipid and Cholesterol Levels

SCAP releases its bound cholesterol when cholesterol levels in the ER membrane drop to less than 5 percent of ER lipids, a value that reflects normal cellular cholesterol levels. Consequently, Insig-1(2) can no longer bind to the cholesterol-free SCAP, and the SCAP-SREBP complex moves from the ER to the Golgi complex via COPII vesicles (Figure 21-6b). SREBP is then cleaved sequentially at two sites by two proteases that are resident in the Golgi membrane. Site-1 protease (S1P) cleaves SREBP in its luminal loop, separating the protein into two membrane-bound halves. S1P-mediated cleavage renders the N-terminal half of SREBP a substrate for the Site-2 protease (S2P), which releases the N-terminal bHLH-containing domain of SREBP into the cytosol by cleaving the intermediate near the cytosol-membrane boundary, an example of *regulated intramembrane proteolysis*. This protein, also called *nSREBP (nuclear SREBP)*, is rapidly translocated into the nucleus, where it activates transcription of genes containing sterol regulatory elements (SREs) in their promoters.

SREBP-2 primarily activates transcription of genes required for cholesterol synthesis and uptake such as those encoding the LDL receptor and HMG-CoA reductase. Thus a reduction in cellular cholesterol, by activating the *Insig-1(2)/SCAP/SREBP pathway*, triggers expression of genes encoding proteins that both import cholesterol into the cell (the LDL receptor) and synthesize cholesterol from small precursor molecules (HMG-CoA reductase).

SREBP-1a and SREBP-1c preferentially activate transcription of genes required for fatty acid and triglyceride synthesis. The level of unsaturated (but not saturated) fatty acids also regulates cleavage of SREBP-1, but by a slightly different mechanism. In cells deprived of unsaturated fatty acids, Insig-1 becomes a substrate for polyubiquitination, extraction from the ER membrane into the cytosol by the process known as *ERAD (ER-a*ssociated *d*egradation; end of Section 13.3), and subsequent degradation by proteasomes. In the absence of Insig-1, SCAP-1 assumes a conformation enabling it to bind the Sec24 subunit of COPII vesicles, similar to the low-cholesterol situation diagrammed in Figure 21-6b. This initiates transport of SCAP-bound SREBP from the ER to the Golgi, where it undergoes proteolytic cleavage by S1P and S2P that releases its bHLH-containing SREBP-1 transcription factor domain, enabling it to activate expression of genes required for synthesis of fatty acids and triglycerides. In all cases, after cleavage of SREBP in the Golgi, SCAP apparently recycles back to the ER, where it can interact with insig-1(2) and another intact SREBP molecule.

High-level transcription of SRE-controlled genes requires the ongoing generation of new nSREBP because this transcription factor is degraded fairly rapidly by the ubiquitin-mediated proteasomal pathway (see Chapter 3). The rapid generation and degradation of nSREBP also helps cells respond quickly to changes in levels of intracellular cholesterol and fatty acids.

Under some circumstances (e.g., during cell growth), cells need an increased supply of all essential membrane lipids and their fatty acid precursors, processes that require coordinate regulation of many metabolic pathways. In the liver, mTOR, after it becomes activated following stimulation by many growth factor receptors (Section 21.2), directly activates transcription of the *SREB-1c* gene. This in turn results in increased transcription of all of its downstream target genes that encode proteins catalyzing steps in fatty acid synthesis.

KEY CONCEPTS OF SECTION 21.3

Responding to Changes in the Levels of Cholesterol and Unsaturated Fatty Acids

- Vertebrate cells maintain a constant level of cholesterol in their membranes.
- The cellular pool of cholesterol is monitored in the ER membrane by the SCAP protein, which is in a stable complex with full-length SREBP.
- At high-cholesterol levels, cholesterol is bound to the SREBP-SCAP complex, which in turn is bound to the ER protein Insig-1(2), anchoring the entire SREBP-SCAP-cholesterol complex in the ER membrane and preventing its transport to the Golgi and subsequent release of its active nSREBP domain, a transcription factor.
- At low-cholesterol levels, cholesterol dissociates from the SREBP-SCAP complex, triggering conformational changes that lead to the dissociation of SCAP from Insig-1(2), the binding of SREBP and SCAP to the COPII coat, and transport of the SREBP-SCAP complex to the Golgi.
- In the Golgi, cleavage of SREBP by two proteases releases the N-terminal transcription factor, bHLH domain of SREBP (nSREBP), which translocates to the nucleus and activates expression of genes containing sterol regulatory elements (SREs) in their promoters. Among these genes are those that encode proteins that import cholesterol into the cell and enzymes that catalyze key steps in the biosynthesis of cholesterol.
- SREBP-2 primarily regulates cholesterol import and synthesis while SREBP-1a and SREBP-1c primarily regulates unesterified fatty acid synthesis.

21.4 Responding to Low Oxygen

Oxygen is essential for metazoan life, and in Chapter 12 we learned how mitochondria produce ATP from ADP and P_i by oxidizing metabolites derived from glucose, fatty acids, and amino acids. Many organs, such as the brain and heart, cannot use anaerobic glycolysis to produce sufficient ATP and thus are completely dependent on production of ATP through oxidative phosphorylation. In this section, we describe how cells sense drops in the level of oxygen in their environment and respond to this stress, as happens when we ascend to a high altitude or during disease states such as cardiopulmonary failure, where the heart or lungs cannot generate sufficient levels of oxygen-bound hemoglobin in the bloodstream. We will learn how, in metazoans, one transcription factor, Hif-1α coordinates the many cellular responses to low oxygen, how Hif-1α is produced but rapidly degraded by cells under ambient oxygen levels, and how as oxygen levels decrease Hif-1α becomes progressively more stable. Hif-1α activates transcription of hypoxia-induced genes by binding to *h*ypoxia-*r*esponsive *e*lements (HRE) in the promoter or enhancer of oxygen-responsive genes. Hif-1α is found only in metazoan animals, but recent work has uncovered a different family of oxygen-sensing transcription factors that are conserved in both animals and plants. Like Hif-1α these proteins are degraded at ambient oxygen, but by a different mechanism.

Induction of the Erythropoietin Gene at Low Oxygen Levels

Red blood cells transport oxygen from the lungs to all body cells, and maintaining a precise level of red blood cells of about 40–50 percent of blood volume is important for normal body functions. We learned in Chapter 16 (Figure 16-18a) that the cytokine erythropoietin (Epo) is the principal hormone that stimulates red blood cell formation, and that Epo is synthesized by certain cells in the liver and kidney as oxygen levels drop. At one time these cells were thought to be one of the few in the body that respond to low oxygen.

In the early 1990s, the key discovery was made that when liver cells were subjected to low oxygen, a then-novel protein bound to the HRE, with a core sequence 5′-CGTG-3′, in the 3′ enhancer segment of the erythropoietin gene. Characterization of the protein, now called *h*ypoxia-*i*nducible *f*actor 1 alpha or simply **Hif-1α** showed that it was present only in cells cultured in a low-oxygen environment and that Hif-1α bound to HREs as a homodimer with a second protein, called *Hif-1β*, whose level was not regulated by oxygen (Figure 21-7a).

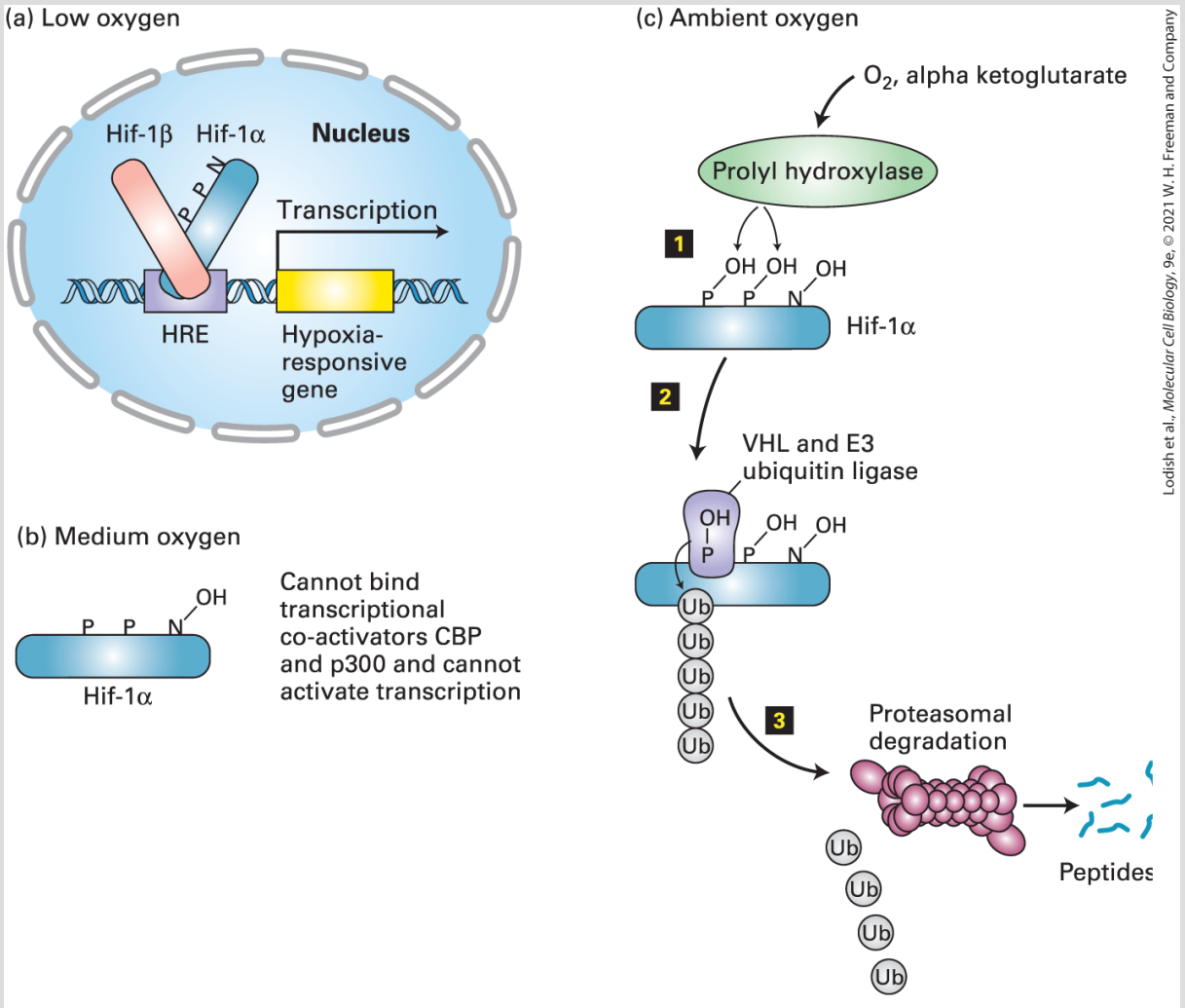


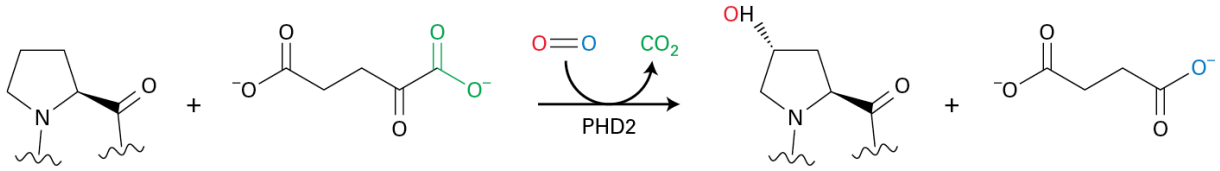
FIGURE 21-7 Degradation of the Hif-1α transcription factor at ambient oxygen levels. (a) Low (0.5 percent) oxygen. Hif-1α is stable and functional. It forms a heterodimer with Hif-1β, binds to HRE elements in the promoters or enhancers of many hypoxia-responsive genes, and induces their transcription. (b) Medium (1–2 percent) oxygen. The asparaginyl hydroxylase FIH is able to bind sufficient oxygen to become catalytically active and transfer an OH group to an asparagine residue on the Hif-1α protein. Hydroxylation blocks the ability of Hif-1α to bind the transcriptional co-activators CBP and p300 and thus activate gene transcription. (c) High (above 4 percent) oxygen. (step 1) The prolyl hydroxylase PHD2 is able to bind sufficient oxygen to become activated and transfer an OH group to two proline residues in Hif-1α protein. (step 2) Each of these hydroxyproline residues form a part of a binding site for the VHL subunit of an E3 ubiquitin ligase. (step 3) The E3 ligase then adds a polyubiquitin chain that triggers immediate degradation of the Hif-1α protein by the proteasome. See text for details.

Oxygen Sensing and Regulated Hif-1α Expression Is a Property of All Nucleated Mammalian Cells

Soon after this discovery, it was shown that the Hif-1α gene is transcribed into mRNA in all body cells regardless of oxygen levels, but that the Hif-1α protein is present only when oxygen levels are very low. As an example, in cultured HeLa cells, the levels of Hif-1α protein and Hif-1α DNA-binding activity increased exponentially as the cellular oxygen concentration decreased, with maximum values at 0.5 percent oxygen and half-maximal values at 1.5 to 2 percent oxygen, compared to ~20 percent oxygen at ambient levels. Indeed, it was soon demonstrated that Hif-1α expression triggered synthesis of *v*ascular *e*ndothelial *g*rowth *f*actor (VEGF), a secreted hormone that stimulates formation of capillaries, tiny blood vessels that connect arteries and veins and increase oxygen delivery to the surrounding tissues. HRE sequences were also found in the enhancers of many genes encoding glycolytic enzymes, and Hif-1α stabilization triggered expression of these genes. In particular, genes encoding several proteins that catalyze regulated and rate-limiting steps in the glycolytic pathway [phosphofructokinase, pyruvate kinase, and lactate dehydrogenase (see Section 12.1)] are strongly induced by Hif-1α at low oxygen levels and promote ATP synthesis in an oxygen-independent manner.

Hif-1α Function and Stability Are Blocked at Ambient Oxygen Levels

Hif-1α protein is synthesized at almost the same rate at low and ambient oxygen levels, but it is rapidly degraded in ambient oxygen. The key to understanding how this happens came with the discovery of a set of proline and asparagine hydroxylation enzymes, iron-containing enzymes that utilize as substrates molecular oxygen and the citric acid cycle intermediate alpha ketoglutarate (see Figure 12-13) and that add a hydroxyl group to the 4-carbon atom of a proline residue or to the 3-carbon atom of an asparagine residue in a protein (as shown below).



Clearly the *p*roline *hyd*roxylase (PHD) enzyme as shown above and the similar asparagine hydroxylase can function only when there is sufficient oxygen in the cell!

There are two checkpoints, both based on oxygen-dependent hydroxylase enzymes, which prevent Hif-1α activity at ambient oxygen. As oxygen levels rise above 0.5 percent, the asparaginyl hydroxylase FIH is the first to become activated (Figure 21-7b). FIH transfers an OH group to a specific asparagine residue in the C-

terminal transactivation domain of the Hif-1α protein; hydroxylation blocks the ability of the Hif-1α protein, bound to an HRE, to bind the transcriptional co-activators CBP and p300 (Chapter 8) and thus blocks the ability of Hif-1α to activate transcription.

At slightly higher oxygen concentrations, the PHD2 enzyme becomes activated and adds a hydroxyl group to two proline residues in the Hif-1α protein (Figure 21-7c). Each of these hydroxyl proline residues on Hif-1α forms a part of a separate binding site for the VHL subunit of a specific E3 ubiquitin ligase. The ligase then adds a polyubiquitin chain to the Hif-1α protein, triggering immediate degradation of Hif-1α by the proteasome. Thus the Hif-1α protein is present in cells only when the oxygen level is very low, and any residual Hif-1α protein remaining at moderate oxygen levels becomes inactive because of asparagine hydroxylation.

Interestingly, Hif-1α is the only known target of the PHD enzyme PHD2, and inactivation of PHD2 in mice causes polycythemia, a production of excess numbers of red blood cells. However, the two other human proline hydroxylase enzymes, PHD1 and PHD3, use molecular oxygen as a substrate to add hydroxyl groups to many other proteins with known functions, and current research is uncovering the regulation of these targets by proline hydroxylation. Flies and worms have a single prolyl hydroxylase, attesting to the evolutionary conservation of this oxygen-dependent signaling pathway in defense against hypoxic stress.

The E3 ubiquitin ligase that adds the ubiquitin chain to Hif-1α was discovered not in a study of oxygen sensing but of kidney tumors. Early work showed that certain kidney tumors had inactivating mutations in a tumor suppressor gene, the von Hippel-Lindau (VHL) protein and that these tumors had an abnormally increased number of capillaries surrounding them because they were producing enhanced amounts of VEGF. Trying to identify the function of the VHL protein, investigators soon found that it was in a complex with other proteins that were known to be part of an E3 ubiquitin ligase enzyme. This complex turned out to be the ubiquitin ligase that adds the ubiquitin chains to Hif-1α in which one or both proline residues was modified to hydroxyproline. Indeed, VHL was shown to directly bind to the two segments of the Hif-1α protein that contain a hydroxyproline, and tumor cells with VHL deficiency exhibited up-regulation of all known Hif-1α target proteins, including many enzymes in the glycolytic pathway. An enhancement of glycolysis was known to be important for the growth of many cancer cells; as we learn in Chapter 25, many cancer cells carry out glycolysis even in the presence of high oxygen levels. While producing less ATP per mole of glucose metabolized than does mitochondrial oxidative phosphorylation, aerobic glycolysis produces many intermediates that are siphoned off the glycolytic pathway as building blocks for amino acids and other molecules that are necessary for the rapid growth of cancer cells. ■

A Conserved Family of Oxygen-Sensitive Transcription Factors Found in Plants and Animals Is Regulated by Post-Translational Addition of an Arginine Residue

Like animal cells, plant cells can also become anoxic; often during flooding or waterlogging, cells inside seeds or meristems (see Chapter 22) can also suffer from reduced O_2 diffusion. Recent work has shown that plants such as rice, *Arabidopsis*, and barley respond to hypoxia by increasing the level of a group of oxygen-sensing transcription factors termed *group VII ethylene response factors (ERFs)*. Similar oxygen-sensing factors have been detected in human and other animal cells, indicating that this signaling pathway evolved very early in eukaryotic evolution. Like Hif-1α, ERF proteins promote the synthesis of proteins that enhance survival under low oxygen, including key enzymes in the glycolytic pathway that facilitate anaerobic metabolism, such as alcohol dehydrogenase and pyruvate decarboxylase (see Chapter 12).

ERF proteins all contain an N-terminal cysteine residue that undergoes an unusual series of enzyme-catalyzed reactions at ambient oxygen levels that targets them for degradation by the proteasome (Figure 21-8). The first step, the oxygen-sensing step, is catalyzed by an enzyme termed *cysteine dioxygenase* that binds O_2

and transfers its two oxygen atoms to the N-terminal cysteine residue, forming cysteine sulfinic acid. Note that this is a different reaction than catalyzed by the PHD2 enzyme, but like PHD2 this reaction only occurs at high oxygen levels. The N-terminal cysteine sulfinic acid, in turn, serves as a binding site for an enzyme termed *arginyl-tRNA-protein transferase (ATE)*. This enzyme catalyzes the transfer of the arginine residue from an arginyl tRNA to the amino group of the cysteine sulfinic acid. (Note that this is an unusual reaction in which an aminoacyl tRNA does not participate in ribosome-catalyzed protein synthesis.) Proteins with N-terminal arginine residues are unstable and are immediately subjected to ubiquitination and degradation in the proteasome. In fact, well before this oxygen-sensing pathway was discovered, researchers had defined the so-called N-end rule, in which the in vivo half-life of a protein is determined by the identity of its N-terminal residue. Primary destabilizing N-terminal residues were shown to be recognized directly by the ubiquitin targeting machinery, and proteins with an N-terminal arginine were known to be very unstable.

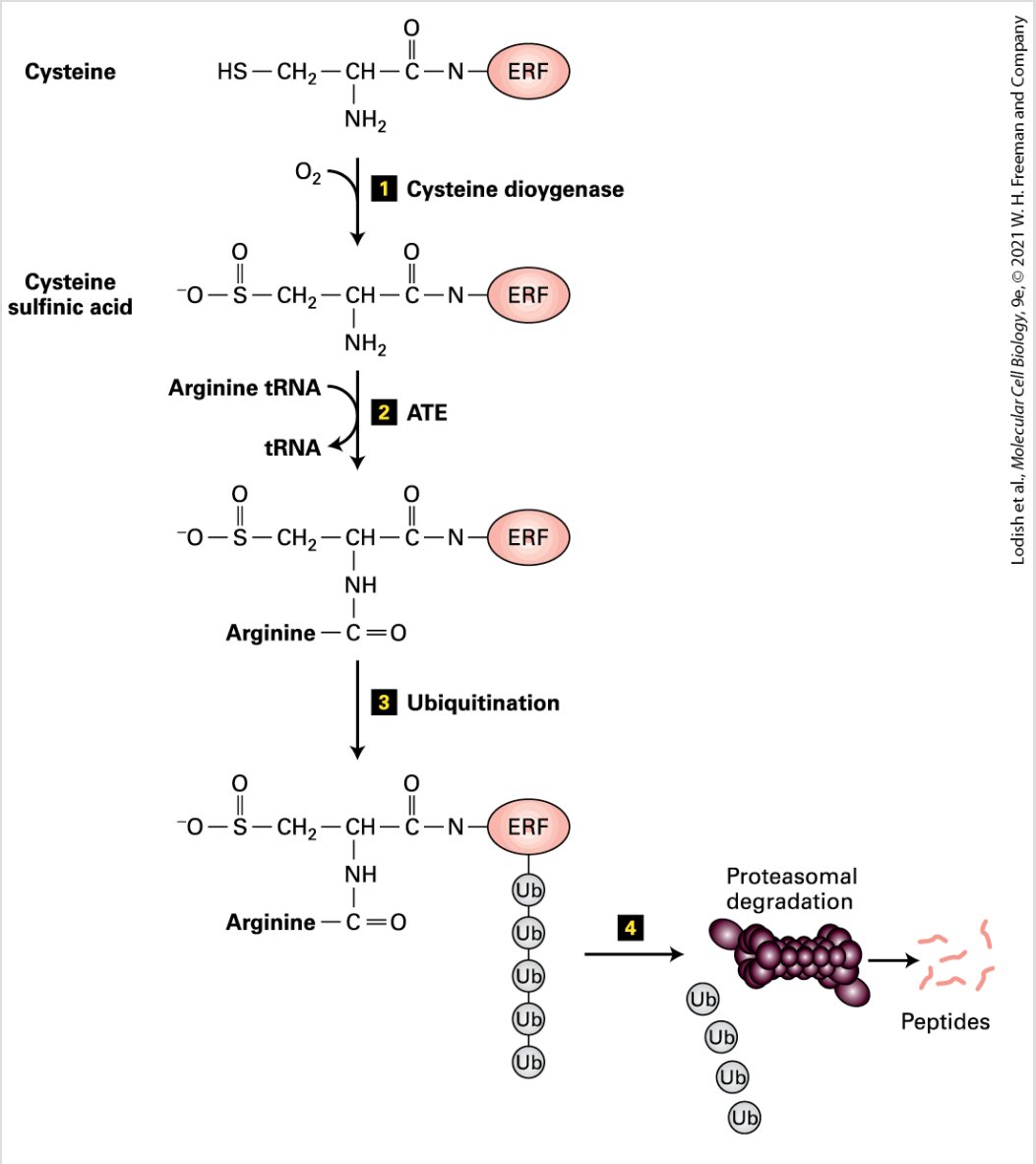


FIGURE 21-8 Degradation of ERF transcription factors at ambient oxygen levels. This pathway occurs in a similar fashion in both plants and animals. Step **1**: At ambient oxygen, but not at reduced oxygen levels, cysteine dioxygenase adds the two oxygen atoms from O_2 to the cysteine residue that is found on the N-terminus of all ERF proteins, forming an N-terminal cysteine sulfinic acid residue. Step **2**: The N-terminal cysteine sulfinic acid serves as a binding site for an arginyl-tRNA-protein transferase (ATE), which then transfers the arginine residue from an arginyl tRNA to the amino group of the cysteine sulfinic acid. Step **3**: An E3 ubiquitin ligase enzyme that uses as substrates

proteins with N-terminal arginine residues adds a ubiquitin chain to the ERF protein that (step **4**) triggers immediate degradation of it by the proteasome.

[Data from D. J. Gibbs and M. J. Holdsworth, 2020, *Cell* **180**(1):22–24.]

Thus at least two classes of oxygen-sensing transcription factors have evolved, ERFs and HIF-1α's. Both induce synthesis of proteins that help the cell or organism survive at low oxygen levels, and both are degraded at ambient oxygen levels by processes initiated by enzyme-catalyzed attachment of O atoms from O_2.

KEY CONCEPTS OF SECTION 21.4

Responding to Low Oxygen

- In metazoans, at low oxygen levels the transcription factor Hif-1α accumulates and induces the transcription of many genes that enable the cell or organism to survive this stress. Proteins that catalyze regulated and rate-limiting steps in the glycolytic pathway are induced in many body cells, and erythropoietin in certain kidney cells.
- At high oxygen levels the prolyl hydroxylase PHD2 transfers an OH group to two proline residues in Hif-1α leading to binding of the tumor-suppressor protein VHL, polyubiquitinylation by an E3 complex, and destruction by the proteasome (see Figure 21-7).
- Many plants and animals express a different family of oxygen-sensitive transcription factors that involves cysteine oxidation and post-translational addition of an arginine residue, leading to polyubiquitinylation and degradation (see Figure 21-8).

21.5 Responding to Elevated Temperatures

The heat-shock response is part of a general cellular response to the accumulation of unfolded proteins caused by one or more cellular stresses, including elevated temperature (giving rise to the name of the response). Other stresses that cause protein denaturation also induce the heat-shock response, such as oxidative stress (which causes oxidation of sulfur atoms in cysteine and methionine), hypoxia, heavy metals, ethanol, and other toxic substances. A similar mechanism induces high-level expression of roughly 50–200 genes in different organisms from archaea to cultured human cells, indicating that this response evolved early in eukaryotic evolution. The **heat-shock proteins** (or **HSPs**) encoded by these genes perform several processes that protect cells from the consequences of unfolding a large fraction of cellular proteins.

In this section, we will discuss two main actors in the heat-shock response. The accumulation of unfolded proteins is sensed by HSP70 proteins expressed at moderate levels in unstressed cells. If these molecular chaperones are unable to restore correct protein folding, a pathway is activated to induce further expression of HSP70 genes and other heat-shock genes that function to prevent aggregation of unfolded proteins. This gene expression program is mediated by the transcription factor HSF (*h*eat-*s*hock *f*actor). This protein is normally kept in check and is released when high

concentrations of unfolded proteins are present in the cell. Below, we will see how HSPs interact with unfolded proteins to restore the correct conformations, how the HSF transcription factor is activated, and how the heat-shock response is induced.

Before describing the details of the heat-shock response, we should note one of its most remarkable and surprising aspects: it is triggered by a temperature increase of just a few degrees. For human cells cultured at 37 °C (310 K), a shift in temperature to 42 °C (315 K) activates transcription of heat-shock protein genes to maximum levels, >100 times the rate in unstressed cells. Since the kinetic energy of molecules in solution is proportional to their temperature in K, this represents only a 1 percent increase in the kinetic energy of protein molecules. How does such a small increase in kinetic energy lead to maximum activation of heat-shock gene transcription? The answer to this puzzle is thought to result from the fact that proteins have evolved to be conformationally flexible at the normal growth temperature of the organisms in which they occur. This flexibility is necessary for the functions of many proteins, both structural and enzymatic. Consequently, most proteins have been optimized by evolution of their primary sequence to be only marginally stable at the normal growth temperature. A small increase in temperature can cause protein unfolding. Without the defenses against protein unfolding from induced heat-shock proteins, these regions of unfolded polypeptide can become entangled, leading to nonspecific protein aggregation and cell death. The heat-shock response that protects against this is triggered by the

large increase in the amount of unfolded polypeptide that results from an increase in temperature of only a few degrees.

The Heat-Shock Response Is Induced by Unfolded Polypeptide Chains

The most conserved heat-shock proteins are molecular chaperones that prevent the formation of nonspecific protein aggregates and assist unfolded proteins in refolding into their native structures. Humans express 11 small HSPs (<45 kDa) exemplified by HSP27. These bind to unfolded, hydrophobic polypeptides, inhibiting them from forming large multimolecular aggregates. The human genes most highly induced in response to elevated temperature encode a group of closely related proteins called *HSP70s*. These are among the most highly conserved protein chaperones throughout all kingdoms of life. The bacterial version, called *DnaK*, shares about 60 percent sequence identity with eukaryotic HSP70 proteins. HSP70s are found principally in the cytosol and in organelles, such as the ER, mitochondria, and chloroplasts. Under normal physiological conditions, HSP70s assist in the de novo folding of nascent and newly synthesized proteins (see Figure 3-19a). In humans, *HSPA8* (also known as *HSC70, "C"* for constitutive) serves this function and is expressed constitutively in all cells; it is not induced by heat shock. The two major HSP70s induced by heat shock, encoded by *HSPA1A* (also known as *HSP70-1*) and *HSPA1B (HSP70-2)*, are nearly

identical in sequence and are ~90 percent identical to HSC70. The two HSP70s are thought to function similarly; in the discussion below they will be referred to together as simply HSP70.

HSP70s (including HSC70) contain two domains, an ATPase domain and a protein-binding domain (see Figure 3-19a). When ATP is bound to the nucleotide-binding pocket in the ATPase domain, the protein-binding domain is held in an open conformation that binds a sequence of seven mainly hydrophobic amino acid residues in an extended conformation. Stretches of hydrophobic amino acids this long generally occur only in the hydrophobic core of properly folded proteins and in transmembrane α helices. Consequently, HSP70s bind unfolded polypeptides or unfolded regions in a polypeptide where such hydrophobic stretches are exposed. This initial binding is low affinity and reversible.

In the following discussion about the control of heat-shock gene transcription, we will be concerned primarily with this ATP-bound form of HSP70s. As discussed in Chapter 3, with the assistance of an *HSP40* co-chaperone, an HSP70 weakly bound to an unfolded region of the *client protein* hydrolyzes the ATP to ADP, resulting in a conformational change of the HSP70 protein-binding domain, so that it unfolds the bound region of the client protein into a conformation that is amenable to proper spontaneous refolding when it is released. ATP hydrolysis also results in very high affinity binding of the client protein. Then, with the assistance of a

nucleotide exchange factor, both the ADP and the client protein are released as ATP becomes bound by the HSP70 ATPase domain. This moves the protein-binding domain into its open conformation. The released, unfolded client polypeptide then has another opportunity to fold into the correct native conformation. Repeated cycles of binding, ATP-dependent unfolding, and spontaneous refolding of the released polypeptide may occur and even lead to the proper refolding of polypeptides in large aggregates.

The Heat-Shock Response Is Regulated Primarily by Related Transcription Factors in All Eukaryotes Called *Heat-Shock Factors*, Including HSF1 in Humans

When the amount of unfolded protein in the cell increases in response to heat shock, HSC70 and any HSP70 induced by heat shock associate with the unfolded regions of polypeptide generated and with any protein aggregates that have formed (Figure 21-9b). Nascent polypeptides still associated with ribosomes are generally the first proteins to unfold because the incomplete polypeptides have fewer stabilizing interactions than the complete protein (see Figure 21-9b). As a result, most of the HSC70 and HSP70s in the cell become associated with unfolded nascent polypeptides. It is this

depletion of free HSC70s and HSP70s that triggers the heat-shock response.

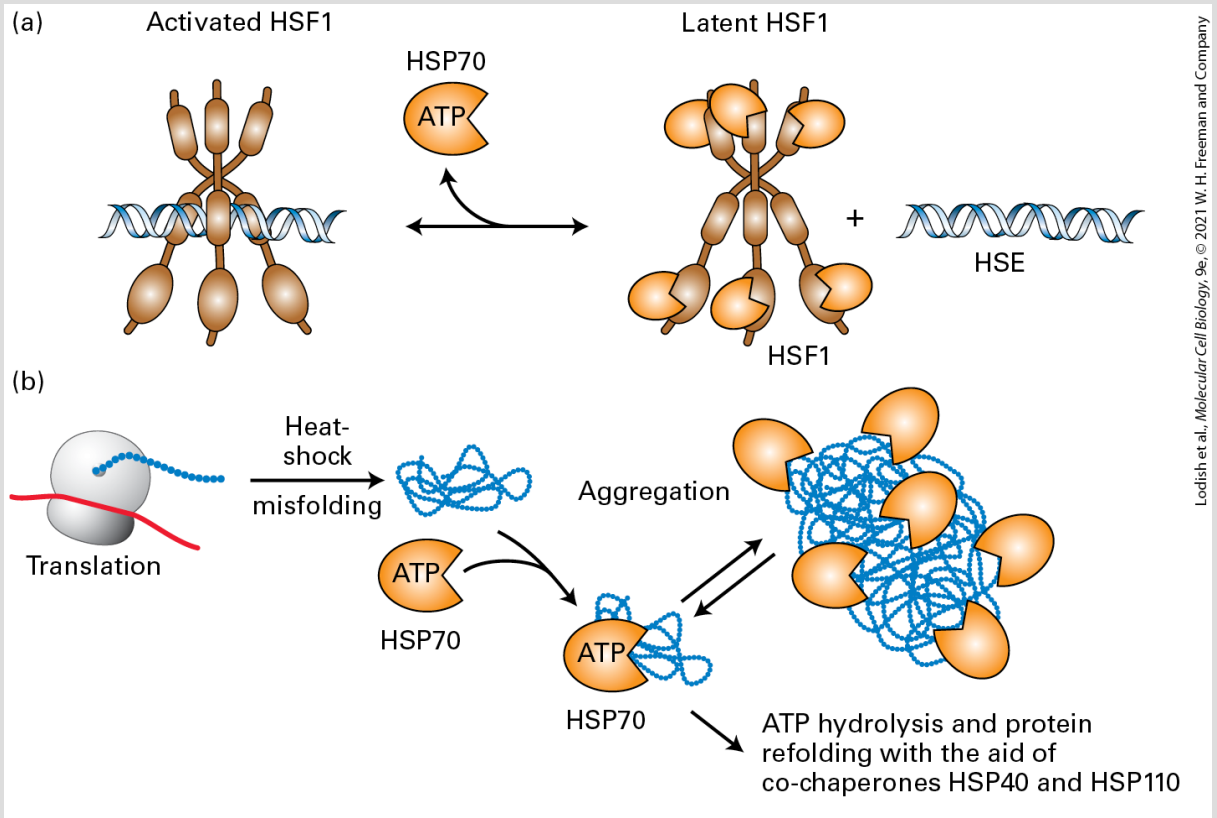


FIGURE 21-9 Regulation of the heat-shock response. HSF1 is inactivated by binding to free HSP70 that has a bound ATP. (a) Regulation of HSF1 by HSP70. Under normal physiological conditions of temperature and the absence of protein denaturants (*right*), most HSF1 is in a latent form complexed with HSC70–ATP in the cytoplasm. If the amount of unfolded protein in the cell increases substantially, most of the cellular HSC70 becomes associated with the denatured proteins. This releases HSF1 to form trimers that are imported into nuclei and activate transcription of genes with promoter proximal HSEs (*left*). (b) During heat shock, nascent polypeptides associated with ribosomes are generally the first to denature because they have fewer interactions stabilizing the correct fold than in full-length proteins. Regions of exposed stretches of hydrophobic amino acids and denatured protein aggregates are bound by the open protein-binding domain of HSP70-ATP, titrating HSP70-ATP away from HSF1 by mass action. The released HSF1 forms trimers that are bound by importins and transported into nuclei where they bind to HSEs and activate transcription of HSP genes, as diagrammed in (a). See A. E. Masser et al., 2019, *eLife* **8**:e47791.

As mentioned above, the heat-shock response involves the sharp increase in the levels of chaperone and other proteins. This is achieved via the release of the HSF1 transcription factor. In its activated form, HSF1 forms a trimer, and an α helix of each monomer interacts with the major groove of three successive alternating inverted repeats with the sequence nGAAn, where *n* denotes any nucleotide (e.g., nTTCnnGAAnnTTCn; Figure 21-9a). HSF-binding sites are known as *heat-shock elements (HSEs).* Virtually all genes activated by the heat-shock response have one or more HSEs in their promoter proximal region. An activation domain at the C-terminus of HSF1 activates transcription through interactions with the mediator complex and other co-activators, as described in Chapter 8. Studies of the *Drosophila Hsp70* gene led to the discovery of a pause in transcription approximately 50 bases downstream from the transcription start site of many genes in metazoans (see Figure 8-16). In *Drosophila* cells cultured at normal temperature, RNA polymerase II remains paused at this site. Following heat shock, HSF is activated and binds to promoter proximal HSEs of the heat-shock genes and stimulates polymerase release through activation of cyclin T-Cdk9 (see Figure 8-16). This results in a very rapid transcriptional response to protein denaturation since RNA polymerase is already engaged in transcription.

As the concentration of unfolded proteins increases following heat shock, the concentration of free HSC70 and HSP70 available to

interact with HSF1 drops precipitously. Since the interaction between HSP70s and HSF1 is reversible, HSF1 is released from HSC70 and HSP70s and is free to trimerize and interact with importins. Then it is transported into nuclei and binds HSEs, stimulating transcription of all genes with promoter proximal HSEs.

Recent studies using purified recombinant yeast HSP70 and derivatives of HSF1 strongly support the following model for heat-shock regulation of HSF1: before activation by heat shock, most HSF1 is located in the cytosol in a latent form associated with HSC70s, which are in excess over HSF1 (Figure 21-9a, *right*). HSC70 in its ATP-bound conformation binds reversibly to several regions in HSF through HSC70's open protein-binding domain that associates with extended, largely hydrophobic regions of polypeptide (Figure 3-19, step **1**). This reversible association of HSC70 with HSF1 inhibits HSF1 interactions with importins required for HSF1 nuclear import. In vertebrates, HSC70 also dissociates a large fraction of the HSF1 trimers into inactive HSF1 monomers bound by HSC70.

Importantly, there are HSEs in the promoter proximal region of the HSF1 gene itself, so *HSF1* is also a heat-shock activated gene. HSF1 mRNA and protein accumulate during extended heat shock, leading to increasing expression of heat-shock genes to the point that they become major cellular proteins. The process is reversed when human cells incubated at 42 °C for several hours, to maximally induce the heat-shock response, are returned to 37 °C. Transcription

of heat-shock genes returns to the level in unshocked cells within a few hours. This occurs because renaturation of unfolded polypeptides catalyzed by HSP70s and other protein chaperones induced by heat shock releases free HSP70s that then associate with activated HSF1, inhibiting its DNA-binding and transcriptional activation and thus repressing transcription of all heat-shock genes.

In the immediate period following extended heat shock, HSF1, HSP70s and other protein chaperones induced by heat shock remain at high concentration. This makes the previously heat-shocked cells resistant to higher temperatures and resistant to higher concentrations of denaturing agents, such as ethanol, than naive cells cultured in normal medium continuously at 37 °C. In this way, the heat-shock response helps cells to both survive acute protein denaturation and also adapts cells to an environment in which they are exposed to denaturing conditions repeatedly. This also explains why cells initially treated with a denaturing agent, such as ethanol, are more resistant to alternative denaturing agents, such as elevated temperature, than cells that were continuously cultured at 37 °C and that did not induce expression of HSPs.

KEY CONCEPTS OF SECTION 21.5

Responding to Elevated Temperatures

- The heat-shock response is induced by unfolded polypeptide chains generated by various conditions, including elevated temperature.

- Most proteins are evolutionarily optimized to be only marginally stable at the normal growth temperature so that they have the flexibility to perform their functions. Consequently, an increase in temperature of only a few K is sufficient to unfold a large fraction of polypeptide chains. Nascent polypeptides associated with ribosomes are especially sensitive to denaturation because they generally have fewer interactions holding them in the proper conformation than the full-length protein (see Figure 21-9b).
- Closely related heat-shock transcription factors (HSF1 in humans) are expressed in all eukaryotes and activate transcription of heat-shock genes when a substantial fraction of unfolded polypeptides accumulate. The heat-shock proteins encoded by these genes bind to and prevent aggregation of unfolded regions of polypeptides and assist in refolding proteins into their proper conformations. HSF1 binds to heat-shock response elements in the DNA sequences of promoter proximal regions and activates high levels of transcription.
- At normal temperature, RNA polymerase II remains paused at approximately +50 in the many heat-shock genes, until HSF is activated, binds to promoter proximal HSEs, and stimulates polymerase release through activation of cyclin T-Cdk9 (see Figure 8-16). This results in a very rapid transcriptional response to protein denaturation, since RNA polymerase is already engaged in transcription and does not have to go through the steps of chromatin decondensation and preinitiation complex assembly before transcription begins.
- At elevated temperature, HSP70s are sequestered away from HSF1 because most HSP70s bind to unfolded regions of polypeptides (see Figure 21-9b). This allows HSF1 to trimerize and become transported into nuclei, where it binds to HSEs in promoter proximal regions of heat-shock genes. HSF1 then stimulates release of the paused RNA polymerase II and high rates of transcription re-initiation of heat-shock genes.

21.6 Sensing Night and Day: Circadian Rhythms

The Earth rotates around an axis passing through the North and South Poles on a 24-hour cycle, with night facing away from and day facing toward the sun. To adapt to this, all organisms have evolved endogenous circadian rhythms that are modulated by dark and light. The word *circadian* derives from *circa diem*, Latin for "about a day," reflecting the fact that circadian rhythms repeat approximately every 24 hours. The study of the biology of circadian rhythms is called *chronobiology*, from the Greek *chronos*, which means "time."

Perhaps the most familiar behavioral circadian rhythm for human and nonhuman animals is the sleep-wake cycle, which for humans and other diurnal animals follows night (sleep) and day (wake). Circadian rhythms are driven by internal molecular clocks within cells that oscillate over a 24-hour period. In addition, there is a process by which these endogenous clocks can be reset by the external cues, called *zeitgeber*, German for "time giver." In fact, the definition of a circadian rhythm requires that it include an endogenous, self-running clock that can be entrained by external cues. The sleep-wake cycle in humans meets both of these criteria. In 1938, two researchers did an experiment in which they spent 32 days in a cave in Kentucky and monitored their sleep-wake cycle. They found that even in the absence of any light, they continued to exhibit circadian sleep-wake rhythms, although over time the

rhythms shifted relative to the actual onset of dark and light in the outside world. Once they left the cave, their sleep-wake cycle resynchronized to the external dark-light cues. As this early experiment demonstrated, human circadian rhythms are not only intrinsic and self-running, but they are also entrainable by external cues. Jet lag provides another commonly experienced example of our internal yet modifiable biological clock: after traveling to a different time zone, individuals usually continue to follow their usual, intrinsic sleep-wake pattern, but within a few days adapt to the dark-light cycle of their new environment.

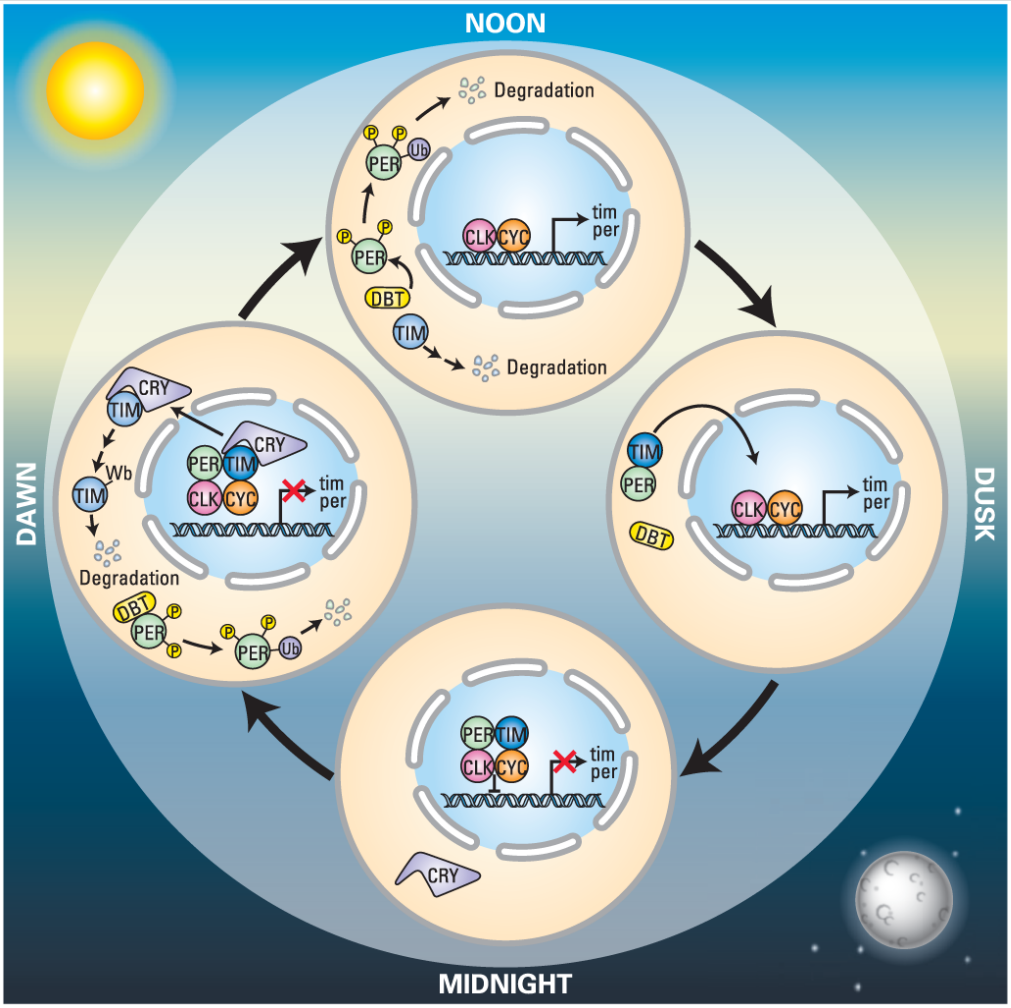
All eukaryotes and some prokaryotes exhibit circadian rhythms. In addition to sleep-wake cycles in humans and other organisms, examples of biological processes that exhibit circadian rhythms include the release of hormones such as insulin and glucagon (described in Section 21.1), feeding behavior in rodents, leaf movement in plants, locomotor activity in flies, and even nitrogen fixation in cyanobacteria. In each case, the behavior involves an intrinsic molecular clock within cells that can be entrained by the external environment. While the most powerful external cue, or zeitgeber, is the dark-light cycle, other cues that impact circadian rhythms include medication, temperature, social interactions, exercise, eating and drinking, and weather. Remarkable progress has been made in understanding the genes and molecules that constitute the molecular clock and the pathways and cues that entrain that clock, together giving rise to circadian rhythms. In this section, we will first consider the mechanisms that give rise to the

molecular clock and then turn our attention to the question of how external cues modulate the molecular clock to entrain circadian rhythms.

The Circadian Clock in Most Organisms Relies on a Negative Feedback Loop

The molecular dissection of circadian clocks provides one of biology's most powerful demonstrations of the genetic basis of behavior, advances that were recognized by the 2017 Nobel Prize in Physiology or Medicine. The initial discoveries on circadian rhythms were conducted in the fruit fly *Drosophila melanogaster*, which, like all organisms, exhibits circadian rhythms in many behaviors. In a series of genetic screens done in the early 1970s, scientists searched for mutant flies that had alterations in their circadian rhythms. This led to the identification of a gene they called *period* or *per*, because mutations altered the periodicity of the circadian rhythm without impacting any other aspect of *Drosophila* physiology. The discovery of the *per* gene opened a field of study that in turn elucidated an entire oscillatory molecular feedback loop that takes ~24 hours to complete. This feedback loop involves transcriptional and post-translational regulatory steps that together form a biological clock that is present in almost all eukaryotic cells and in some prokaryotic cells.

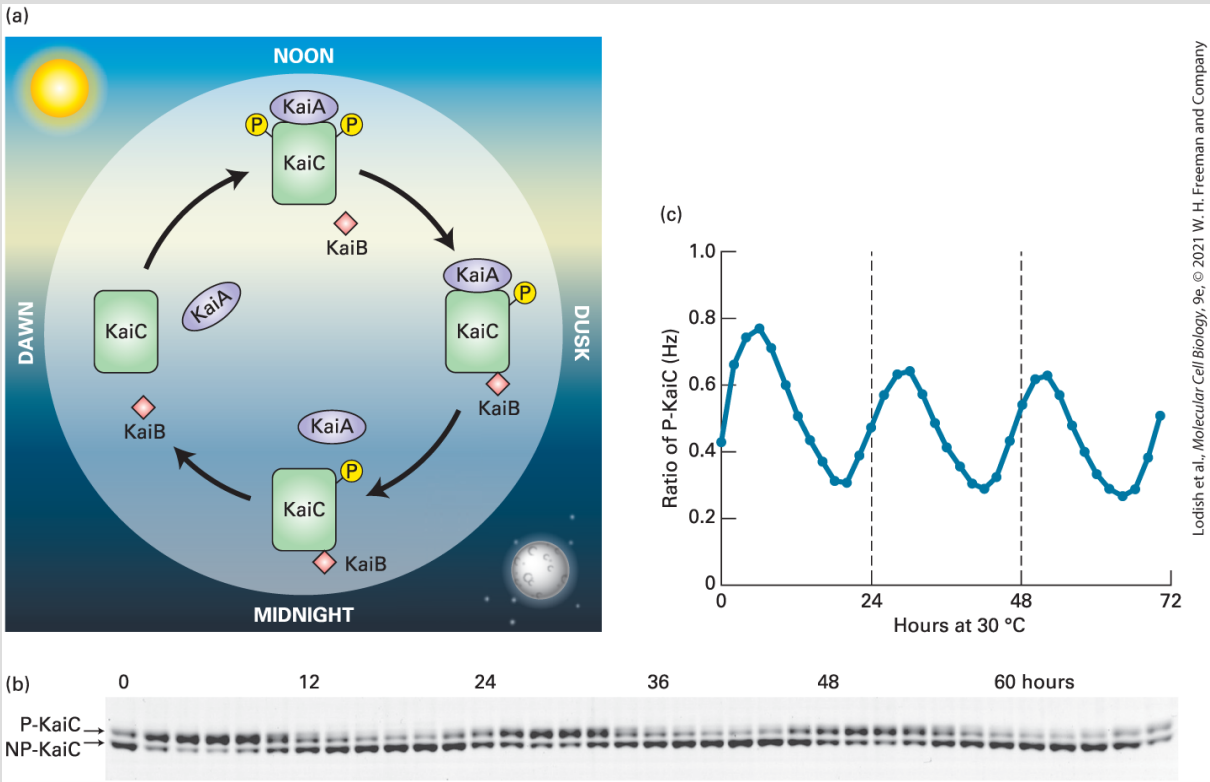
As shown in Figure 21-10, the central components of the clock in *Drosophila* cells constitute a negative feedback loop in which clock genes are regulated by their own protein products. The components of this loop include: two transcriptional repressors, PER and TIMELESS (TIM), encoded by the *per* and *timeless (tim)* genes; two transcriptional activators, CLOCK (CLK) and CYCLE (CYC), encoded by the *clock (clk)* and *cycle (cyc)* genes; and several kinases and E3 ubiquitin ligases. CRY, a cryptochrome (a protein with an attached flavin group that undergoes a conformation change when it absorbs blue light), encoded by the *cry* gene, plays a key role. TIM and PER levels cycle every 24 hours; they accumulate during the day and are degraded during the night.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 21-10 Molecular circadian clock in eukaryotes. The circadian clock in *Drosophila melanogaster* cells consists of two transcriptional repressors, PER and TIM, two transcriptional activators, CLK and CYC, several kinases (including the DBT kinase) and E3 ligases, and the light-sensitive protein CRY. During the day CLK and CYC dimerize to stimulate the transcription of PER and TIM, but TIM and PER proteins both undergo regulated degradation. In the case of PER, it is phosphorylated by the DBT kinase during the day, leading to ubiquitylation by E3 ligases and degradation. Concentrations of PER and TIM peak at dusk, when TIM binds to PER and protects it from DBT phosphorylation and subsequent degradation. During the night, the TIM-PER dimer is transported into the nucleus where it binds to CLK and CYC, blocking their ability to stimulate transcription of the *per* and *tim* mRNAs throughout the night. At dawn, the light triggers a conformational change in the CRY-photosensitive

protein, which then binds to TIM, releasing it from PER, and from CLK and CYC. CRY binding to TIM triggers its ubiquitylation and degradation, and in the absence of TIM binding, PER also undergoes DBT-mediated phosphorylation, ubiquitylation, and degradation. This now allows CLK and CYC to bind to the *per* and *tim* promoters to induce their transcription. This entire cycle takes approximately 24 hours to complete.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 21-11 Molecular circadian clock in prokaryotes. (a) The circadian clock in the cyanobacteria *Synechococcus elongatus* consists of three proteins, KaiA, KaiB, and KaiC. KaiC possesses intrinsic kinase and phosphatase activity. At dawn, KaiC exists in a de-phosphorylated state. After binding of KaiA, however, its kinase activity is activated and it autophosphorylates itself at two residues. One of the residues in KaiC slowly becomes de-phosphorylated, leading to the binding of the KaiB protein and release of KaiA. KaiB binding to KaiC inhibits its kinase activity, leading to fully de-phosphorylated of KaiC and release of KaiB. (b) This entire cyanobacteria cycle can be reconstituted in a test tube, where the phosphorylation state of KaiC undergoes a 24-hour oscillation, as shown in an immunoblot, in which the phosphorylation form of KaiC (P-KaiC) is separated from its nonphosphorylated

form (NP-KaiC). (c) Quantification of the ratio of P-KaiC to NP-KaiC in part (c) reveals a 24-hour circadian rhythm.

[Parts (b) and (c) republished with permission from American Association for the Advancement of Science, from M. Nakajima et al., 2005, "Reconstitution of Circadian Oscillation of Cyanobacterial KaiC Phosphorylation In Vitro," *Science* **308**(5720):414–415; permission conveyed through Copyright Clearance Center, Inc.]

CLK and CYC are both basic helix-loop-helix transcription factors (see Section 8.4) that form a heterodimer and during the day bind to the promoter regions of the *per* and *tim* genes to induce their expression. The levels of PER and TIM proteins are low when their transcription is initially induced because they undergo protein degradation. In the case of the PER protein, its stability is regulated by phosphorylation by the DOUBLETIME (DBT) kinase, which in turn leads to the recruitment of an E3 ubiquitin ligase that targets PER for proteasome-mediated degradation.

The concentrations of PER and TIM peak in the evening, with TIM binding to and stabilizing PER. The TIM-PER complex is then phosphorylated by another set of kinases, which promotes nuclear translocation of the TIM-PER complex. In the nucleus, the TIM-PER complex associates with the CLK-CYC dimer, thereby repressing its transcriptional ability and inhibiting expression of their own *per* and *time* genes throughout the night.

CRY is a flavin-containing photosensitive protein (see Figure 2-33), and light exposure at dawn triggers a conformational change in CRY,

followed by CRY binding to TIM. Upon association with CRY, TIM is ubiquitinated and degraded by the proteasome, freeing up PER to once again be phosphorylated by DBT, ubiquitinated, and degraded. Degradation of TIM and PER relieves the inhibition of CLK-CYC and thereby reactivates the transcriptional activation of PER and TIM during the day.

This feedback mechanism occurs over an approximately 24-hour period, constituting the molecular circadian clock. In simplistic terms, the clock involves two transcriptional repressors, PER and TIM, whose stability is regulated by phosphorylation, ubiquitination, and degradation. PER and TIM repress expression of their own genes by binding to and inactivating the CLK and CYC transcriptional activators (which promote expression of the *per* and *tim* genes).

The Circadian Clock in Bacteria: A Different Solution

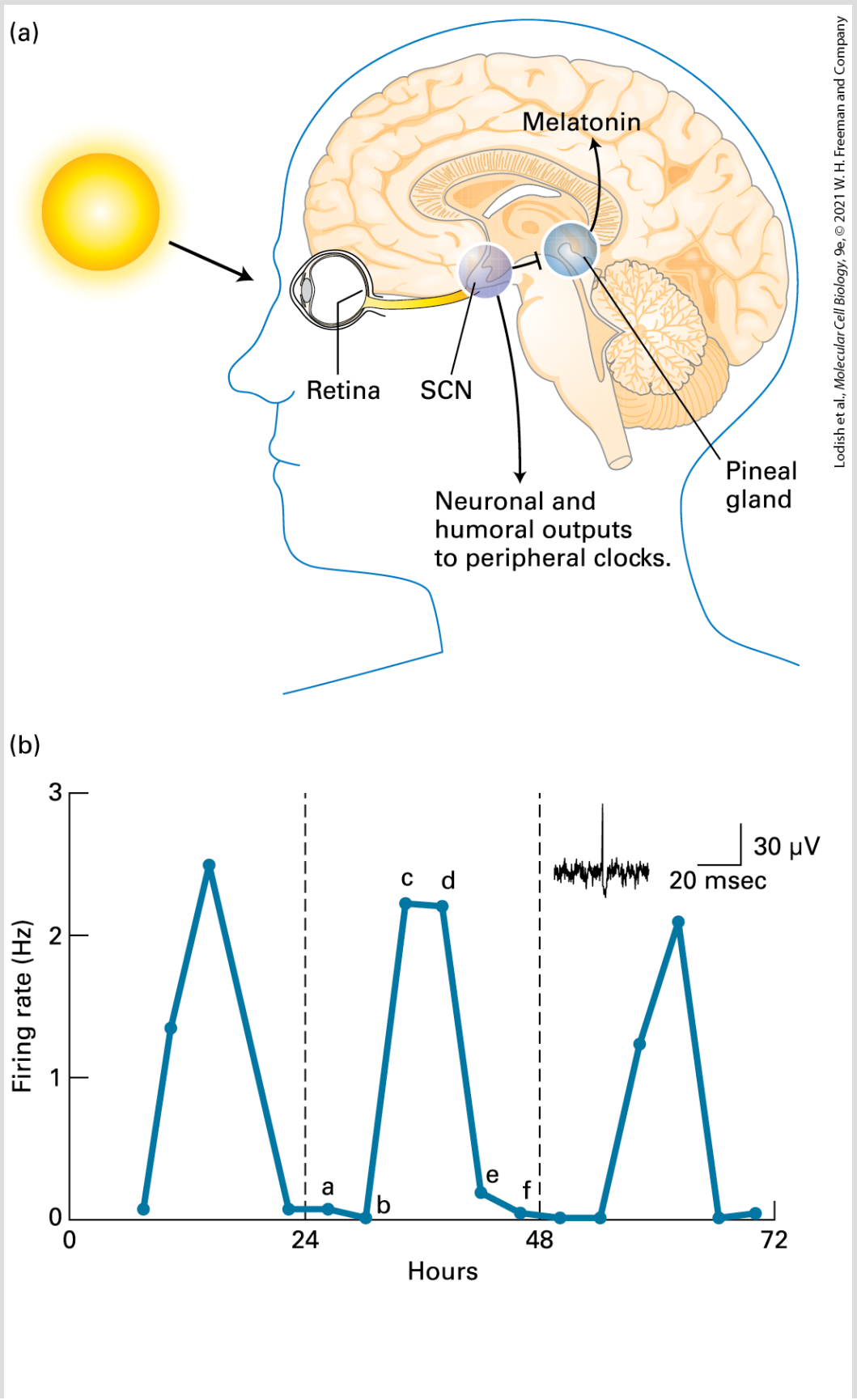
Most, but not all, molecular clocks achieve oscillation through the type of transcriptional feedback loop shown in Figure 21-10. In the cyanobacteria *Synechococcus elongatus,* the clock is composed of three proteins, *KaiA, KaiB,* and *KaiC*, and the oscillatory feedback loop involves a series of post-translational modifications to KaiC (Figure 21-11a). KaiC has intrinsic kinase and phosphatase activities.

As an isolated protein, the phosphatase is most active and KaiC exists mostly in its de-phosphorylated state. KaiA binding to KaiC inhibits the phosphatase activity and promotes the kinase activity, leading to autophosphorylation of KaiC at two sites. KaiC then slowly de-phosphorylates itself at one of the two sites, and the singly-phosphorylated KaiC now binds to KaiB. KaiB inhibits the kinase activity of KaiC, such that over time KaiC returns to its de-phosphorylated state. This simple feedback loop, consisting of a single protein (KaiC) that alternately phosphorylates and de-phosphorylated itself depending on its protein-protein interactions (with KaiA and KaiB) can actually be fully reconstituted in a test tube (Figures 21-11b and c), creating a 24-hour cycle of phosphorylated and de-phosphorylated KaiC!

Suprachiasmatic Nucleus: The Master Clock in Mammals

While molecular clocks are distributed in cells throughout the body, the **suprachiasmatic nucleus (SCN)**, a collection of neurons in a region of the brain called the *hypothalamus* (Figure 21-12a), functions as a master clock or central pacemaker for circadian rhythms in mammals. Indicative of the central role of the SCN in synchronizing all clocks, destruction of the SCN abolishes all regular circadian rhythms and transplantation of an SCN into SCN-ablated animals restores circadian rhythms. Neurons in the SCN receive

input from photoreceptor-containing ganglion cells in the retina. This information is then relayed by SCN neurons to downstream neurons and glands to trigger the release of hormones that in turn regulate an array of biological rhythms. For example, light and dark regulate the release of melatonin from the pineal gland. Melatonin production, a hormone believed to promote sleep onset, is inhibited by light and stimulated by darkness, with the light and dark information being relayed from the retina to the SCN to the pineal gland. As another example, light activates release of cortisol from the adrenal gland. Cortisol, a stress hormone, is highest in daytime and regulates many physiological processes, including blood pressure, blood glucose levels, and inflammation.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 21-12 The suprachiasmatic nucleus (SCN) is the master clock in mammals. (a) Light is detected by cells in the retina and relayed via a neuronal pathway (lavender) to the suprachiasmatic nucleus (SCN) in the brain. The SCN in turn sends signals to other neuronal centers and glands to induce signals that regulate molecular clocks throughout the body. One of these signals is to the pineal gland, which releases melatonin, a hormone that modulates the onset of sleep. (b) Individual neurons isolated from the SCN show 24-hour oscillating rhythms in action-potential firing (see Chapter 23 for discussion of action potentials) that persist for as long as the neurons are maintained in culture.

[Part (b) Data from D. K. Welsh et al., 1995, *Neuron* **14**:697–706.]

In addition to relaying information about light and dark, the SCN possesses a particularly robust intrinsic molecular clock. The SCN consists of a heterogeneous cluster of about 10,000 neurons on each side of the brain. When SCN neurons are isolated and placed in culture, they show circadian rhythms in levels of free calcium in the cytosol and electrical firing that follows a 24-hour period for as long as they are maintained in culture (Figure 21-12b). In addition to these cell-autonomous circadian rhythms, signaling between neurons in the SCN demonstrate circadian rhythms, such that the circuit properties of the SCN are critical to the generation of circadian rhythms throughout the body.

Recent research has revealed the impact of circadian rhythms on many aspects of cell biology, from transcriptional, translational, and post-translational regulation to the dynamics of intracellular organelles, to cell differentiation, and to cell death. In research

laboratories, this realization has motivated many scientists to consider the time of day during which they conduct their experiments. Circadian rhythms are also clinically important. In addition to primary diseases of circadian rhythms, which include narcolepsy (a syndrome characterized by daytime sleepiness and involuntary, uncontrollable sleep episodes) and sleep disorders in which the phase of sleep is altered, many other diseases are accompanied by alterations in circadian rhythms. For example, nearly 25 percent of individuals with Alzheimer's disease have significant sleep disorders, and up to 75 percent of children with autism spectrum disorders have sleep disturbances. Increased molecular and cell biological insights into the functioning of molecular clocks and the mechanisms by which external cues entrain these clocks are likely to have important therapeutic implications for these clinical disorders. ■

KEY CONCEPTS OF SECTION 21.6

Sensing Day and Night: Circadian Rhythms

- All organisms exhibit ~24-hour circadian rhythms that are driven by molecular clocks within cells and that can be reset by external cues such as light and dark.
- A negative feedback loop in which clock genes are regulated by their own protein products is conserved from flies to humans. This feedback loop consists of two transcriptional repressors, PERIOD and TIMELESS, two transcriptional activators, CLOCK and CYCLE, and several kinases and E3 ubiquitin ligases. PERIOD and TIMELESS levels cycle every 24 hours, accumulating during the day and being degraded at night.
- Bacteria have a different circadian clock, consisting of three proteins, KaiA, KaiB, and KaiC, that form a transcriptional feedback loop that oscillates on a 24-hour cycle.

- In mammals, the suprachiasmatic nucleus (SCN) in a region of the brain called the *hypothalamus* serves as a master clock or central pacemaker. Ablation of the SCN abolishes all regular circadian rhythms.

21.7 Sensing and Responding to the Physical Environment

We learned in Chapters 16 and 17 how cells coordinate their activities by communicating with each other through receptors that bind soluble molecules secreted by a signaling cell or signaling proteins bound to the plasma membrane of a neighboring cell. Earlier in this chapter, we saw how cells regulate their metabolism and gene-expression patterns by sensors that respond to changes in many small molecules in their environment, as well as to changes in temperature and light. Cells also perceive their microenvironment through physical and mechanical cues (mechanotransduction) from the extracellular matrix (ECM) in which they are imbedded and by tension across intercellular adherens junctions generated from the indirect connections of these junctions to the actin cytoskeleton (see Figures 20-1 and 20-38). In this section, we discuss the Hippo pathway, a protein kinase signaling cascade conserved in all metazoans that senses and integrates these cues about the physical properties of the ECM and modulates cell growth and differentiation in response. As an example, during development, cells proliferate and differentiate to generate organs of specific sizes, and we will see how mutations in this pathway can result in excessive organ growth and how the Hippo pathway regulates the expression of multiple genes involved in growth control and differentiation. Abnormalities in this pathway contribute to many developmental disorders and generation of many types of cancers.

The Hippo Kinase Cascade Pathway in *Drosophila* and Mammals

The Hippo kinase pathway (Figure 21-13a) is conserved in all metazoans and their closest unicellular relatives, suggesting that it evolved before multicellularity. The discovery of the Hippo pathway came from studies by *Drosophila* geneticists searching for genes that control cell replication, genes now known as *tumor suppressors* (see Chapter 25). Examination of F_2 progeny derived from flies treated with a mutagen (see Figure 6-4) revealed that homozygous recessive mutations in a gene named *Hippo* resulted, as mentioned previously, in abnormally fat larvae that resemble a small hippopotamus, leading to the name of the gene. Further examination revealed that this phenotype was due to the growth of an abnormally large fat body, the *Drosophila* larva physiological equivalent of the vertebrate liver. Isolation and characterization of other *Drosophila* mutants with a similar homozygous recessive phenotype led to identification of genes encoding the core components of the Hippo pathway, heterodimeric protein kinases Hpo-Sav and Wts-Mats, which regulate the nuclear import of the transcriptional co-activator Yki (Figure 21-13a). Once imported into the nucleus, Yki activates transcription by binding to a sequence-specific DNA-binding protein encoded by *Sd,* which does not affect transcription in the absence of Yki. Later in this section, we learn about some of the many genes involved in cell proliferation and differentiation induced by the

human orthologs of Yki, called *YAP* and *TAZ*, and the human counterparts of *Drosophila Sd*, DNA-binding proteins TEAD 1 to 4.

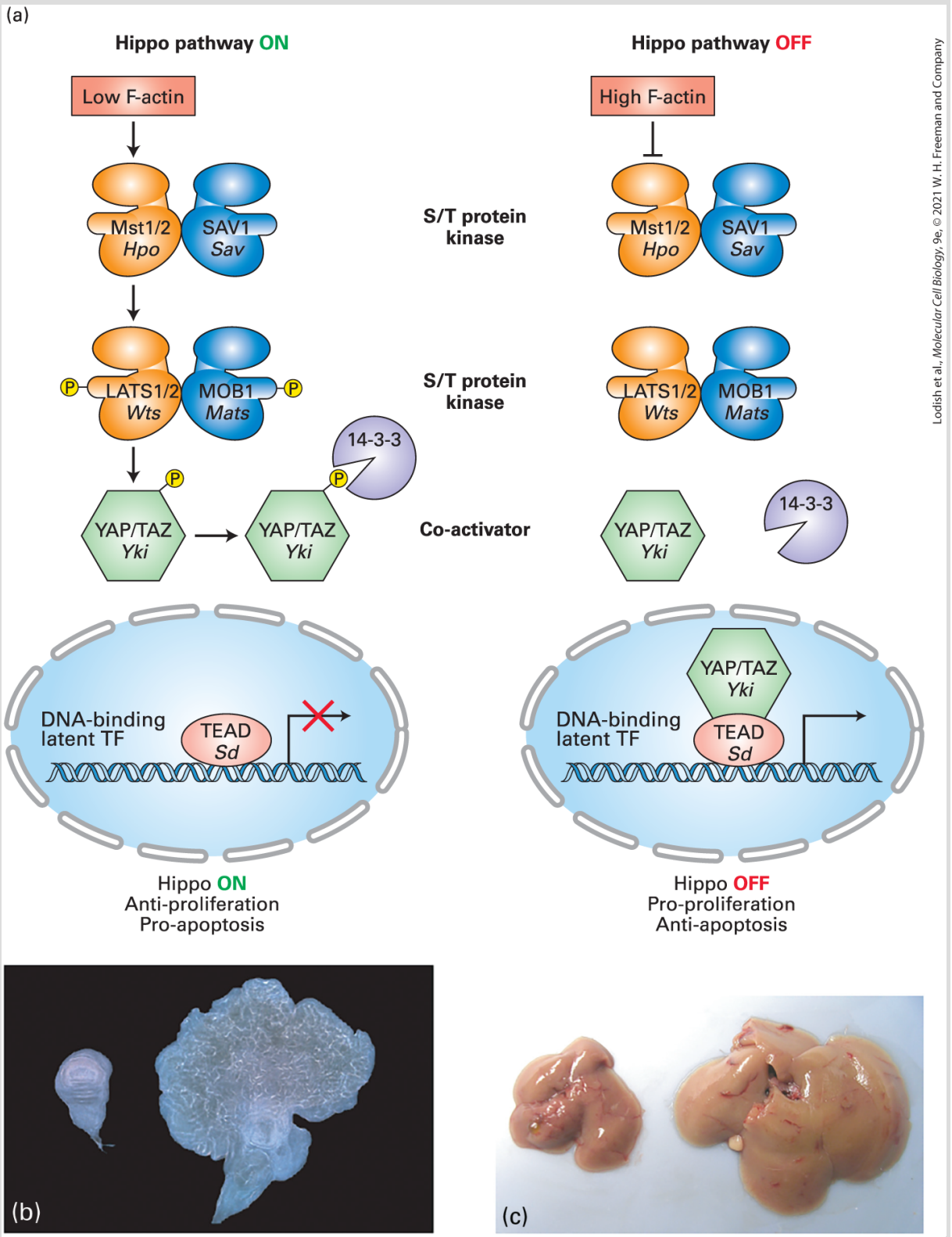


FIGURE 21-13 The Hippo pathway has similar components and functions in *Drosophila* and mammals. (a) Proteins with orthologous functions in the Hippo pathway in *Drosophila*

and humans are indicated in italics and regular fonts, respectively. Arrows indicate activation of the downstream signaling component; see text for details. (b, c) Consequences of overexpression during embryogenesis of Yki in *Drosophila* wing imaginal disk (b), or of YAP in mouse liver (c) during embryogenesis. Normal tissues are shown to the left and tissues overexpressing Yki or YAP are shown to the right.

[Parts (b) and (c) republished with permission from Elsevier, from D. Pan, 2010, "The Hippo Signaling Pathway in Development and Cancer," *Dev. Cell* **19**(4):491–505; permission conveyed through Copyright Clearance Center, Inc.]

When the Hippo pathway is active, activated Hpo-Sav kinase phosphorylates and activates the Wts-Mats protein kinase, which then phosphorylates and inactivates the Yki transcriptional co-activator. Yki phosphorylated at specific serine residues is inactive as a transcriptional co-activator because it is bound by a class of *14-3-3 proteins* anchored in the cytoplasm (see Chapter 16) that bind phosphoserine and closely related phosphothreonine residues, preventing nuclear import of Yki.

The Hippo pathway is highly conserved in mammals (see Figure 21-13a), except that there are duplicates of the orthologous *Drosophila* genes. Human genes *MST1* and *MST2* are closely related to the orthologous *Drosophila Hpo* gene, encoding the catalytic subunit of the first protein kinase in the pathway. *LATS1* and *LATS2* are both orthologous to *Drosophila Wts* and encode the catalytic subunit of the second protein kinase in the pathway. And *YAP* and *TAZ* are ~60 percent identical in amino acid sequence and are both orthologs of the *Drosophila* transcription factor gene *Yki*. As discussed in Chapter

7 for β-globin gene family members, the proteins encoded by these duplicated and diverged Hippo pathway genes probably have similar but somewhat specialized functions. Inactivation of the Hippo pathway leads by default to nuclear import of the YAP/TAZ transcriptional co-activators, directed by their nuclear localization signals (NLSs). This leads to transcriptional activation of genes that have TEAD transcription factors prebound to enhancers that regulate them. This requirement for pre-binding of TEAD factors for gene activation by the imported YAP and TAZ co-activators determines which genes respond to regulation of Hippo signaling in different cell types.

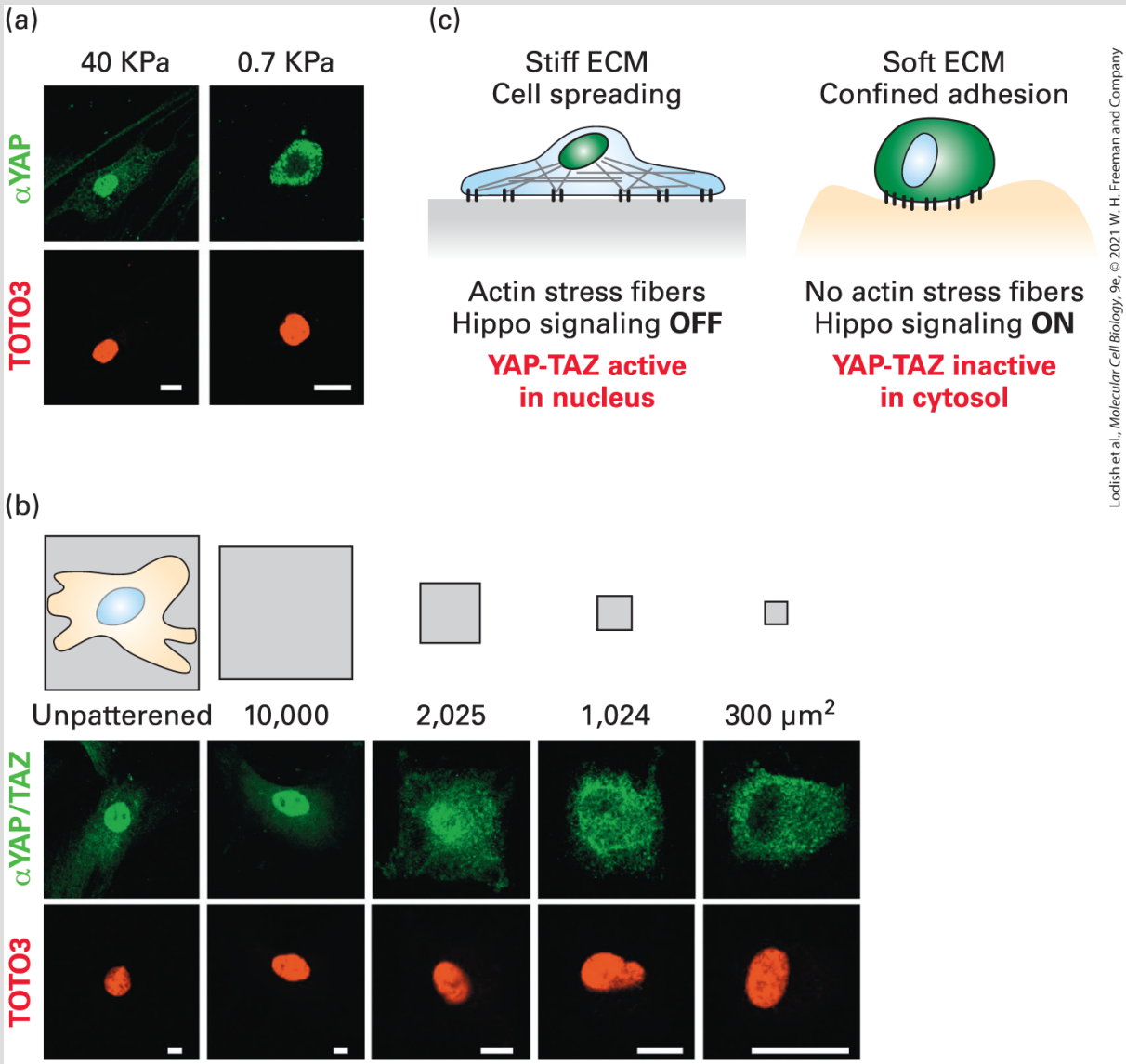
The functional similarity between the Hippo pathway in *Drosophila* and mammals was dramatically demonstrated by experiments in which *Drosophila Yki* or the human *YAP* gene were overexpressed in cells in developing embryos. Overexpression of *Yki* in the *Drosophila* wing imaginal disc that normally develops into a wing during larval morphogenesis led to a great increase in the size of the imaginal disk, while the overall morphology of the developing organ was much less affected (Figure 21-13b). Analogously, overexpression of YAP in the developing mouse liver led to a greatly increased size of the liver compared to normal, but the overall morphology of the developing organ was near normal (Figure 21-13c). Further analysis showed that increased organ size in both *Drosophila* and mice was due to an increase in the number of cell divisions and a decrease in apoptosis in the developing organs. Further, YAP and TAZ are

considered to be oncogenes (see Chapter 25) because they are expressed at much higher than normal level in several types of human cancers where they likely contribute to oncogenesis by stimulating cell cycling (see Chapter 19) and inhibiting apoptosis (see Chapter 22).

Regulation of the Hippo Kinase Cascade by Cell Interactions with the Extracellular Matrix and by Tension on Actin Filaments

We learned in Chapter 4 that most types of animal cells will grow in culture only when attached to a solid surface. Many cells secrete extracellular matrix proteins such as fibronectin and collagen that stick to the surface of the plastic or glass culture dish and that in turn bind to integrins on the cell surface and allow the cell to spread on the culture dish. Evidence that the Hippo pathway, and consequently the nuclear localization of YAP and TAZ, is regulated by cellular interactions with the ECM was observed when wild-type human epithelial and mesenchymal stem cells were cultured on either a stiff or a soft surface (Figure 21-14a). Stiffness of the substratum is measured in units of pressure, kilopascals (KPa), required to deform its shape. At high stiffness (40 KPa) comparable to a glass coverslip routinely used to culture a monolayer of cells for immunofluorescence studies, the Hippo pathway was off and YAP

was nuclear. However, at low-substrate stiffness (0.7 KPa) the Hippo pathway was on and YAP was primarily cytoplasmic (see Figure 21-14a).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 21-14 Regulation of Hippo pathway activity by physical properties of the ECM. (a) YAP localization in mesenchymal stem cells attached to a stiff (40 KPa) or soft (0.7 KPa) substratum, as detected by immunofluorescence. TOTO3 is a fluorescent stain for DNA and shows the location of nuclei. (b) Influence of the area of a square of fibronectin on nuclear localization of YAP and TAZ (both immunostained green).

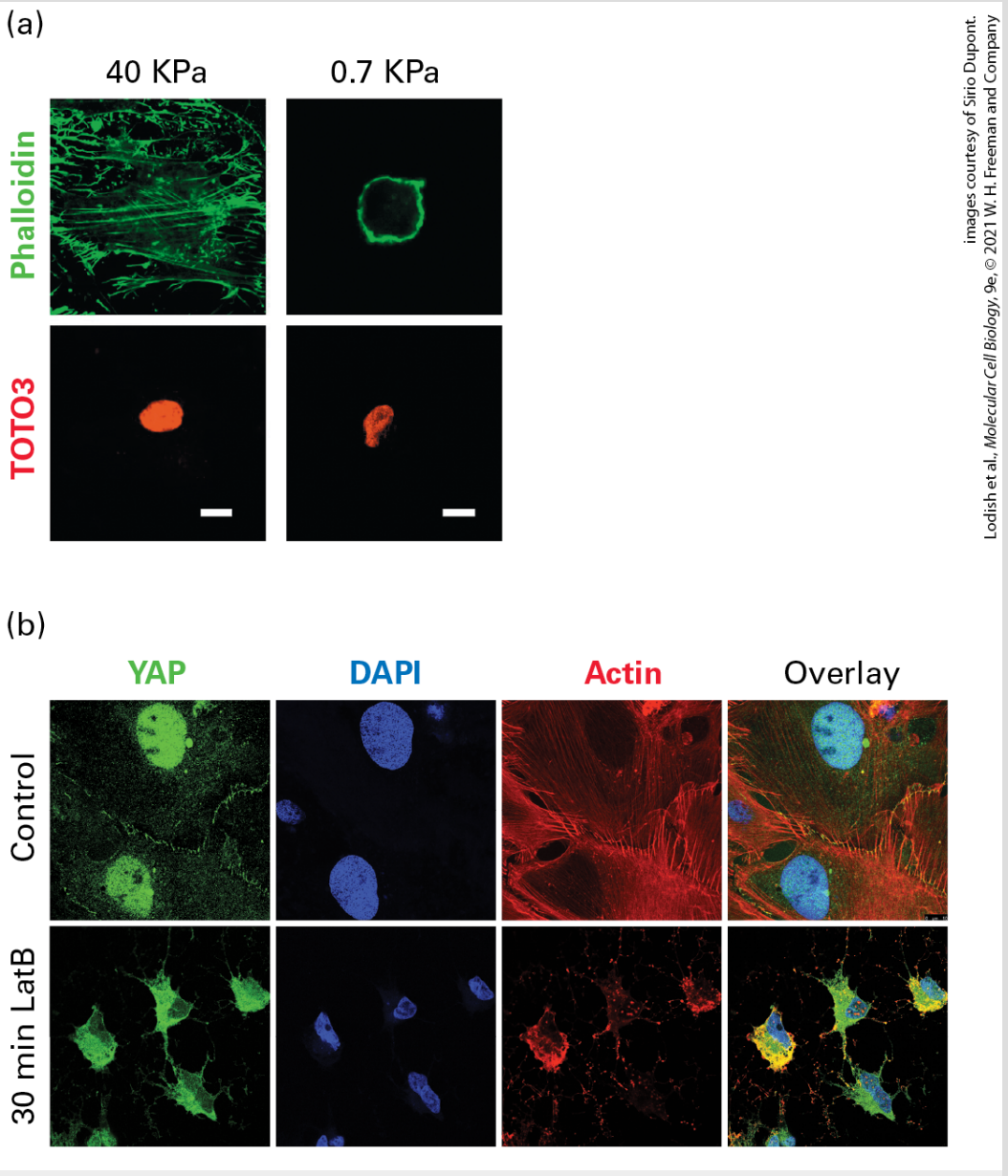
(c) Summary of the influence of substrate stiffness and the area of cell spreading on assembly of actin stress fibers and YAP/TAZ subcellular localization.

[Parts (a) and (b) republished with permission from Springer Nature, from S. Dupont et al., 2011, "Role from YAP/TAZ in Mechanotransduction," *Nature* **474**:179–183; permission conveyed through Copyright Clearance Center, Inc.; Images courtesy of Sirio Dupont.]

In another revealing experiment, mesenchymal stem cells cultured from an early mouse embryo were plated on successively smaller squares of fibronectin dried onto a surface that cells do not otherwise attach to (Figure 21-14b). Individual cells spread only on the patches of fibronectin. At areas $>2025\ \mu m^2$ where the cell could spread out to near the area it covers on a continuous surface of fibronectin, YAP was nuclear, indicating that the Hippo pathway was off. However, in cells plated on smaller fibronectin squares of only $300\ \mu m^2$ where they could not spread out, YAP was primarily cytoplasmic (see Figure 21-14b), indicating that the Hippo pathway was active, leading to phosphorylation of YAP and TAZ and their retention in the cytoplasm. These experiments revealed that the activity of the Hippo pathway responded to the physical properties of the ECM and the area of cell adhesion (Figure 21-14c).

We learned in Chapter 17 that interactions of actin filaments with proteins attached to the plasma membrane is important for many aspects of cell adhesion. Indeed, cells that spread out on a stiff substratum contained extensive actin stress fibers visualized by staining with the fluorescently labeled fungal toxin phalloidin (Figure 21-15a). These are the cells in which the Hippo pathway was

off, as indicated by their nuclear YAP (see Figure 21-14a). In contrast, no actin stress fibers were observed in cells plated on a soft substratum (see Figure 21-15a); in these cells Hippo signaling was active, as indicated by cytoplasmic YAP (see Figure 21-14a). This correlation suggested that actin stress fibers somehow inhibit the Hippo pathway, leading to nuclear import of YAP that is not bound by cytoplasmic 14-3-3 proteins. Further experiments using the fungal toxin Latrunculin, which disrupts actin filaments, showed that actin fibers are indeed required for inhibition of Hippo signaling and nuclear localization of YAP (Figure 21-15b).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company
images courtesy of Sirio Dupont.

EXPERIMENTAL FIGURE 21-15 Regulation of Hippo pathway activity by tension on actin filaments. (a) Visualization of actin stress fibers with fluorescently labeled (green) phalloidin, a fungal toxin that binds actin filaments, but not actin monomers, in mesenchymal stem cells attached to a stiff (40 KPa) or soft (0.7 KPa) substratum. (b) Actin stress fibers are required for YAP nuclear localization. E1AKD293 cells are flattened against the substratum by tension on actin fibers running from apical to basal adherens junctions. The tension results from myosin motor proteins that move actin fibers extending from apical adherens junctions past actin fibers extending from basal adherens

junctions (see Chapter 17). This flattens the whole cell against the substratum, including the nucleus, making it appear large when viewed from the top (control). Thirty minutes after addition of Latrunculin B (LatB) to the media, actin cables were dissociated into actin monomers and nuclei took on a globular morphology that appeared much smaller when viewed from the top. In the absence of actin fibers, the Hippo pathway was on and YAP was sequestered in the cytoplasm.

[Part (a) republished with permission from Springer Nature, from S. Dupont et al., 2011, "Role of YAP/TAZ in Mechanotransduction," *Nature* **474**:179–183; permission conveyed through Copyright Clearance Center, Inc.; Part (b) republished with permission of the authors, from N. R. Zemke, D. Gou, and A. J. Berk, 2019, "Dedifferentiation by Adenovirus E1A Due to Inactivation of Hippo Pathway Effectors YAP and TAZ," *Genes Dev.* **33**(13–14):828–843.]

The Hippo Pathway and Early Embryogenesis

The Hippo pathway is essential for many steps in development including the very first example of cellular differentiation during embryogenesis in mammals, the differentiation of cells on the surface of a 16–32 cell embryo, called a *morula*, into trophectoderm cells (see Figure 22-2). The inner cells of the morula remain undifferentiated and divide to form the inner cell mass of the early blastocyst that later develops into the embryo itself. The outer trophectoderm is required for implantation of the blastocyst in the uterine epithelium; it will later form extraembryonic tissues including the embryonic portion of the placenta (see Figure 22-3).

The Hippo pathway is active only in the inner cells of the morula; these are the cells that are surrounded on all sides by other cells and connected to them by adherens junctions (Figure 20-14e; Figures 21-16a and b). In these cells, YAP and TAZ are phosphorylated by the LATS kinase and retained in the cytoplasm (Figure 21-16a). In contrast, the outer cells of the morula have a free surface not attached to other cells. Under this segment of the plasma membrane are long filaments of F-actin, and as we will see this causes the Hippo pathway to become inactive. As a result, YAP/TAZ are imported into the nucleus where, together with a TEAD (see Figure 21-13a), they activate transcription of genes that initiate the trophectoderm differentiation program.

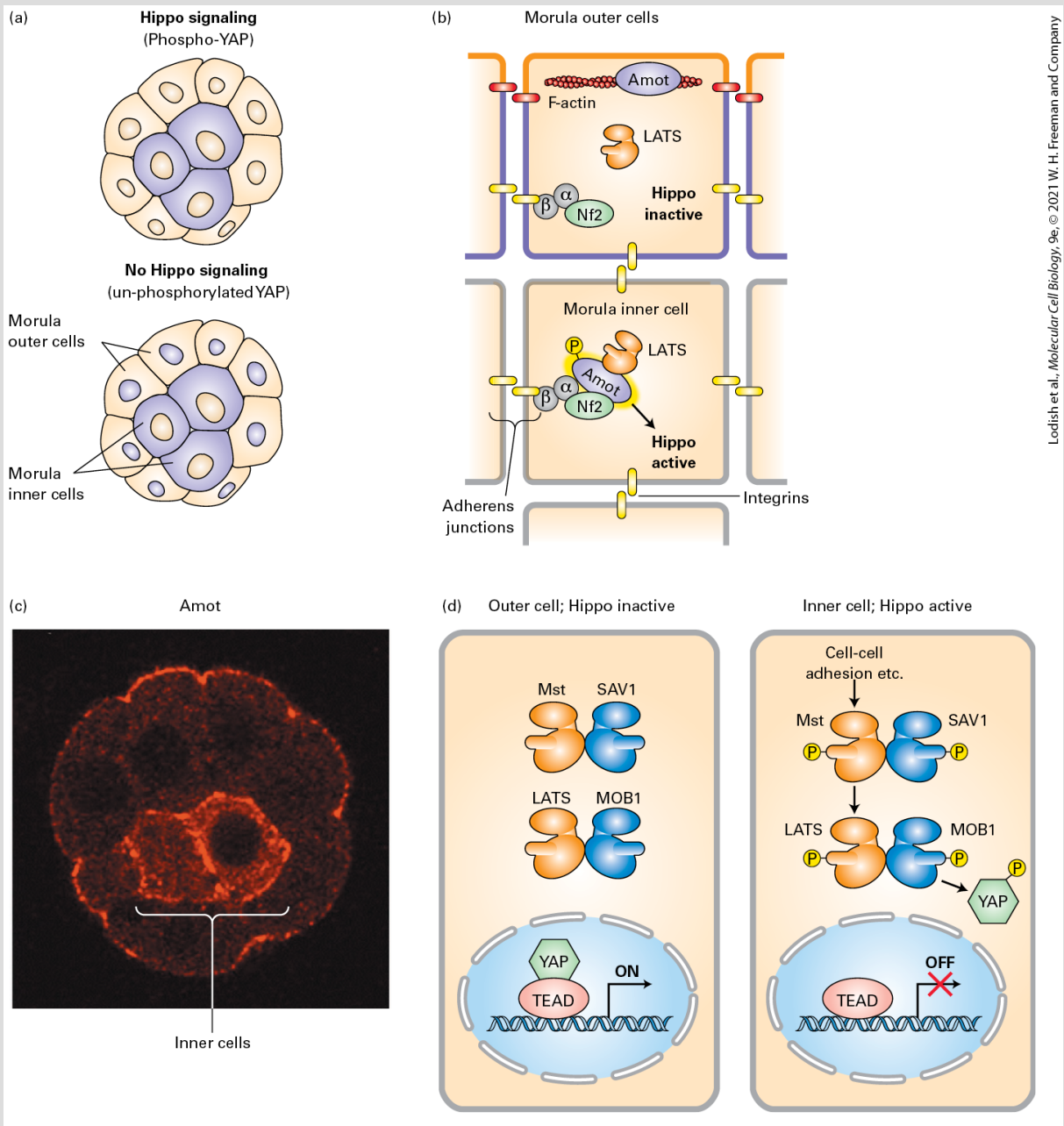


FIGURE 21-16 Regulation of the Hippo pathway in the morula by the subcellular localization of angiomotin. (a) Active Hippo signaling in the inner cells of the morula is indicated by the cytoplasmic localization of both phospho-YAP (*upper* panel) and total YAP (*lower* panel). In the outer cells, the Hippo pathway is off and YAP is unphosphorylated and nuclear. (b) Diagram of angiomotin (Amot) associated with adherens junctions on all cell surfaces of the inner cells (*lower panel*) and with the apical actin fibers in the outer cells (*upper panel*). In inner cells, association of Amot with Nf2 at adherens junctions stimulates Amot phosphorylation by LATS1 and 2, activating the Hippo pathway. In the outer cells,

sequestration of Amot at the apical surface inactivates Hippo signaling. (c) Immunofluorescent staining of Amot in the morula demonstrating its localization to all surfaces of the two inner cells shown and its sequestration to the apical surface of outer cells. (d) Summary of Hippo pathway activity and YAP nuclear localization in inner and outer cells of the morula.

[Part (c) republished with permission from John Wiley & Sons, Inc., from H. Sasaki, 2017, "Roles and Regulations of Hippo Signaling During Preimplantation Mouse Development," *Dev., Growth Differ.* **59**:12–20; permission conveyed through Copyright Clearance Center, Inc.]

In the morula, Hippo pathway signaling is controlled by an actin-binding protein called *angiomotin* (*Amot*; Figure 21-16b). In the inner cells, Amot is associated with the adherens junctions present on all surfaces of the cell (Figure 21-16c). The interaction between Amot and adherens junctions requires another junction-associated Hippo pathway component, neurofibromatosis 2 (Nf2), discussed below. In this adherens-junction-associated protein complex, a specific Amot serine residue becomes phosphorylated by the Hippo pathway protein kinases LATS1 and 2. In these preimplantation embryos, the phosphorylation of Amot stimulates phosphorylation of YAP/TAZ by LATS1/2, causing retention of YAP/TAZ in the cytosol.

Amot proteins (three paralogs in mice and humans) have an F-actin binding and polymerizing activity, and the Amot serines phosphorylated by LATS1/2 are in their actin-binding domain. Phosphorylation of Amots at these sites inhibits their ability to bind F-actin and enhances their phosphorylation by LATS1/2. This enhanced Amot phosphorylation stimulates Hippo signaling at adherens junctions (see Figure 21-16b), maintaining cytoplasmic

sequestration of YAP/TAZ in the inner cells of the morula (see Figure 21-16a). However, in the outer cells of the 16–32 cell embryo, cell polarization through the Par-aPKC system (see Chapter 22) induces assembly of a thick band of cortical actin filaments at the apical surface of the cell, where adherens junctions do not form because there are no neighboring cells. In these outer cells unphosphorylated Amot is sequestered away from the adherens junctions by binding to the cortical actin filaments. Consequently, Amot does not get phosphorylated by LATS, the Hippo pathway is inactive, and YAP/TAZ are imported into the nucleus (Figure 21-16d). In these cells nuclear YAP/TAZ together with TEAD (see Figure 21-13a) activate transcription of the homeodomain gene *Cdx2* and the zinc-finger transcription factor *Gata3,* which in turn initiate the developmental program of trophectoderm cells. In this way, cell-cell interactions via E-cadherins on the outside of the plasma membrane control Hippo signaling in the cytoplasm.

The experiments in Figures 21-14 and 21-15 indicate that the signals that control the activity of the Hippo pathway are regulated by physical properties of the ECM, such as stiffness, the area of the interacting adhesive cell surfaces, and tension on the actin cytoskeleton. These in turn are controlled by the changing cellular and tissue environment during embryogenesis and later stages of development to control Hippo signaling and the proliferation and differentiation of progenitor cells in the developing embryo. Such adhesive and mechanical forces ultimately control the sizes and

morphology of tissues and organs during development and during tissue regeneration after injury.

Mutations that inactivate one allele of human NF2, a component of the Hippo pathway (see Figure 21-16b), cause abnormal growth of peripheral nervous system tissue in the dominant genetic disease neurofibromatosis type II. NF2 associates with the cadherin-associated protein α-catenin on the cytoplasmic side of adhesion junctions (Figure 20-14e). As noted above, a specific Amot serine residue becomes phosphorylated by Hippo pathway protein kinases LATS1 and 2 in these adherens junction-associated protein complexes (Figure 21-16b). In preimplantation embryos, this phosphorylation of Amot stimulates phosphorylation of YAP/TAZ by LATS1/2, causing sequestration of YAP/TAZ in the cytoplasm. Phosphorylation at this site also blocks Amot association with actin filaments. In patients with one inactivated NF2 gene, half the normal amount of the NF2 protein is expressed. This decreases the activity of the Hippo pathway leading to abnormally high activation by YAP/TAZ and excessive cell proliferation. ■

Much remains to be learned about how regulation of the Hippo pathway in response to tension on the actin cytoskeleton contributes to tissue size and morphology, although it is clear that the overall process helps to sculpt the size and shape of metazoan organs. Understanding Hippo pathway regulation is medically important because abnormalities in the Hippo pathway function

contribute to oncogenic transformation and the invasiveness and metastasis of human tumors. Deeper understanding of Hippo pathway regulation may allow the logical design of therapeutic interventions in these diseases.

KEY CONCEPTS OF SECTION 21.7

Sensing and Responding to the Physical Environment

- The Hippo protein kinase pathway controls the sizes and morphologies of several organs during embryonic development of metazoans.
- The Hippo pathway is also required for regeneration of damaged tissues and organs in many metazoan tissues capable of regenerating after injury.
- The core components of the pathway are two heterodimeric protein kinases (Hpo-Sav and Wts-Mats in *Drosophila*). The activity of the pathway is regulated by signals from adherens junctions, mechanical forces acting on the actin cytoskeleton, and other aspects of the physical properties of the cellular environment.
- The principal targets of the pathway are the transcriptional co-activators YAP and TAZ (in mammals). Phosphorylation of YAP/TAZ causes their cytoplasmic sequestration and consequently inactivation of their transcriptional co-activator function.
- In the absence of this phosphorylation, YAP/TAZ are imported into the nucleus where they bind to inactive TEAD transcription factors (in mammals) prebound to multiple cell-type-specific enhancers, leading to transcriptional activation of these genes. Because of this, nonoverlapping sets of genes are regulated by the Hippo pathway in different cell types, depending on where TEAD factors are bound.
- In many cell types, activation of the Hippo pathway promotes apoptosis and inhibits cell cycling, limiting tissue growth. Inactivation of the Hippo pathway stimulates cell replication and inhibits apoptosis, stimulating tissue growth.
- Abnormal regulation of the Hippo pathway contributes to the uncontrolled growth of several types of human cancers and also to invasion of normal surrounding tissue and metastasis by cancers because of abnormal regulation of the actin cytoskeleton.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

autophagy
circadian clock (circadian oscillator)
effector proteins
feedback repression
glucagon
gluconeogenesis
GLUT4 storage vesicle (GSV)
heat-shock proteins (HSP)
heat-shock response
Hif-1α
Hippo pathway
homeostasis

insulin
mechanoreceptors
mTOR protein kinase
mTORC1
sensors
signal amplification
signal transduction pathways
SRE-binding proteins (SREBPs)
suprachiasmatic nucleus (SCN)

Review the Concepts

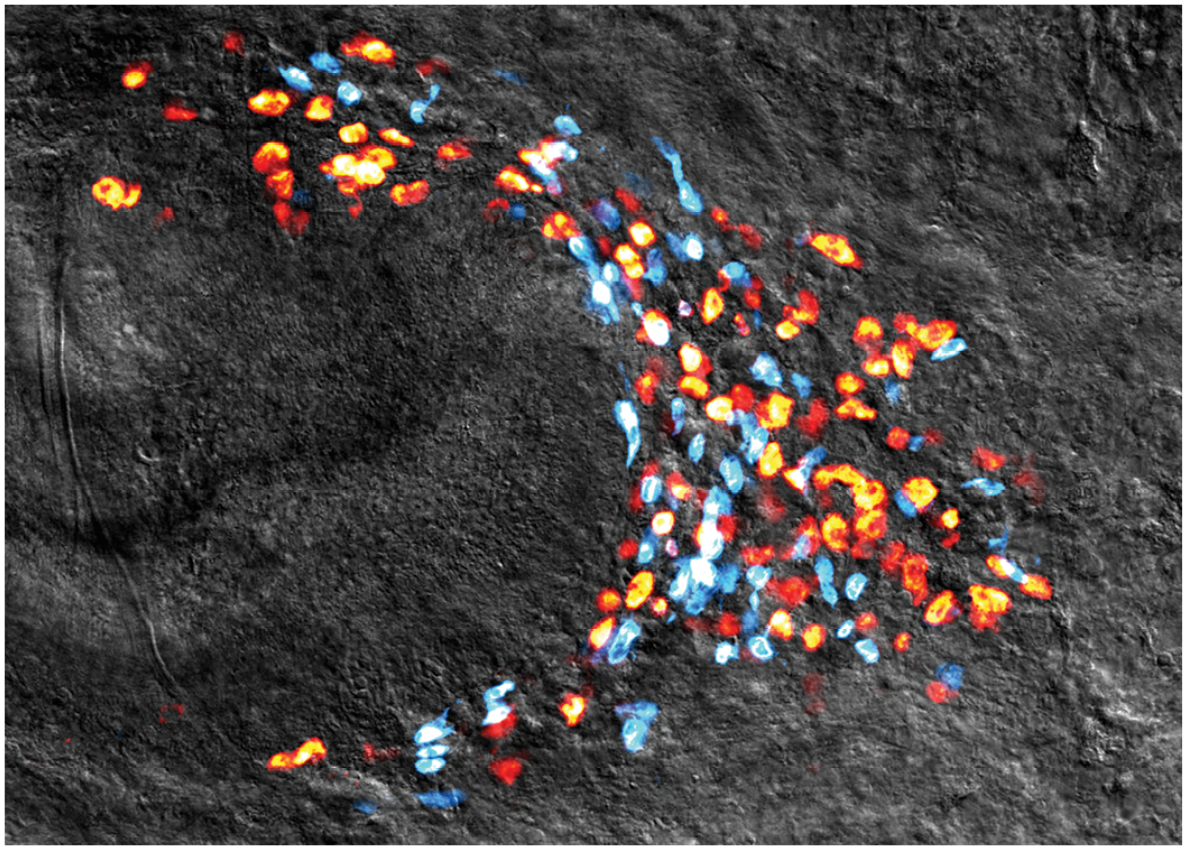
1. How are insulin and glucagon secretion in the pancreas regulated? How does each hormone affect blood glucose levels?
2. What is GLUT4, and how is it regulated by insulin?
3. What parts of the blood glucose regulation pathways are affected in type 1 diabetes? In type 2 diabetes?
4. List and briefly describe the anabolic signal transduction pathways activated by mTORC1. How do these pathways contribute to cell growth?
5. One of the signals that activates mTOR are levels of cytosolic amino acids. How does this pathway function?
6. To what class of proteins does Rheb belong? Where is Rheb found in cells? Explain the mechanisms that regulate Rheb and describe the downstream consequences when a cell accumulates active Rheb.

7. How do SCAP and SREBP proteins work to prevent the expression of genes containing sterol regulatory elements when sufficient cholesterol is present in the endoplasmic reticulum (ER) membrane?
8. Describe the changes that are triggered in the cell when the cholesterol levels in the ER membrane drop to less than 5 percent of the total ER lipids.
9. What is nSREBP? Where is it generated? What is its function?
10. The expression of what two types of proteins is induced in mammals by Hif-1α? How do these proteins contribute to survival under low-oxygen conditions?
11. How do an aspargine hydroxylase, a proline hydroxylase, and an E3 ubiquitin ligase keep Hif-1α in check under normal oxygen levels?
12. How are ERF proteins in plants kept in check at ambient oxygen levels?
13. Why are proteins so sensitive to very small changes in temperature?
14. What is the role of heat-shock proteins (HSPs) in the protein-folding process?
15. In unstressed cells, how do HSPs and heat-shock factor (HSF) proteins interact to keep the heat-shock response in check?
16. What are the key genes or proteins in the circadian clock in *Drosophila melanogaster?* How does a negative feedback loop play a role in the circadian clock?
17. Describe the circadian clock in cyanobacteria.

18. What is a zeitgeber? How would a zeitgeber affect an organism's circadian rhythms? What are some examples of a zeitgeber?
19. What role does the Hippo pathway play in flies and mammals?
20. Based on the experiments described in this section, explain how input about the extracellular matrix (ECM) is translated to the on or off state of the Hippo pathway. How is subcellular localization of pathway components involved?
21. How does the Hippo pathway play a role in differentiation of the inner and outer cells of the morula?

CHAPTER 22

Stem Cells, Cell Asymmetry, and Regulated Cell Death



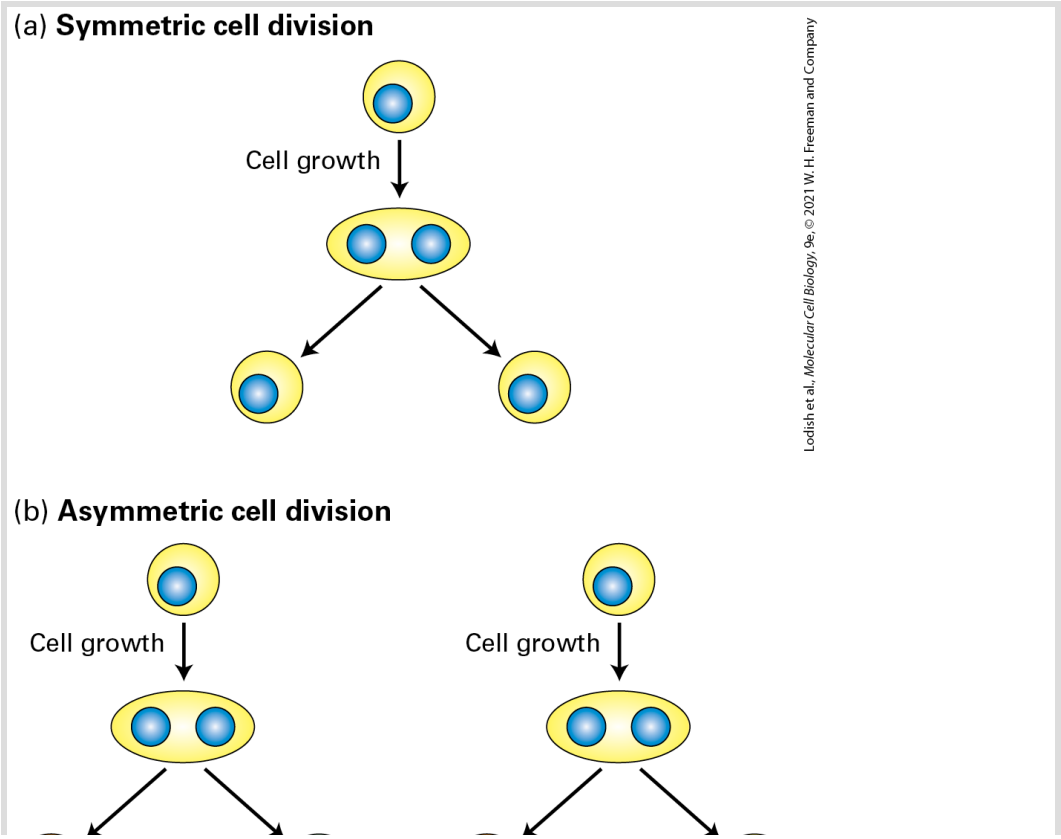
Daniel E. Wagner and Peter W. Reddien, MIT, Whitehead Institute.

Pluripotent stem cells called *neoblasts* provide the cellular basis for regeneration in planarian flatworms. Shown is a colony of neoblasts (yellow, orange, and red), all derived from a single neoblast 14 days after regeneration of the tail was initiated by amputation; differentiating cells (blue), also derived from the single neoblast, are shown as well.

OUTLINE

22.1 Early Mammalian Development, Embryonic Stem Cells, and Induced Pluripotent Stem Cells
22.2 Stem Cells and Niches in Multicellular Organisms
22.3 Mechanisms of Cell Polarity and Asymmetric Cell Division
22.4 Cell Death and Its Regulation

Many descriptions of cell division imply that the parent cell gives rise to two daughter cells that look and function exactly like the parent cell. In other words, they imply that cell division is **symmetric** and that the progeny have properties similar to those of the parent (Figure 22-1a). Many yeasts, fungi, and other single-celled eukaryotes do divide this way. Mature liver cells — hepatocytes — also divide symmetrically, each giving rise to two daughter hepatocytes.



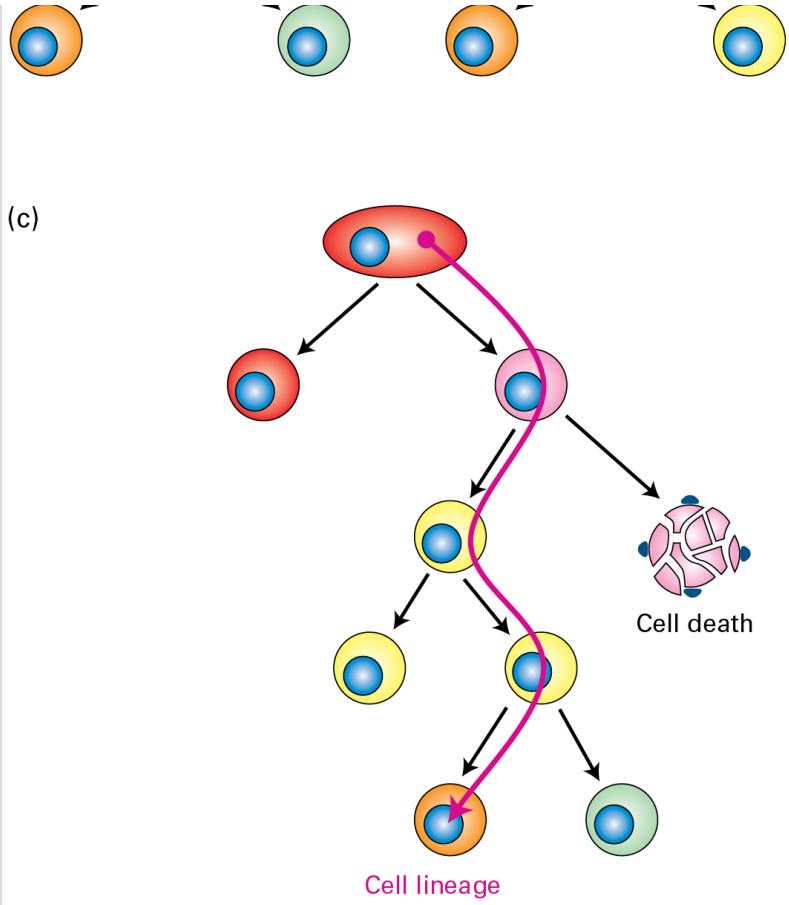


FIGURE 22-1 Overview of the birth, lineage, and death of cells. Following growth, daughter cells are born as the result of symmetric or asymmetric cell division. (a) The two daughter cells resulting from symmetric division are essentially identical to each other and to the parent cell. Such daughter cells can subsequently have different fates if they are exposed to different signals. (b) The two daughter cells resulting from asymmetric cell division differ from birth and consequently have different fates. In some cases (*left*), both daughter cells are different from each other and from the parent cell. In others (*right*), one daughter cell is essentially identical to the parent and the other assumes a different fate. Asymmetric division is common when the parent cell is a stem cell and one of the daughters is also a stem cell; this allows the number of stem cells (yellow) to remain constant while the other daughter generates other cells (orange) that divide and then mature into one or more differentiated cell types. (c) A series of symmetric and asymmetric cell divisions, called a *cell lineage*, gives birth to each of the specialized cell types found in a multicellular organism and is under tight genetic control.

But if this were always the case, none of the hundreds of differentiated cell types and functioning tissues present in complex multicellular plants and animals would ever be formed. Differences among cells can arise when two initially identical daughter cells diverge upon receiving distinct developmental or environmental signals. Alternatively, the two daughter cells may differ from birth, with each inheriting different portions of the parent cell (Figure 22-1b). Daughter cells produced by such **asymmetric cell division** may differ in size, shape, or protein composition, or their genes may be in different states of activity or potential activity. The differences in these internal signals confer different fates on the two cells. In certain asymmetric cell divisions, one of the daughters is similar to the parent cell and the other forms a different type of cell.

In multicellular organisms, the formation of working tissues and organs, during both development and cell replacement, depends on specific patterns of mitotic cell divisions. A series of such cell divisions akin to a family tree is called a **cell lineage**. A cell lineage traces the birth order of cells as they progressively become more restricted in their developmental potential and *differentiate* into specialized cell types such as skin cells, neurons, or muscle cells (Figure 22-1c).

The development of a new metazoan organism begins with the *egg*, carrying a set of chromosomes from the mother, and the **sperm**,

carrying a set of chromosomes from the father. These *gametes*, or sex cells, are haploid because they have gone through meiosis (see Chapter 19). In the process called *fertilization*, they combine to create the initial single cell, the zygote, which has two sets of chromosomes and is therefore diploid. During embryogenesis, the zygote undergoes numerous cell divisions, both symmetric and asymmetric, ultimately giving rise to an entire organism. As we will see later in the chapter, many of the early divisions of the nematode *Caenorhabditis elegans* follow a *mosaic development* strategy, in which all of the early cell divisions are asymmetric and each daughter cell gives rise to a discrete set of differentiated cell types (see Figure 22-26).

However, the focus of the first section of this chapter is early mammalian development, quite different from that of *Caenorhabditis,* and the generation of embryonic stem cells. Both mouse and human embryos pass through an eight-cell stage; each of these cells can still form every type of tissue — both the embryo itself and the extraembryonic tissues — making all eight cells **totipotent**. At the 16-cell stage, this is no longer true: some of the cells have become committed to particular differentiation paths.

A group of cells called the *inner cell mass* will ultimately give rise to all tissues of the embryo proper, and another set of cells, the extraembryonic trophectoderm, will form the placental tissue (see Figures 22-2 and 22-3). Cells such as those in the inner mass that can

generate all embryonic tissues, but not extraembryonic tissues, are called **pluripotent**. As we will learn in Section 22.1, cells of the inner cell mass can be cultured in defined media, forming **embryonic stem (ES) cells**. ES cells can be grown indefinitely in culture, where they divide symmetrically, so that each daughter cell remains pluripotent and can potentially give rise to all of the tissues of an animal. We will discuss the use of ES cells in uncovering the transcriptional network of gene expression underlying pluripotency as well as in forming specific types of differentiated cells for research purposes or, potentially, as replacement parts for worn-out or diseased cells in patients.

For many years, animal cell differentiation was thought to be unidirectional, but recent data reveal that differentiation can be reversed experimentally. Through recombinant expression of specific transcription factors, several types of specialized, differentiated cells can be converted into another type of differentiated cell. Strikingly, introducing just a small number of the transcription factors that control the pluripotency of ES cells, under defined conditions, can over time convert at least some types of **somatic cells** into **induced pluripotent stem (iPS) cells** that have properties seemingly indistinguishable from those of ES cells. We will learn that, like ES cells, iPS cells have profound utility for experimental biology and medicine.

Many types of **stem cells** are important during both metazoan development and adult life. They are unspecialized cells that can

reproduce themselves as well as generate specific kinds of more specialized cells (see Figure 22-1b). Their name comes from the image of a plant stem, which grows upward, continuing to form more stem, while also sending off leaves and branches to the side. In the second section of this chapter, we explore several types of stem cells that differ in the variety of specialized cell types they can form, including those that give rise to germ cells, intestinal cells, and the variety of cell types found in blood. Unlike ES cells, the stem cells in adults are **multipotent**: they can give rise to multiple types of differentiated cells found in the organism but not all of them.

Some stem cells can undergo symmetric divisions in which both of the daughters are stem cells, thus increasing the number of stem cells. Many stem cells in animals and plants undergo asymmetric divisions in which one of the daughter cells is a stem cell; thus the numbers of stem cells can remain constant or can increase during the early stages of the organism's life. Many cells have life spans much shorter than that of the organism as a whole and so need to be constantly replaced. In mammals, for instance, cells lining the intestine and phagocytic macrophages live for only a few days. Stem cells are therefore important not only during development but also for replacement of worn-out cells in adult organisms. However, the numbers of certain stem cells often decreases as the organism ages, limiting their ability to replace certain outworn cells or tissues.

The zygote is totipotent in that it has the capacity to generate every cell type in the body as well as the supportive placental cells that are

required for embryonic development, but because the zygote does not *self-renew* (make more of itself), it is not considered a stem cell.

We have already mentioned that the diversity of cell types in an animal requires asymmetric cell divisions in which the compositions and fates of the two daughter cells differ. This process requires the parent cell to become asymmetric, or **polarized**, before cell division, so that the cell contents are unequally distributed between the two daughters. This process of polarization is critical not only during development but also for the function of many types of differentiated cells. For example, transporting epithelial cells, such as those that line the intestine, are polarized, with their free apical surface facing the lumen to absorb nutrients and their basolateral surface contacting the extracellular matrix to transport nutrients toward the blood (see Figures 11-30 and 20-1). Other examples include cells that migrate up a chemotactic gradient (see Figure 18-54) and neurons, which have multiple dendrites extending from one side of the cell body that receive signals and a single axon extending from the other side that transmits signals to target cells (see Chapter 23). Thus the mechanisms that cells use to polarize are important and general aspects of their function. Not surprisingly, these mechanisms integrate elements of cell signaling pathways (see Chapters 15 and 16), cytoskeletal reorganization (see Chapters 17 and 18), and membrane trafficking (see Chapter 14). In the third section of this chapter, we discuss how cells become polarized as

well as the mechanism of asymmetric cell division for maintaining stem cells and generating differentiated cells.

Death — of both organisms and the cells within them — is an essential part of life; without carefully regulated death, there could not be life. In the last section of this chapter, we discuss the regulation of cell death, termed **programmed cell death**. We learn that apoptosis, one type of evolutionarily conserved programmed cell death, is absolutely crucial for the formation and maintenance of many tissues (see Figure 22-1c). We also learn about a second type of programmed cell death, necroptosis, that vertebrates employ to kill virus-infected cells and thus block the spread of infection. Cells also die from environmental stresses or because they are no longer needed by the organism. Precise genetic regulatory systems, each with checks and balances, control cell death — just as other genetic programs control cell division and differentiation.

These aspects of cell biology — cell birth, the establishment of cell polarity, and programmed cell death — converge with developmental biology, and they are among the most important processes regulated by the signaling pathways discussed in earlier chapters.

22.1 Early Mammalian Development, Embryonic Stem Cells, and Induced Pluripotent Stem Cells

Fertilization Unifies the Genome

It is remarkable that a mammalian sperm is ever able to reach and insert its nucleus into an egg; in humans, each sperm is competing with more than 100 million others for a single oocyte. After the first sperm succeeds in fusing its membrane with the oocyte membrane, a flux of calcium flows into the oocyte cytosol, spreading outward from the site of sperm binding. As in other regulated secretory pathways, one of the effects of the rise in calcium is to trigger fusion of vesicles located just under the plasma membrane of the egg, called *cortical granules*, with the plasma membrane, releasing their contents to the outside of the plasma membrane and forming a shielding fertilization membrane that blocks other sperm from binding to the egg. Finally the sperm nucleus enters the egg cytoplasm, and the egg and sperm nuclei fuse to create the diploid zygote nucleus.

Oocytes contain multiple mitochondria, and in mammals and many other species inheritance of mitochondrial DNA is exclusively

maternal; little if any sperm mitochondrial DNA enters the oocyte (see Chapter 12). Female-specific mitochondrial DNA inheritance has been used to trace maternal heritage in human history, for example, to follow early humans from their origins in Africa. The egg cytoplasm is also packed with *maternal mRNA*: transcripts of genes whose products are essential for the earliest stages of development. There is little or no transcription during oocyte meiosis and the first embryonic cleavages, so during this time the oocyte's mRNA is crucial for protein production.

Cleavage of the Mammalian Embryo Leads to the First Differentiation Events

The fertilized egg, or zygote, does not remain a single cell for long. Fertilization is quickly followed by *cleavage*, a series of cell divisions that take about one day each (Figure 22-2); these divisions happen before the embryo is implanted in the uterine wall. Initially, the cells are fairly spherical and loosely attached to one another. As demonstrated experimentally, each cell at the eight-cell stage is totipotent and has the potential to give rise to a complete animal when implanted into the uterus of a pseudopregnant animal (one treated with hormones to make her uterus responsive to embryos). It is likely that each of the cells in eight-cell human and other mammalian embryos are also totipotent and that totipotency is lost

during later cleavages, though the timing may be different from that of a mouse embryo.

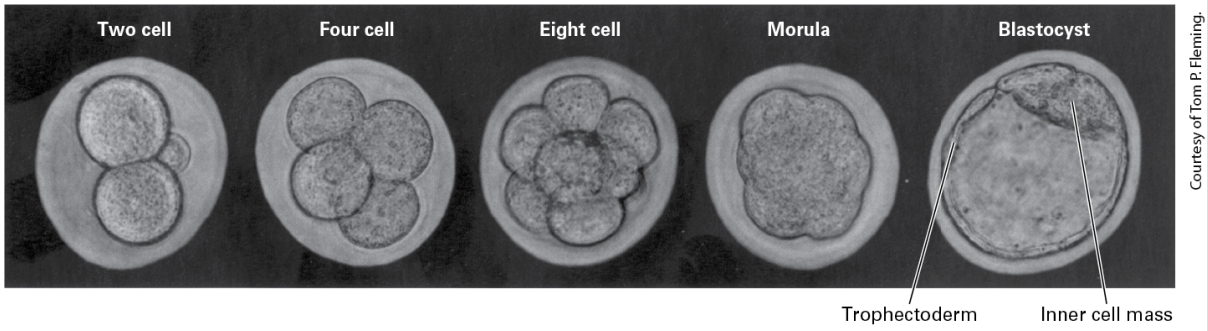


FIGURE 22-2 Early cleavage divisions in the mouse embryo. There is little cell growth during these early divisions, so the cells become progressively smaller. See text for discussion.

The eight-cell embryo undergoes compaction — a process that depends in part on induction of the cell-surface homotypic cell-adhesion protein E-cadherin (see Figure 20-14) that both causes substantial increases in the affinities of the blastomere cells for one another and polarizes the cells. Three days after fertilization, the eight-cell embryo divides to form the 16-cell *morula* (from the Greek for "raspberry"; see Figure 22-2); during this division one of the daughters of each blastomere ends up in an internal position of the 16-cell embryo whereas the other remains on the surface. The outer cells of the 16-cell embryo are fated to become the **trophectoderm (TE)** of the later 64-cell stage termed the **blastocyst** (see Figures 22-2 and 22-3) and will form extraembryonic tissues such as the placenta. The inner cells of the 16-cell embryo are fated to become the **inner cell mass (ICM)** of the blastocyst (just 10–15 cells in a mouse), which

gives rise to the embryo proper (Figure 22-3). During these divisions, fluid flows into an internal cavity called the *blastocoel*.

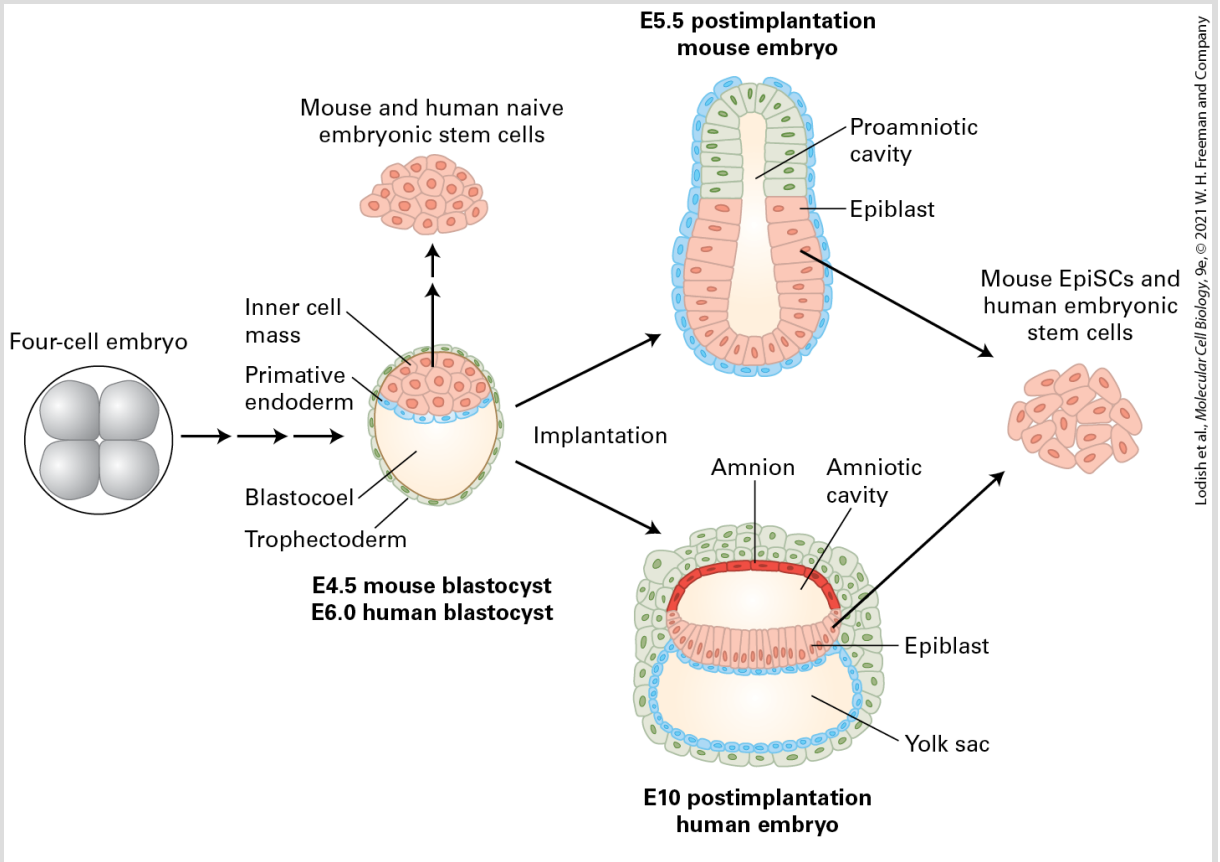


FIGURE 22-3 Schematic representation of mouse and human pre- and postimplantation embryos and the stem-cell lines that can be derived from them. In both mice and humans, a four-cell embryo normally develops into a blastocyst consisting of trophectoderm (TE) cells on the outside and inner cell mass (ICM) cells inside (see Figure 22-2); the ICM gives rise to cells that will form the extraembryonic primitive endoderm (PE) and the remaining cells differentiate into epiblast cells that will give rise to the embryo proper. After implantation into the wall of the uterus the detailed development of the mouse and human embryo differ, but in both cases the epiblast cells form the embryo proper; these cells and the earlier ICM cells can form embryonic stem cells when placed in culture. The extraembryonic cells form part of the placenta together with cells from the uterine wall. Extraembryonic tissues are shown in different shades of teal and epiblast derivatives in pink.

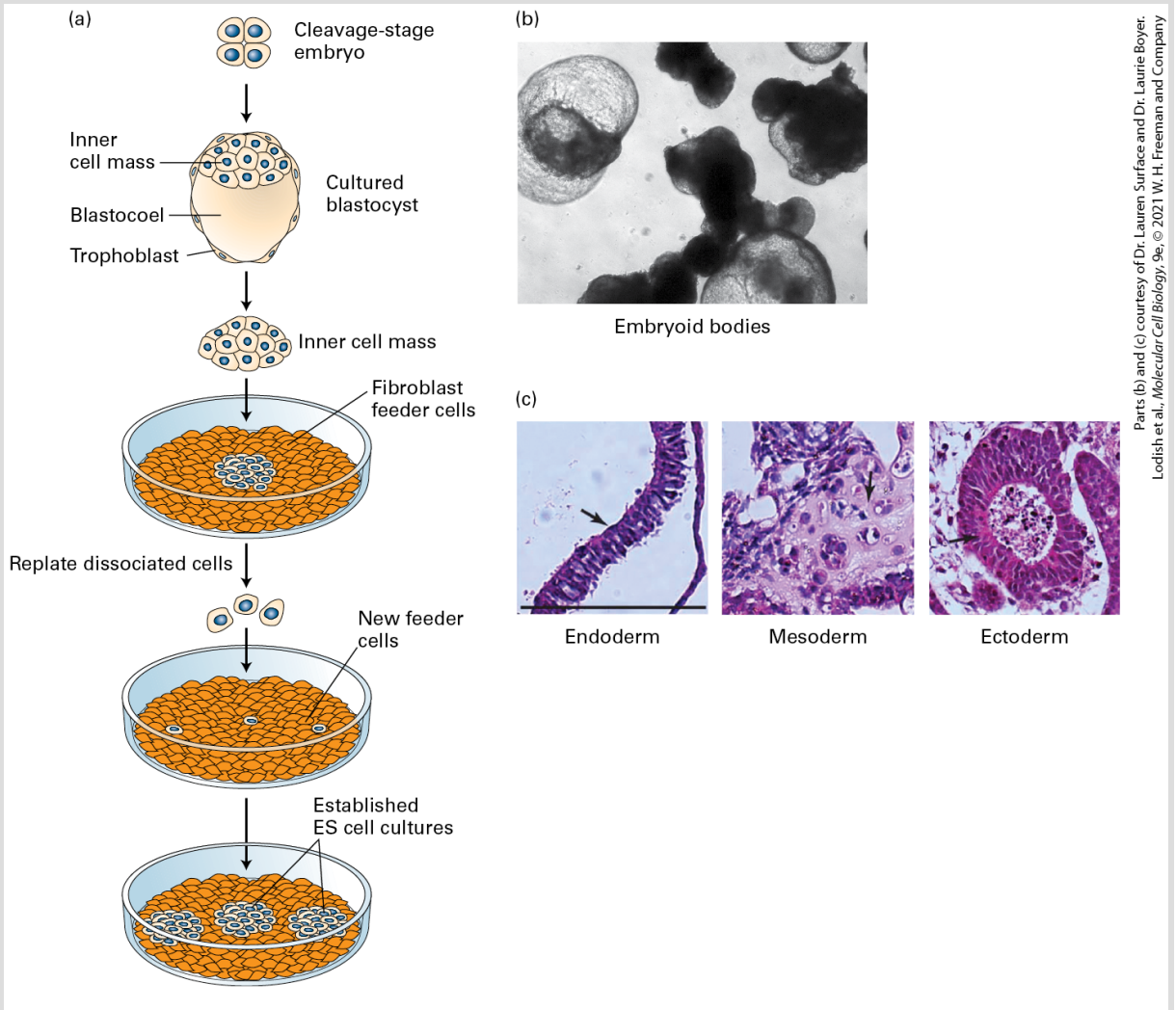
[Information from M. N. Shahbazi, E. D. Siggia, and M. Zernicka-Goetz, 2019, *Science* **364**:948.]

In the blastocyst, the ICM is found on one side of the blastocoel, while the TE cells form a hollow ball around the ICM and blastocoel. At this point, the TE cells are in an epithelial sheet, while the ICM cells are a loose mass that can be described as mesenchyme. *Mesenchyme*, a term most commonly applied to mesoderm-derived cells, refers to loosely organized and loosely attached cells. The fate of a cell in the early embryo — TE or ICM — is determined by the cell's location. If a labeled cell is placed on the outside of a very early embryo, it is likely to form extraembryonic tissues, while a cell placed inside an embryo is likely to form embryonic tissues. Both ICM and TE cells are stem cells: each starts its own distinct lineage and divides prolifically to produce diverse populations of cells. Gene expression measurements of each stage of early development show dramatic changes in which genes are expressed in ICM and TC cells and their descendants. Even these very early embryos use Wnt, Notch, and TGF-β signals to regulate gene expression (see Chapter 16).

Pluripotent Cells of the Inner Cell Mass Are the Source of ES Cells

ES cells can be isolated from the inner cell mass of early mammalian embryos and grown indefinitely in culture when attached to a feeder-cell layer that provides certain essential growth

factors (Figure 22-4a). As mentioned in the chapter introduction, cultured ES cells are pluripotent: they can differentiate into a wide range of cell types of the three primary germ layers, either in culture or after reinsertion into a host embryo. More specifically, mouse ES cells can be injected into the blastocoel of an early mouse embryo and the cell aggregate surgically transplanted into the uterus of a pseudopregnant female. The injected ES cells will participate in forming most, if not all, tissues of the resultant chimeric mice. Furthermore, the injected ES cells will often give rise to functional sperm and eggs that, in turn, can generate normal live mice.



Parts (b) and (c) courtesy of Dr. Lauren Surface and Dr. Laurie Boyer. Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 22-4 ES cells can be maintained in culture and can form differentiated cell types. (a) Human or mouse blastocysts are grown from cleavage-stage embryos produced by in vitro fertilization. The ICM or the developmentally later epiblast (see Figure 22-3) is separated from the surrounding extraembryonic tissues and plated onto a layer of fibroblast cells, which help to nourish the embryonic cells by providing specific growth factors. When individual cells are replated, they form colonies of ES cells, which can be maintained for many generations and can be stored frozen. ES cells can also be cultured without a fibroblast feeder layer if specific cytokines are added; leukemia inhibitory factor (LIF), for instance, supports growth of mouse ES cells by triggering activation of the Stat3 transcription factor; see J. S. Odorico et al., 2001, *Stem Cells* **19**:193. (b) Embryonic stem cells allowed to differentiate in suspension culture form multicellular aggregates termed *embryoid bodies*. (c) Hematoxylin- and eosin-stained sections of embryoid bodies that contain derivatives of all three germ layers formed from the ICM during embryogenesis.

Arrows in the images point to the following tissue types: (*left*) gut epithelium (endoderm), (*middle*) cartilage (mesoderm), and (*right*) neuroepithelial rosettes (ectoderm). Black bar = 100 μm.

In a more recent variation on these experiments, a host zygote is treated with drugs that transiently block mitosis so that it and the blastocyst it forms are tetraploid (with four copies of each chromosome), incapable of forming differentiated cells and tissues. When normal diploid ES cells are injected into such a blastocyst, all the cells in the live mice that are born after transplantation of the blastocyst aggregate derive from the donor ES cells. This finding is powerful evidence that single mouse ES cells are indeed pluripotent. Because ethical considerations and, in many countries, legal restrictions preclude similar transplantation experiments with human ES cells, formal proof that human ES are pluripotent is lacking.

Importantly, both human and mouse ES cells can differentiate into a wide range of cell types in culture. When cultured in suspension, ES cells form multicellular aggregates, called *embryoid bodies* (Figure 22-4b), that resemble early embryos in the variety of tissues they form. When embryoid bodies are subsequently treated with various combinations of growth factors or transferred to a solid surface, they produce a variety of differentiated cell types, including gut epithelia, cartilage, and neural cells (Figure 22-4c). Under other conditions, ES cells have been induced to differentiate in culture into precursors for multiple specific cell types, including blood cells

and pigmented epithelia; for this reason, ES cells have proved extremely useful in identifying the factors that commit a pluripotent cell to differentiating down a particular cell lineage.

What properties give these cells of the early embryo their remarkable plasticity? As we will see in the next section, a variety of actors play a role: DNA methylation, transcription factors, chromatin regulators, and micro-RNAs all affect which genes become active.

Multiple Factors Control the Pluripotency of ES Cells

During the earliest stages of embryogenesis, as the zygote begins to divide, both the paternal and maternal DNA become demethylated (see the discussion of DNA methylation in Chapter 8). This happens in part because a key maintenance methyl transferase, Dnmt1, is transiently excluded from the nucleus and in part because demethylase enzymes actively remove (or "erase") methylation marks from 5-methyl cytosine residues during early development. As a result, the pattern of DNA methylation is reset during the first few cell divisions, erasing earlier epigenetic marking of the DNA and creating conditions in which cells have greater potential for diverse pathways of development. Mice engineered to lack Dnmt1 die as early embryos with drastically undermethylated DNA. ES cells

prepared from such embryos are able to divide in culture, but in contrast to normal ES cells, cannot undergo in vitro differentiation.

ES cell properties are also critically dependent on the action of master transcription factors produced shortly after fertilization. The transcription factors Oct4, Sox2, and Nanog have essential roles in early development and are required for the specification of ICM cells in the embryo as well as for the specification of ES cells in culture. The expression of Oct4 and Nanog is exclusive to pluripotent cells such as the cells of the ICM and cultured ES cells. Sox2 is expressed in pluripotent cells, but its expression is also necessary in the multipotent neural stem cells that give rise exclusively to neuronal and glial cell types (discussed in Chapter 23). Genetic studies in mice suggest that these three regulators have distinct roles but function in related pathways to maintain the developmental potential of pluripotent cells. For example, disruption of Oct4 or Sox2 results in the inappropriate differentiation of ICM and ES cells into trophectoderm. Thus knowledge of the set of genes regulated by these transcription factors might reveal their essential roles during development.

The genes that are bound by these three transcription factors have been identified using chromatin immunoprecipitation experiments (see Chapter 8); each protein is found at more than a thousand chromosomal locations. The target genes encode a wide variety of proteins, including the Oct4, Nanog, and Sox2 proteins themselves, forming an autoregulatory loop in which each of these three

transcription factors induces its own expression as well as that of the others (Figure 22-5). These transcription factors also bind to the transcription-control regions of many genes encoding proteins and micro-RNAs important for the proliferation and self-renewal of ES cells.

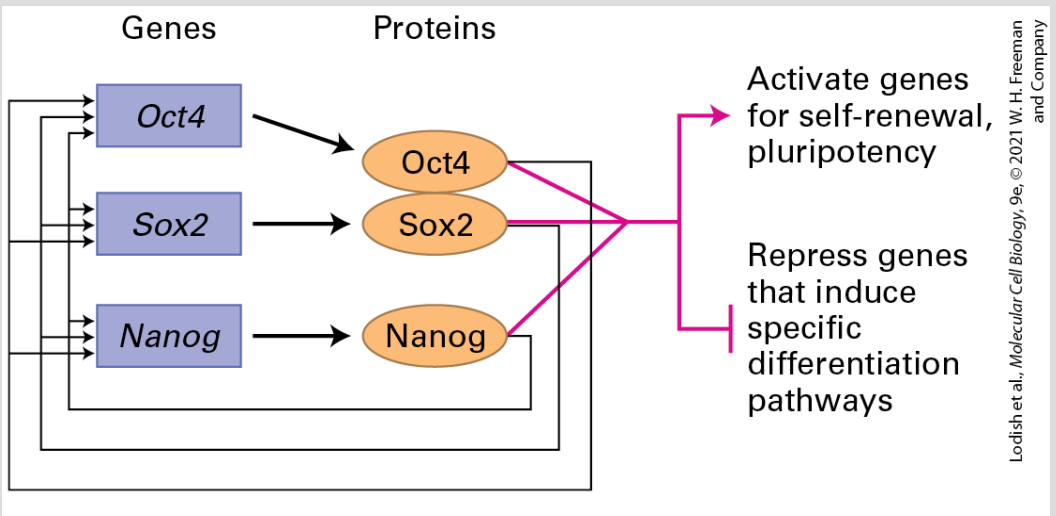


FIGURE 22-5 Basic transcriptional network regulating pluripotency of ES cells. The master transcription factors Oct4 and Sox2 form a dimer, and each of the three transcription factors, Oct4, Sox2, and Nanog, binds to its own promoter as well as to the promoters of the other two (black lines), forming a positive autoregulatory loop that activates transcription of each of these three genes. These transcription factors also bind to the transcription-control regions of many active genes encoding proteins and micro-RNAs important for the proliferation and self-renewal of ES cells as well as to those of many genes that are silenced in undifferentiated ES cells and that encode proteins and micro-RNAs essential for the formation of many differentiated cell types (magenta lines). See M. Li and J. C. Izpisua Belmonte, 2018, *Nat. Cell Biol.* **20**:382.

Several protein hormones are provided by feeder cells or added to culture media to prevent differentiation of ES cells. These hormones

include leukemia inhibitory factor (LIF), which activates Stat3; Wnt, which activates the β-catenin transcription factor; and bone morphogenetic protein 4 (BMP4), which activates the Smad1 transcription factor (see Chapter 16). In ES cells, these three transcription factors bind at multiple genomic sites co-occupied by Oct4, Nanog, and Sox2 proteins. Thus signaling pathways activated by cell-surface receptors are directly coupled to regulation of genes in the core pluripotency circuitry; this observation reinforces a point made in Chapter 16 that transcription factors activated by cell-surface receptors frequently bind at sites in the genome occupied by master transcription factors specific to that type of cell.

Chromatin regulators that control gene transcription (see Chapter 8) are also important in ES cells. In *Drosophila*, Polycomb group proteins form complexes to maintain gene repression states that have been previously established by DNA-binding transcription factors. Two mammalian protein complexes related to the fly Polycomb proteins, PRC1 and PRC2 (see Figure 8-48), are abundant in ES cells. Early mouse embryos lacking components of PRC2 display early developmental defects. The PRC2 complex acts by adding methyl groups to lysine 27 of histone H3, thus altering chromatin structure to repress genes. (Note that the methylation here is on an amino acid in a protein, a type of regulation distinct from the methylation of cytosine residues in DNA.) In ES cells, PRC1 and PRC2 both silence genes whose encoded proteins or micro-RNAs (miRNAs) would otherwise induce differentiation into

particular types of cells; the Polycomb proteins also maintain these genes in an epigenetic preactivation state such that they are poised to become activated later as part of the proper execution of specific developmental gene expression programs. Thus ES cells lacking PRC2 functions fail to differentiate properly.

Many other regulators play important roles in controlling gene expression and maintaining pluripotency during very early development. For example, the gene encoding the miRNA *let-7* is transcribed in ES cells, but the precursor RNA transcript is not cleaved to form the mature, functional miRNA. ES cells express a developmentally regulated RNA-binding protein termed *Lin28* that binds to the *let-7* precursor RNA and prevents its cleavage. *Let-7* is essential for differentiation of ES cells, and experimental expression of mature *let-7* miRNA in ES cells blocks their ability to undergo self-renewal. Thus repression of *let-7* processing by Lin28 is essential for pluripotency.

As we will see later in this section, the possibility of using embryonic stem cells therapeutically to restore or replace damaged tissue is fueling much research on how to induce them to differentiate into specific cell types. Apart from their possible benefit in treating disease, ES cells have already proved invaluable for producing mouse mutants useful in studying a wide range of diseases, developmental mechanisms, behavior, and physiology. Using the recombinant DNA techniques described in Chapter 6, one

can eliminate or modify the function of a specific gene in ES cells. The mutated ES cells can then be employed to produce mice with a gene knockout or modification (see Figure 6-40). Analysis of the effects of deleting or modifying a gene in this way often provides clues about the normal function of the gene and its encoded protein.

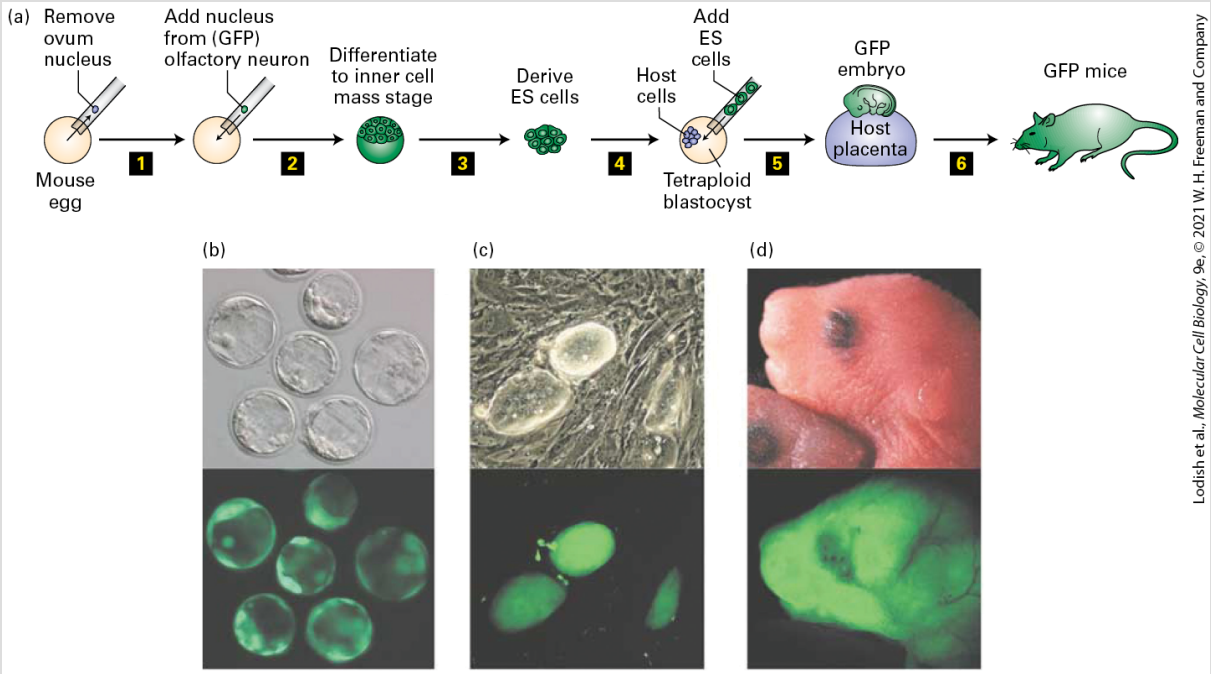
Animal Cloning Shows That Epigenetic Changes During Differentiation Can Be Reversed

Although different cell types may transcribe different parts of the genome, for the most part the genome is identical in all cells. Segments of the genome are rearranged and lost during development of the T and B lymphocytes of the immune system from hematopoietic precursors (see Chapter 24), but most somatic cells appear to have an intact genome, equivalent to that in the germ line. Evidence that at least some somatic cells have a complete and functional genome comes from the successful production of cloned animals by nuclear transfer. In this procedure, often called *somatic-cell nuclear transfer* (SCNT), the nucleus of an adult somatic cell is introduced into an egg whose nucleus has been removed; the manipulated egg, which contains the diploid number of chromosomes and is equivalent to a zygote, is then implanted into a foster mother. Outside of the few genes in mitochondrial DNAs that remain in the egg, the only source of genetic information to guide

development of the embryo is the nuclear genome of the donor somatic cell. The low efficiency of generating cloned animals by SCNT, combined with a high frequency of diseases such as obesity in the animals that are cloned, however, raises questions about how many adult somatic cells do in fact have a complete functional genome and whether those that do can be completely reprogrammed into a pluripotent undifferentiated state. Even the successes, such as the famous cloned sheep "Dolly," have some medical problems. If differentiated cells do have a physically complete genome, clearly only parts of it are transcriptionally active (see Chapter 8). A cell could, for example, have an intact genome, but be unable to properly reactivate specific genes due to inherited chromatin epigenetic states.

Further evidence that the genome of a differentiated cell can revert to having the full developmental potential characteristic of an ES cell comes from experiments in which olfactory sensory neurons — postmitotic cells that normally will not divide again — were genetically marked with green fluorescence protein (GFP) and then used as donors of nuclei (Figure 22-6). When the nuclei from differentiated olfactory sensory neurons were implanted into enucleated mouse oocytes, a small fraction of them developed into blastocysts that produced GFP. The blastocysts were used to derive ES cell lines, which were then injected into tetraploid blastocysts. These manipulated embryos, derived entirely from olfactory sensory neuron genomes, formed healthy mice in which all cells expressed the GFP. Thus, at least in some cases, the genome of a

terminally differentiated cell can be reprogrammed completely to form all tissues of a mouse.



EXPERIMENTAL FIGURE 22-6 Mice can be cloned by somatic-cell nuclear transfer from olfactory neurons. (a) Procedure for generating cloned ES cell lines using nuclei from olfactory sensory neurons and using them to generate cloned mice. Step **1**: A nucleus from an olfactory sensory neuron isolated from a mouse that expresses green fluorescent protein (GFP) only in its olfactory neurons was used to replace the nucleus of a mouse egg. Step **2**: The resultant zygote was cultured to the blastocyst stage. Step **3**: The ICM cells, all of which were clones of the original olfactory sensory neuron, and all of which expressed GFP, were used to generate lines of ES cells. Step **4**: These ES cells were injected into a tetraploid blastocyst. Step **5**: When the blastocyst was transplanted into the uterus of a pseudopregnant mouse, the tetraploid cells from the host blastocyst could form the placenta (violet), but not the embryo proper. Step **6**: Therefore, all of cells in the embryo proper and in the mouse that developed from it expressed GFP. (b–c) Bright-field (*top*) and fluorescence images (*bottom*) of (b) nuclear-transfer blastocysts and (c) the ES cells that were isolated from the ICM. (d) A control 12-hour-old mouse (*top*) and a mouse cloned from an olfactory sensory neuron, all of whose cells expressed GFP (*bottom*).

[Parts (b–d) reprinted with permission from Nature Publishing Group, from K. Eggan et al., 2004, "Mice Cloned from Olfactory Sensory Neurons," *Nature* **428**(6978):44–49; permission conveyed through Copyright Clearance Center, Inc.]

Somatic Cells Can Generate iPS Cells

Because of the inefficiency of SCNT, it remained unclear whether all types of somatic mammalian cells retained an intact genome and whether they could be induced to dedifferentiate into an ES cell–like state. Shinya Yamanaka, who won the 2012 Nobel Prize for this research, used retrovirus vectors to express a wide variety of transcription factors, singly and in combination, in cultured fibroblast cells. Remarkably, he found that both human and mouse fibroblasts could be reprogrammed to a pluripotent state, called an *induced pluripotent stem-cell state*, similar to that of an ES cell, by transformation with retroviruses encoding just four proteins: KLF4, Sox2, Oct4, and Myc. Note that two of these, Sox2 and Oct4, are two of the master transcription factors expressed in ES cells, as discussed previously. In addition to fibroblasts, keratinocytes (skin-forming cells) and other types of differentiated cells have been reprogrammed to iPS cells. Like ES cells, single mouse iPS cells can be experimentally introduced into a blastocyst and form all of the tissues of a mouse, including germ cells, attesting to the fact that somatic cells can indeed be reprogrammed to an embryonic pluripotent state.

Several other transcription factors, and even certain small organic molecules, can replace the Oct4 gene in the Yamanaka reprogramming cocktail. Subsequent analysis led to the discovery that each of these factors directly activates transcription of the endogenous (cellular) Oct4 gene, leading to induction of pluripotency. Thus it was hypothesized that, over time, forced expression of transcription-factor genes activates expression of many cellular genes, including those encoding Oct4 and other pluripotency proteins; over the course of several weeks, this activation reprograms at least some of the somatic cells to an ES-like state. To experimentally establish the point that activation of endogenous genes leads to reprogramming to an ES-like state, cultured keratinocytes were repeatedly transfected with synthetic mRNAs encoding the four canonical Yamanaka transcription factors, KLF4, Sox2, Oct4, and Myc. These cultured cells generated normal iPS cells that had no trace of any of the exogenously added mRNAs, attesting to the reprogramming of keratinocytes into iPS cells by inducing expression of only normal cellular genes.

In fibroblasts, the chromatin of most pluripotency-associated genes is inaccessible to transcription-factor binding, primarily due to the repressive histone H3 lysine 9 trimethylation ($H3K9Me_3$) mark. Among the genes activated by Oct4 are two that encode H3K9 demethylases, which remove these repressive chromatin marks and, over time, result in activation of pluripotency genes. Consistent with this notion, expression of these H3K9 demethylases increases during reprogramming, and their knockdown inhibits efficient iPS-

cell generation. Indeed, reprogramming involves other major changes in epigenetic modifications as well, including DNA methylation and several other types of histone modifications that serve to repress or allow potential activation of hundreds of genes.

Among the genes up-regulated during reprogramming is LIN28. As discussed above, LIN28 promotes pluripotency by blocking biogenesis of members of the *let-7* family of micro-RNAs. Expression of telomerase (see Figure 7-40) is also induced during reprogramming; telomerase restores the normal length of telomeres that are shortened during the many rounds of cell divisions required for certain types of differentiated cells to form. Cellular reprogramming to pluripotency is also accompanied by a rewiring of metabolic pathways: a metabolic shift from a highly respiratory metabolism characteristic of most types of differentiated cells, which depend on mitochondria for ATP production, to a highly glycolytic carbon flux characteristic of ICM cells in a normal embryo.

Patient-Specific iPS Cells Can Be Used to Develop Potential Treatments for Many Diseases

Because iPS cells can be derived from somatic cells of patients with difficult-to-understand diseases, they have already proved

invaluable in uncovering the molecular and cellular basis of several afflictions (Figure 22-7). Consider *amyotrophic lateral sclerosis (ALS)*, often called *Lou Gehrig's disease*, a fatal disease in which the motor neurons that connect the spinal cord to the muscles of the body progressively die off, causing muscle weakness and death, limb paralysis, and ultimately death due to respiratory failure. There is no cure.

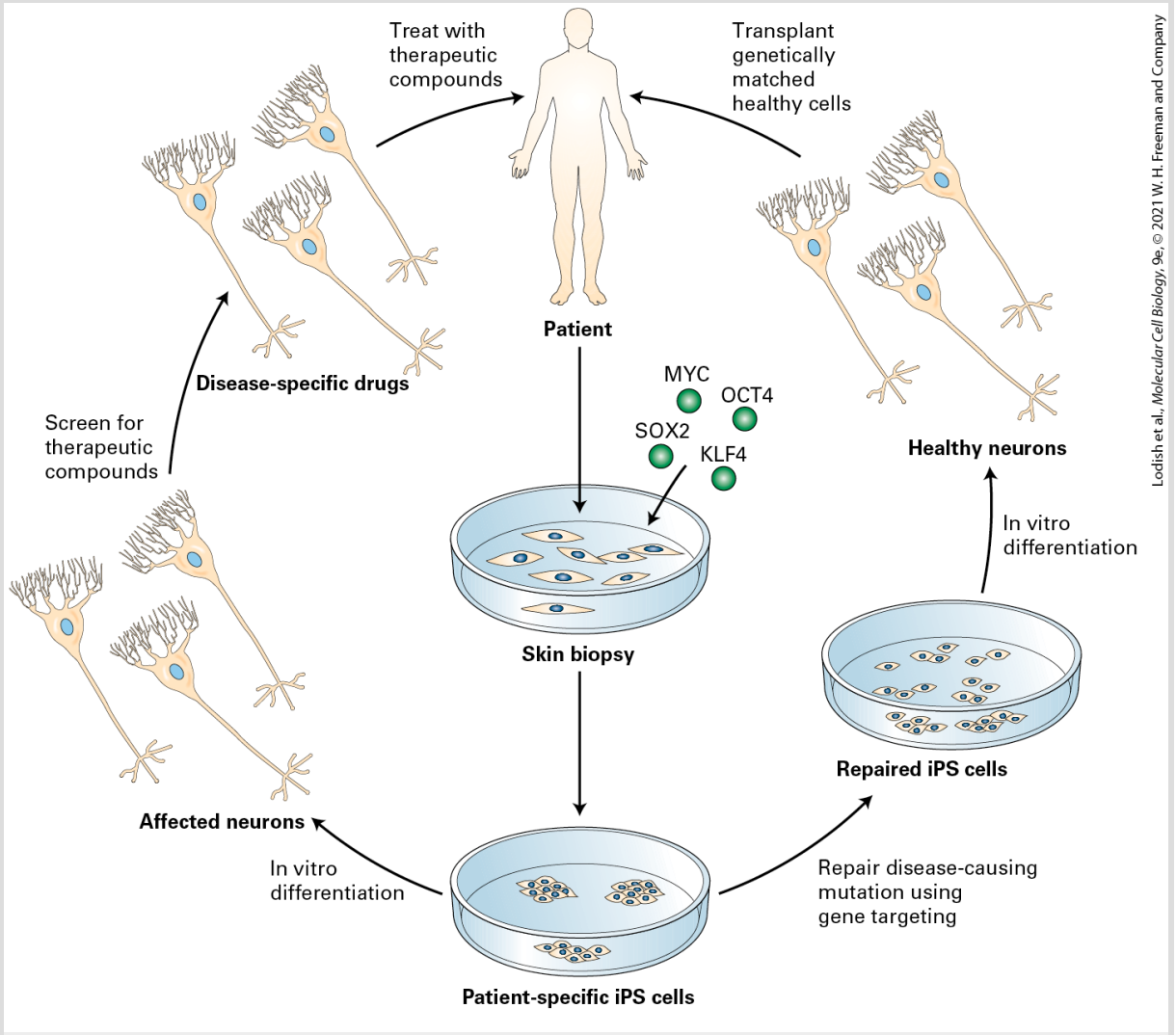


FIGURE 22-7 Medical applications of iPS cells. In this example, the patient has a neurodegenerative disorder caused by abnormalities in certain nerve cells (neurons).

Patient-specific iPS cells — in this case derived by recombinant expression of the four Yamanaka transcription factors in cells isolated from a skin biopsy — can be used in one of two ways. In cases in which the disease-causing mutation is known (e.g., familial Parkinson's disease), gene targeting could be used to repair the DNA sequence (*right*). The gene-corrected patient-specific iPS cells would then undergo directed differentiation into the affected neuronal subtype (e.g., midbrain dopaminergic neurons) and be transplanted into the patient's brain (to engraft the nigrostriatal axis). Alternatively, directed differentiation of the patient-specific iPS cells into the affected neuronal subtype (*left*) will allow the patient's disease to be modeled in vitro, and potential drugs can be screened, aiding in the discovery of novel therapeutic compounds.

[Information from D. A. Robinton and G. Q. Daley, 2012, *Nature* **481**:295.]

In approximately 10 percent of patients, the disease is dominantly inherited (familial ALS), but in 90 percent of patients, there is no apparent genetic linkage (sporadic ALS). An analysis of the underlying causes of the disease at a molecular and cellular level was impossible for many years because one cannot simply extract neurons or the surrounding glial cells (see Chapter 23) to analyze or culture from living humans. In several studies, iPS cells derived from the skin cells of elderly patients with these and other familial and sporadic forms of the disease were successfully differentiated in culture to form motor neurons; this success demonstrated the feasibility of leveraging the self-renewal of iPS cells to generate a potentially limitless supply of the cells specifically affected by ALS. Early studies showed that motor neurons bearing several types of ALS mutations were hyperexcitable, generating more of the electrical signals called *action potentials* (see Chapter 23) than normal. Using these differentiated neurons, researchers screened

thousands of small organic molecules, including many approved as drugs for treatment of other unrelated diseases, for those that could reverse the abnormalities in the ALS iPS cell–derived motor neurons. Several were identified and are in clinical trials to see if they can slow or stop the devastating effects of ALS. These experiments illustrate the value of iPS and ES cells in generating cell culture models of many types of difficult-to-study human diseases that can be used to screen for drugs that could treat many as yet untreatable afflictions. ■

ES and iPS Cells Can Generate Functional Differentiated Human Cells

Neurons and glial cells, as well as other cell types derived from human iPS cells have been implanted into mice with some promising results. Specifically, mouse stem cell–derived cardiomyocytes (heart muscle cells) can correct heart arrhythmias in mice; certain glial cells — oligodendrocytes — show promise in aiding recovery from experimental spinal injury in mice; and retinal epithelial cells can partially correct defects in mouse models of blindness. It is not yet known whether similar differentiated cells derived from human iPS cells can function when transplanted into humans.

One recent advance — the generation of normal insulin-secreting β islet cells from human iPS and ES cells — shows promise for treatment of both type 1 and type 2 diabetes. Type 1 diabetes results from autoimmune destruction of pancreatic β cells, whereas the more common type 2 diabetes results from insulin resistance in liver and muscle (see Figures 21-1 and 21-2), eventually leading to dysfunction and death of β cells. Patients who receive transplants of human islets from cadavers can be made insulin independent for 5 years or longer, but this approach is limited because of the scarcity and quality of donor islets; the possibility of an unlimited supply of human β cells from stem cells could potentially extend this therapy to millions of new patients.

One key to this successful generation of β cells was employing successive treatment with different combinations of growth factors that stimulated iPS or ES cells to traverse the multiple steps of the normal embryonic developmental sequence by which the progeny of undifferentiated ICM cells form mature functional β cells (Figure 22-8a). The so-called *SC-β cells* that resulted have a structure very similar to that of normal β islet cells, including secretory granules filled with almost crystalline insulin; they also secrete normal amounts of insulin in response to elevation of the glucose level in their culture medium. Shortly after their transplantation into mice, these cells secrete human insulin into the blood in a glucose-regulated manner. Most important, after transplantation of these cells into immunocompromised diabetic mice, their high glucose

levels are lowered to normal (Figure 22-8b), indicating the potential use of these β islet cells — which can be produced in culture in essentially unlimited numbers — for the treatment of diabetes. Screening to identify new drugs that improve β cell function, survival, or proliferation can also make use of such a uniform supply of stem cell–derived β cells. As detailed in Chapter 21 (Figure 21-1), normal islets also contain α cells that secrete glucagon as well as other cell types; one goal is to generate from iPS cells islet-like cell aggregates that contain both α and β cells.

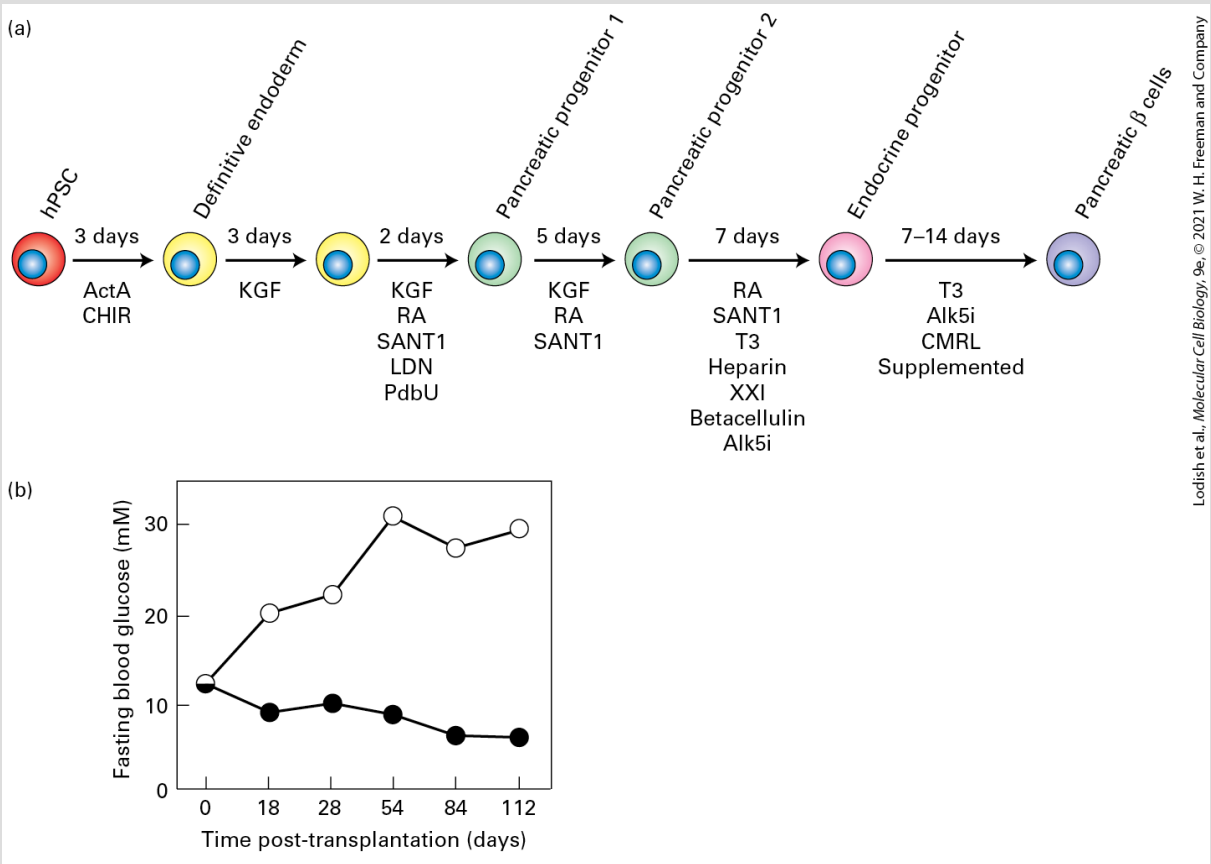


FIGURE 22-8 Production of normal insulin-secreting β islet cells from human iPS or ES cells. (a) Schematic of directed differentiation of human ES or iPS cells into insulin-secreting β islet cells. Clusters of a few hundred human ES or iPS cells were sequentially cultured in

media containing the indicated growth factors for the indicated number of days to first produce definitive endoderm cells, then a series of embryonic pancreatic progenitor cells, then pancreatic endocrine progenitors, and finally stem cell–derived insulin-producing β islet cells (termed *SC-β cells*). Abbreviations indicate the multiple signaling molecules that are employed: Act A, activin A; CHIR, GSK3 inhibitor; KGF, keratinocyte growth factor; RA, retinoic acid; SANT1, Sonic Hedgehog pathway antagonist; LDN, a BMP type 1 receptor inhibitor; PdbU, a protein kinase C activator; Alk5i, Alk5 receptor inhibitor II; T3, triiodothyronine, a thyroid hormone; XXI, γ-secretase inhibitor; betacellulin, an EGF family member. (b) These experiments used a strain of diabetic mice with a mutation in the insulin gene as well as in several immune-system genes such that the animals did not reject transplants of human tissue. Previous work had shown that the elevated glucose levels in these mice could be restored to normal by transplantation with human pancreatic islets. In this experiment, mice were transplanted with SC-β cells (black circles) or a similar number of control pancreatic progenitor cells (white circles). At the start of the experiment, the average blood glucose level in these mice was about 11 mM, well above the normal 5 mM. The average blood glucose level in the control mice rose continuously to about 30 mM, indicating severe diabetes, while in the mice transplanted with the human SC-β cells, blood glucose dropped to nearly the normal 5 mM.

[Part (b) Data from F. W. Pagliuca et al., 2014, *Cell* **159**:428.]

The coming years are certain to see the development of many other types of differentiated cells from human iPS cells that can be used as replacement parts for a variety of maladies. Indeed, fully functional mouse oocytes have been generated in culture from mouse ES cells and also from induced pluripotent stem cells derived from both embryonic fibroblasts and adult tail tip fibroblasts (Figure 22-9). When subjected to in vitro fertilization with wild-type sperm, all of these in vitro–generated oocytes developed to two-cell embryos. Remarkably, several of these two-cell embryos transferred

to pseudopregnant females were successfully delivered as viable pups.

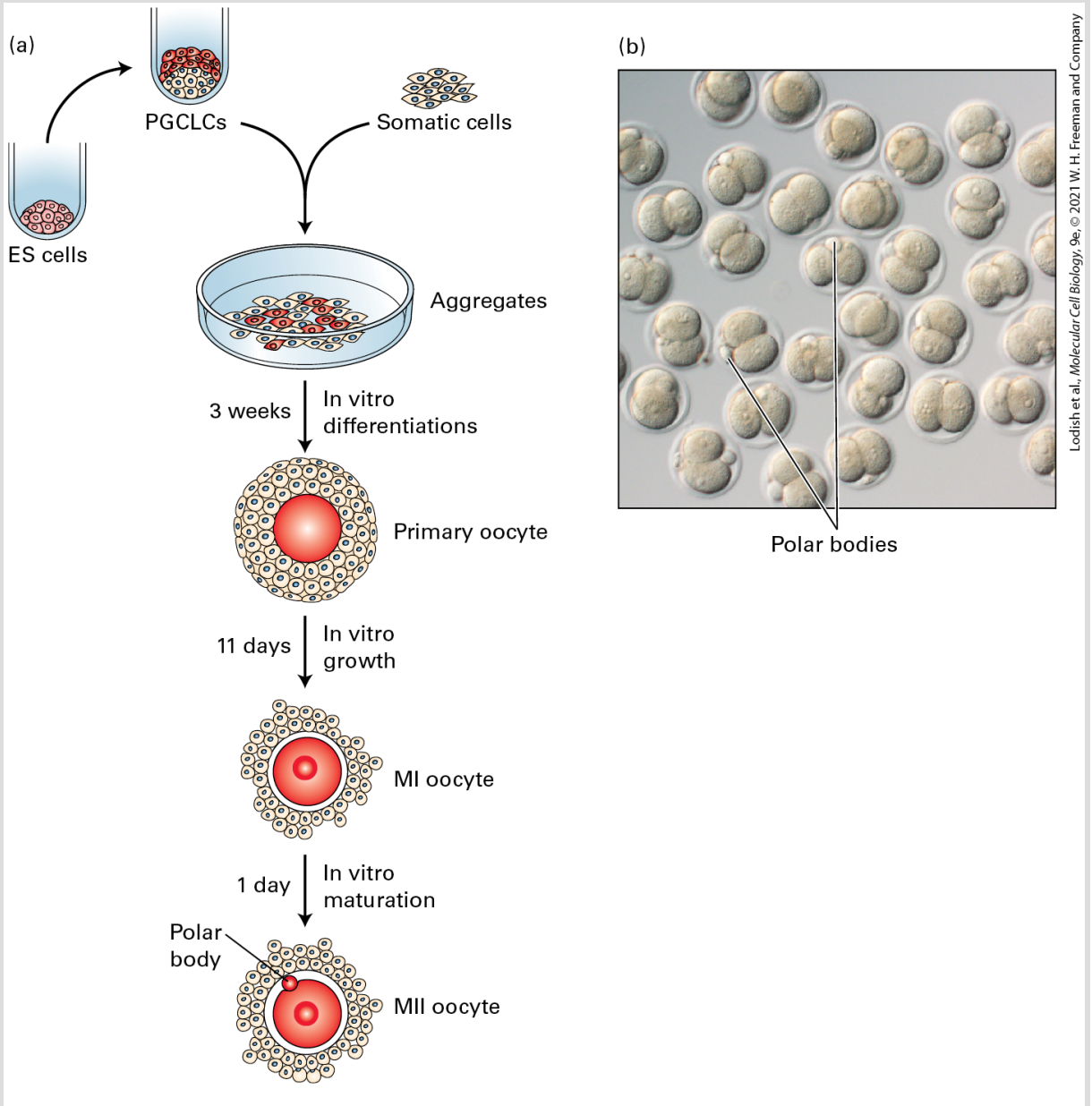


FIGURE 22-9 Production of functional mouse oocytes from pluripotent stem cells. (a) A schematic of oocyte production in vitro. Female mouse ES cells were first differentiated in culture to form primordial germ cell-like cells (PGCLCs) and then cultured together with female gonadal somatic cells isolated from a Day 12.5 mouse embryo to form primitive ovary-like aggregates. After 2–3 weeks in culture, primary oocytes had formed in which the

developing oocytes were surrounded by stromal cells; many oocyte progenitors were undergoing meiosis. After further culture, mature haploid oocytes — metaphase II (MII) — had formed. (b) After the final MII stage of differentiation, in vitro–generated oocytes were subjected to in vitro fertilization with wild-type sperm; all of these in vitro–generated oocytes were fertilized and developed to two-cell embryos. Note the polar bodies, small haploid cells that are the other products of meiosis and do not have the ability to be fertilized.

[Part (b) republished with permission from Springer Nature, from O. Hikabe et al., 2016, "Reconstitution In Vitro of the Entire Cycle of the Mouse Female Germ Line," *Nature* **539**:299–303; permission conveyed through Copyright Clearance Center, Inc.]

Many important questions must be answered, however, before the feasibility of using differentiated cells or eggs derived from human ES or iPS cells for therapeutic purposes can be assessed adequately. For instance, when undifferentiated human or mouse ES or iPS cells are transplanted into an experimental mouse, they form teratomas, tumors that contains masses of partially differentiated cell types. Thus it is essential to ensure that all of the ES or iPS cells used to generate an implant have indeed undergone differentiation and have lost their pluripotency and their ability to induce teratomas or cause other problems. Many groups are trying to generate fully functional oocytes from human ES or iPS cells as a treatment for certain types of female infertility. With the prospect of making babies from in vitro–derived oocytes on the horizon, combined with the possibility of editing the genomes of the ES cells that give rise to these oocytes, it is incumbent on society at large to grapple with the complex ethical issues at stake. Scientists and nonscientists alike should prioritize and expedite these debates.

KEY CONCEPTS OF SECTION 22.1

Early Mammalian Development, Embryonic Stem Cells, and Induced Pluripotent Stem Cells

- In asymmetric cell division, two different types of daughter cells are formed from one parent cell. In contrast, both daughter cells formed in symmetric cell divisions are identical but may have different fates if they are exposed to different external signals (see Figure 22-1).
- Fusion of a haploid sperm and haploid egg nucleus generates a diploid zygote. The initial divisions of the mammalian embryo yield equivalent totipotent cells (see Figure 22-2), but subsequent divisions yield the first differentiation event, the separation of the trophectoderm from the inner cell mass (see Figure 22-3).
- The inner cell mass is the source of the embryo proper as well as of embryonic stem cells (ES cells).
- Cultured ES cells are pluripotent, capable of giving rise to all differentiated cell types of the organism with the exception of extraembryonic tissues. They are useful in the production of genetically altered mice and offer the potential for therapeutic uses.
- The pluripotency of ES cells is controlled by multiple factors, including the state of DNA methylation, chromatin regulators, certain micro-RNAs, and the transcription factors Oct4, Sox2, and Nanog.
- Animal cloning establishes that cell differentiation can be reversed.
- Induced pluripotent stem (iPS) cells can be formed from somatic cells by expression of combinations of key transcription factors, including KLF4, Sox2, Oct4, and Myc.
- As exemplified by ALS, differentiated cells produced in culture from human iPS cells can be used to understand the underlying cause of a disease as well as to screen drugs that could be used to treat the disease.
- β islet cells produced in culture from human iPS cells secrete insulin normally in response to an elevation of glucose in the media and reverse the high glucose levels in diabetic mice.
- Fully functional mouse oocytes can be produced in culture from mouse ES and iPS cells.

22.2 Stem Cells and Niches in Multicellular Organisms

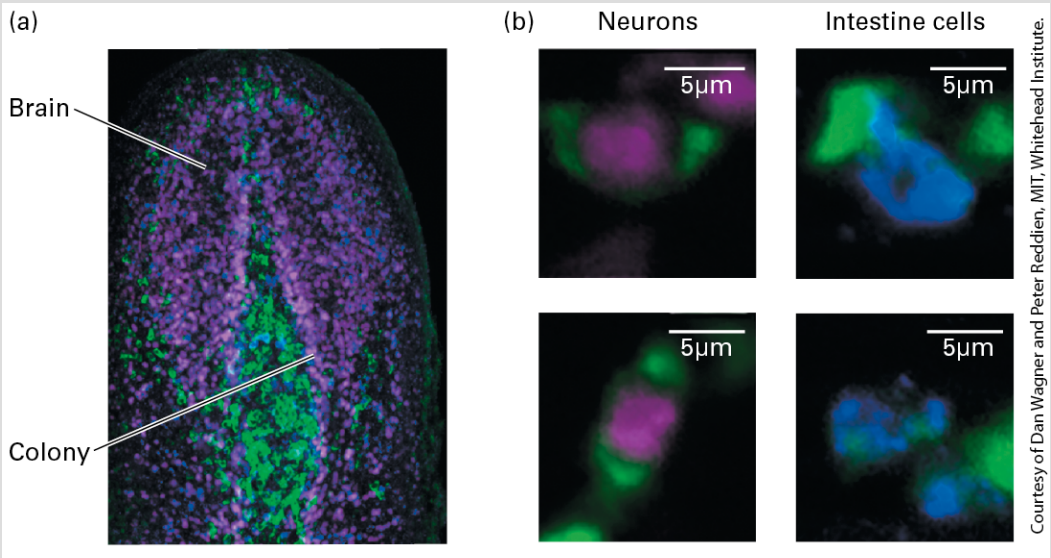
Many types of differentiated cells are sloughed from the body or have life spans that are shorter than that of the organism. Disease and trauma can also lead to losses of differentiated cells. Since most types of differentiated cells do not divide, they must be replenished from nearby somatic stem-cell populations. In vertebrates and most invertebrates, such stem cells, in contrast to pluripotent ES cells, are *unipotent* or multipotent in that they can give rise to one or several, but not all, of the cell types found in the organism. Postnatal (adult) vertebrate animals contain stem cells for many tissues, including the blood, intestines, skin, ovaries, testes, and muscle. Even some parts of the adult brain, where little cell division occurs, have populations of neuronal stem cells (see Chapter 23). In striated muscle, stem cells are most important in healing, as relatively little cell division occurs at other times. Some other cell types, such as liver cells (hepatocytes) and insulin-producing β islet cells, reproduce mainly by division of already differentiated cells, as exemplified by regeneration of the liver when large pieces are surgically removed.

Adult Planarians Contain Pluripotent Stem Cells

We noted in Chapters 1 and 16 that small body segments of planaria can regenerate whole animals. Regeneration was known to require a population of proliferating stem cell–like cells, termed **neoblasts**, that are present throughout the adult body, but a key question was whether regeneration is accomplished, as in many animals with this capability, by the collective activity of multiple lineage-restricted stem or progenitor cells, or whether pluripotent stem cells are involved. Recent experiments showed that adult planarians contain lineage-restricted neoblasts as well as pluripotent stem cells, termed *clonogenic neoblasts*, or *cNeoblasts*.

The key studies used gamma irradiation to inhibit most or all cell division in adult planarians; the treated animals could not regenerate and suffered massive tissue loss because of failed replacement of aged, differentiated cells. The few functional proliferating neoblast cells remaining after irradiation could be identified by a marker gene termed *smedwi-1*. Several days after irradiation, individual neoblasts formed colonies of *smedwi-1*-positive cells that contained multiple types of differentiated body cells, and it was hypothesized that this *smedwi-1*-positive subpopulation of neoblasts was pluripotent (Figure 22-10). To test this hypothesis, single neoblasts were transplanted into lethally irradiated planaria that lacked all of their own neoblasts. Remarkably, several transplant recipients lived past 7 weeks and regenerated, from the single transplanted cell, neuronal, intestinal, and other differentiated cell types that were distributed throughout

the body. The animals eventually regained feeding behavior and regenerated complex tissues, including eyes. These experiments indicated that at least some of the neoblast stem cells in adult planarians are indeed pluripotent, providing a cellular basis for the remarkable regenerative abilities of planarians. Despite much effort, no pluripotent stem cells have ever been reliably identified in any adult vertebrate organism.



EXPERIMENTAL FIGURE 22-10 Broad differentiation potential of individual neoblasts in planarians. Planarians were subjected to subtotal irradiation, which leaves few surviving neoblasts, but one surviving neoblast grew into a colony of neoblasts in the head. (a) Neoblasts are labeled in green with an antibody to the SMEDWI-1 protein. The *smedwi-1* gene is specifically transcribed in neoblasts, and transcription stops when a neoblast progeny cell stops dividing and differentiates. The SMEDWI-1 protein produced in neoblasts slowly degrades during differentiation into mature cell types, but enough remains to detect the progeny of this neoblast. (b) Double-labeling with the SMEDWI-1 antibody and fluorescent in situ hybridization with RNA probes for markers for differentiated cells therefore detects newly produced differentiated cells derived from the single neoblast. Neurons are labeled in magenta with an RNA probe to the gene *chat* and intestine cells are labeled in blue with an RNA

probe to the gene *gata4/5/6*. Anterior is up. See D. E. Wagner, I. E. Wang, and P. W. Reddien, 2011, *Science* **332**:811.

Multipotent Somatic Stem Cells Give Rise to Both Stem Cells and Differentiating Cells

The most common type of stem cells in adult metazoans, multipotent somatic stem cells, give rise to the specialized cells composing body tissues. The two critical properties of these stem cells that together distinguish them from all other cells are the ability to reproduce themselves during many cell divisions (self-renewal) and the ability to generate progeny of more restricted potential. Multipotent somatic stem cells have two other key properties (Figure 22-11):

1. They can give rise to multiple types of differentiated cells. In this sense, they differ from **progenitor cells** (also called *precursor cells*), which generally give rise to only one or two types of differentiated cells. A stem cell has the capability of generating a number of different cell types, but not all cell types: it is not pluripotent like an ES cell. For instance, a multipotent hematopoietic (blood) stem cell will form more of itself plus multiple types of blood and immune cells, but never a skin or a liver cell.

2. They are stem cells in that they are undifferentiated; in general, they do not express mRNAs or proteins characteristic of the differentiated cell types formed by their descendants. Stem cells of a particular type generally appear during embryonic development and rapidly increase in number; their progeny produce the many types of differentiated cells that make up the embryo and juvenile. Once adulthood is reached, the number of stem cells remains relatively constant and often declines with age. In that sense, stem cells are often said to be immortal, although no single stem cell survives for the life of the animal. Indeed, when pushed to divide more frequently than normal by chronic tissue injury, repeated rounds of chemotherapy, or genetic defects that impair genomic integrity, stem cells consistently exhibit a finite replicative capacity.

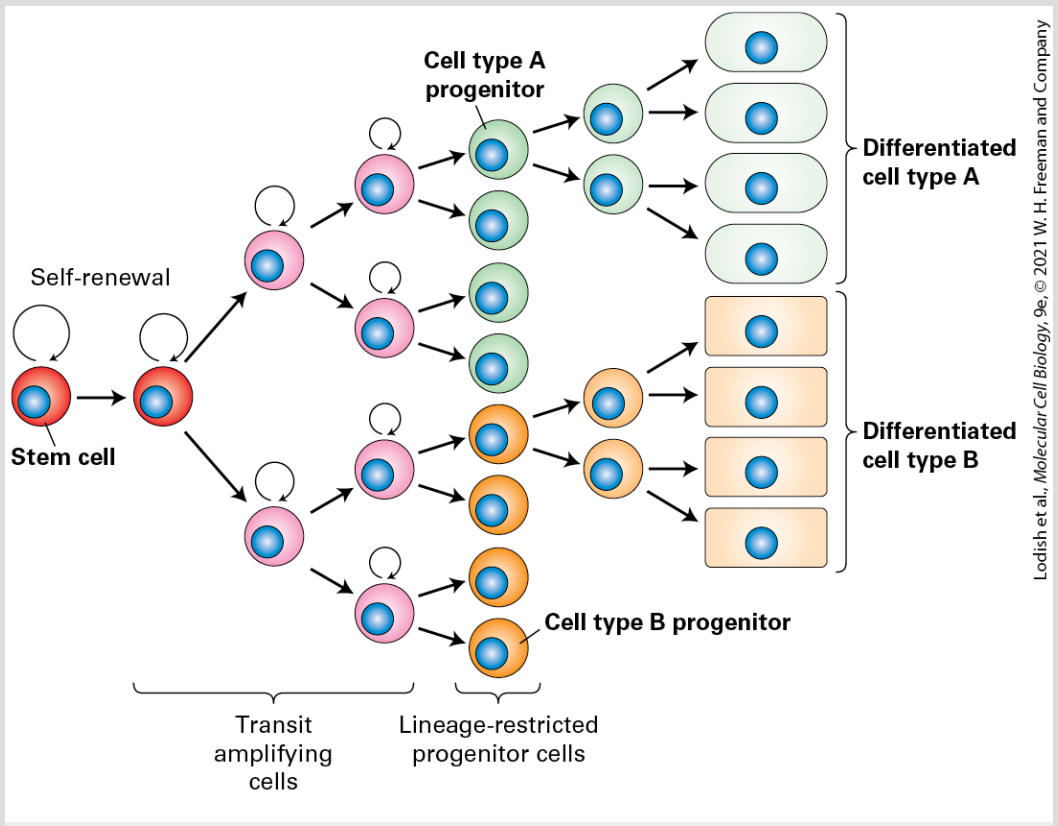


FIGURE 22-11 The pathway from stem cells to lineage-restricted progenitors to differentiated cells. On average, during each division of a multipotent somatic stem cell, at least one of the daughter cells becomes a stem cell like the parent cell. Stem cells thus undergo self-renewal divisions such that the number of stem cells of a particular type can stay constant during the organism's lifetime. Other daughter cells, termed *transit amplifying cells*, divide rapidly and can undergo limited numbers of self-renewal divisions in which the properties of the daughter cells closely resemble those of the parent. These cells ultimately produce lineage-restricted progenitor cells, which cannot undergo self-renewal divisions but can divide and produce differentiated cells of a particular type.

Many types of stem cells in the adult body divide infrequently; they are kept in reserve in case certain types of differentiated cells are required. In contrast, their non-stem-cell daughters frequently undergo many rapid rounds of cell division. Such cells, often called *transit amplifying cells* (see Figure 22-11), can undergo limited numbers of self-renewal divisions, but eventually their many progeny form lineage-restricted progenitor cells. These cells, in turn, can divide and generate very specific types of terminally differentiated cells.

Stem cells can exhibit several patterns of cell division. Some types of stem cells always divide asymmetrically to generate one copy of the parent cell and one daughter cell with more restricted capabilities, such as dividing for a limited time or giving rise to fewer types of progeny than the parent stem cell (Figure 22-12a). This type of stem-cell division is commonly found in invertebrates such as *Drosophila*, discussed below.

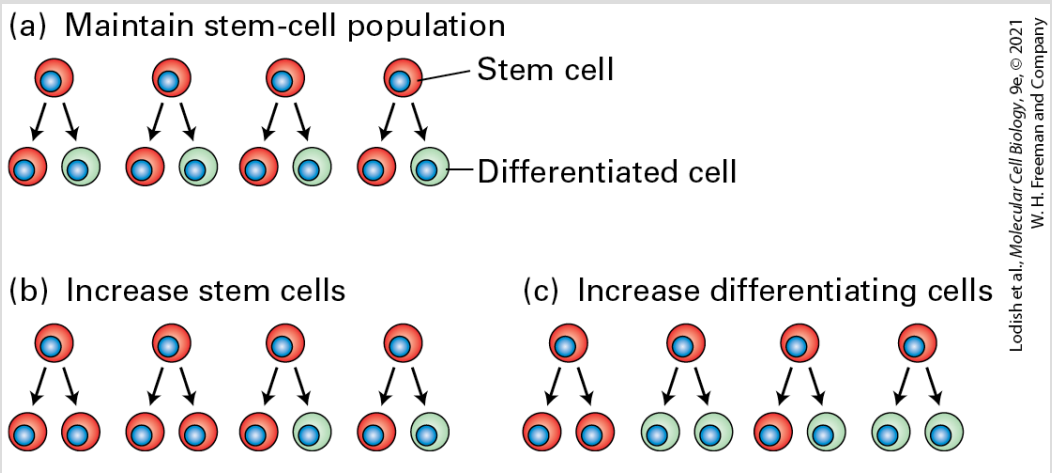


FIGURE 22-12 Patterns of stem-cell differentiation. Different patterns of stem-cell division produce different proportions of stem cells (red) and differentiating cells (green). (a) Stem cells can undergo asymmetric divisions, producing one stem cell and one differentiating cell. This pattern maintains the number of stem cells in the population. (b) Some stem cells can divide symmetrically to increase their numbers, which is often important in normal development or during recovery from injury. At the same time other stem cells in the same population can be dividing asymmetrically as in (a). (c) Some stem cells may divide as in (b) while at the same time other stem cells produce two differentiating progeny. See S. J. Morrison and J. Kimble, 2006, *Nature* **441**:1068–1074.

Other patterns of stem-cell division, commonly found in vertebrates, allow the number of stem cells or differentiated cells to increase or decrease according to the needs of the animal (Figure 22-12b, c). Hormones secreted by or on the surface of adjacent cells frequently regulate these patterns of stem-cell division. For example, a stem cell may divide symmetrically to yield two daughters that undergo different fates: depending on external signals sent by other cells, one may remain a stem cell and the other may become a transit amplifying cell that will undergo several

rounds of cell division and ultimately generate one or more types of differentiated progeny. As we will see in greater detail shortly, this happens in the small intestine: often one of the daughters remains a stem cell identical to its parent while the other daughter divides rapidly and generates multiple types of differentiated intestinal cells. Other stem-cell divisions are symmetric, producing two stem cells and increasing the number of stem cells of a particular type; this pattern of stem-cell division is common during development. Thus mitotic divisions of stem cells can either enlarge the population of stem cells or maintain a stem-cell population while steadily producing a stream of differentiating cells.

Stem Cells for Different Tissues Occupy Sustaining Niches

Stem cells need the right microenvironment to remain multipotent and to regulate the timing and pattern of their divisions. In addition to *intrinsic* regulatory signals — such as the presence of certain transcription factors and other regulatory proteins — stem cells rely on *extrinsic* hormonal and other signals from surrounding cells to maintain their status as stem cells. The location where a stem-cell fate can be maintained is called a **stem-cell niche**, by analogy to an ecological niche — a location that supports the existence and competitive advantage of a particular organism.

In order to investigate or use stem cells, we must find them and characterize them. It is often difficult to identify stem cells precisely; they are very rare among cells and generally lack distinctive shapes. Some stem cells divide rarely, if at all, until stimulated by signals that convey the need for new cells. For example, inadequate oxygen supplies can stimulate blood stem cells to divide, and injury to the skin or muscle can stimulate regenerative cell division starting with the activation of skin or muscle stem cells. Some stem cells, including those that form the continuously shed epithelium of the intestine, are continuously dividing, usually at a slow rate. In the rest of this section, we focus on several types of stem cells in plants and animals that are well characterized; in the coming years, undoubtedly other types of stem cells will also be understood in great detail.

Germ-Line Stem Cells in Many Organisms Produce Sperm or Oocytes

The **germ line** is the cell lineage that produces oocytes and sperm; it is distinct from the somatic cells that make all the other tissues but are not passed on to progeny. The germ line, like somatic-cell lineages, starts with stem cells, but these cells are unipotent in that they make only germ cells. Stem-cell niches have been especially well defined in studies of germ-line stem cells in *Drosophila* and *C.*

elegans. Germ-line stem cells are present in adult flies and worms, and the locations of these stem cells are well known.

In the fly, the niche where oocyte precursors form and begin to differentiate is located next to the tip of the *germarium*, the part of the ovary where eggs are formed (Figure 22-13a). There are two or three germ-line stem cells in this location next to a few cap cells, which create the niche by secreting two proteins in the TGF-β family, Dpp and Gbb, as well as Hedgehog (Hh) protein (Figure 22-13b). (These secreted protein signals were introduced in Chapter 16.) The cap cells create the niche because the TGF-β-class signals they send repress transcription of a key differentiation factor, the *Bag of marbles (Bam)* protein, in the neighboring germ-line stem cells. Repression of the *bam* gene allows germ-line stem cells to undergo self-renewing divisions, whereas activation of *bam* promotes differentiation. When a germ-line stem cell divides, one of the resulting daughters remains adjacent to the cap cells and is therefore maintained as a stem cell, like the parent cell. The other daughter is too far from the cap cells to receive the cap-cell-derived signals Dpp and Gbb. As a result, expression of the *bam* gene is induced, causing that daughter cell to enter the differentiation program. The signals involved were identified in part through the power of *Drosophila* genetics: mutant germ-line stem cells with defects in their Dpp or Gbb receptors, or their downstream signal transduction proteins, are lost prematurely. Conversely,

overexpression of Dpp by cap cells prevents differentiation of germ-line stem cells and causes formation of tumorlike germ cell masses.

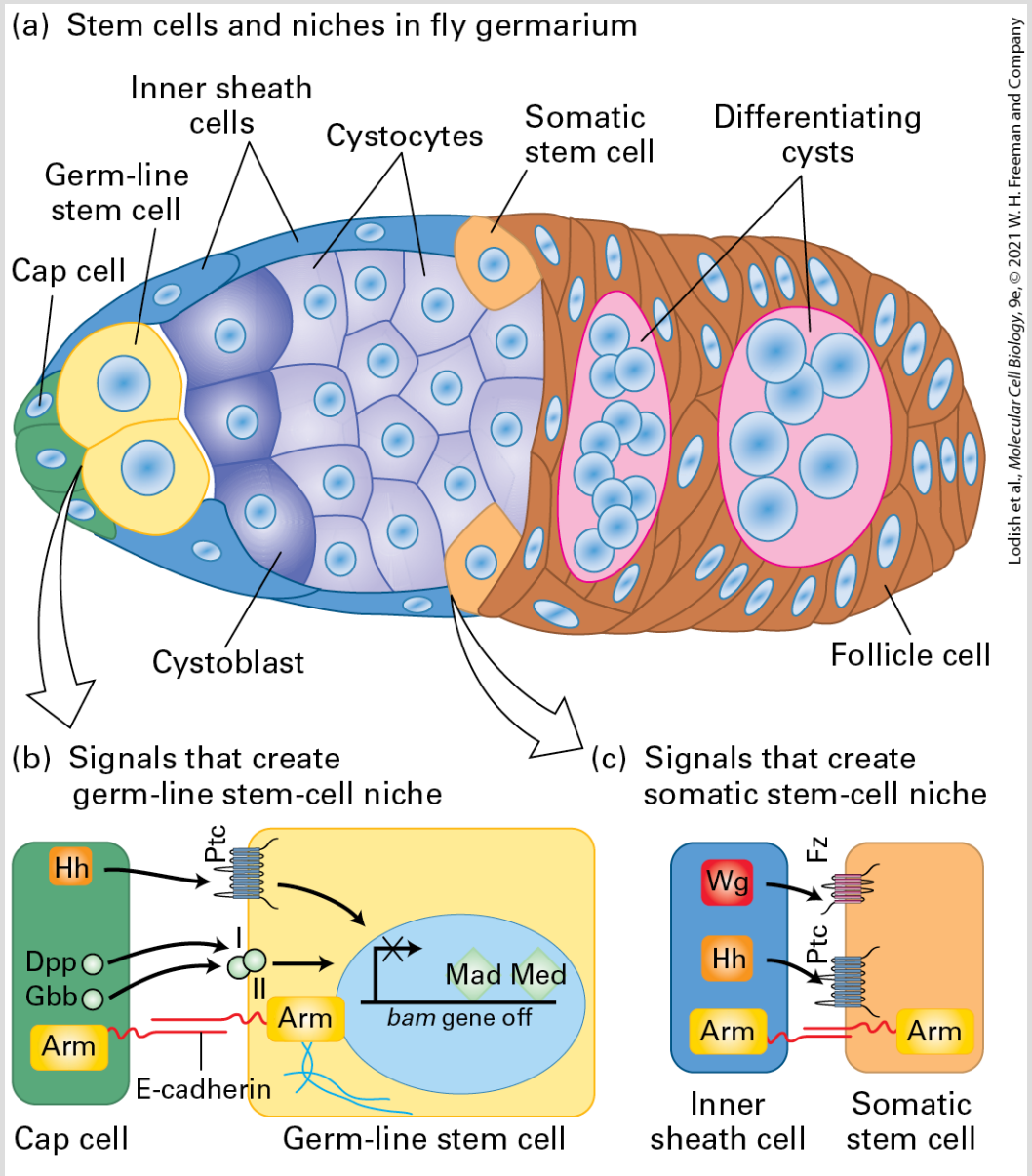


FIGURE 22-13 A *Drosophila* germarium. (a) Cross section of the germarium, showing female germ-line stem cells (yellow) and some somatic stem cells (orange) in their niches and the progeny cells derived from them. The germ-line stem cells produce cystoblasts (dark purple), which undergo four rounds of mitotic division to produce 16 interconnected cystocyte cells (light purple), one of which becomes the oocyte; the somatic stem cells produce follicle cells

(brown), which will make the eggshell. The cap cells (dark green) create and maintain the niche for germ-line stem cells, while the inner sheath cells (blue) produce the niche for somatic stem cells. (b) Signaling pathways that control the properties of germ-line stem cells. The signaling molecules — the TGF-β-family proteins Dpp and Gbb as well as Hedgehog (Hh) — are secreted by the cap cells. Binding of these ligands to receptors on the surface of a germ-line stem cell — the TGF-β receptors I and II and Ptc, respectively — results in repression of the *bam* gene by two transcription factors, Mad and Med. Repression of *bam* allows germ-line stem cells to undergo self-renewal divisions, whereas activation of *bam* promotes differentiation. The transmembrane cell-adhesion protein E-cadherin forms the homotypic adherens junctions (Chapter 20) between germ-line stem cells and cap cells. Arm (Armadillo), the fly β-catenin, connects the cytoplasmic tails of the E-cadherin to the actin cytoskeleton; both E-cadherin and Arm are important in maintaining the stem cells in their niche. (c) Signaling pathways that control the properties of somatic stem cells. The Wnt signal Wingless (Wg) is secreted by the inner sheath cells and binds to the Frizzled receptor (Fz) on a somatic stem cell. Hh is similarly produced and binds to the Ptc receptor. Both of these signals result in self-renewal of somatic stem cells. See R. Lehmann, 2012, *Cell Stem Cell* **10**:729, https://doi.org/10.1016/j.stem.2012.05.016; and E. W. Kahney, J. C. Snedeker, and X. Chen, 2019, *Curr. Opin. Cell Biol.* **60**:27.

The stem cells are held in the niche by the transmembrane cell-surface protein E-cadherin (see Chapter 20), which forms adherens junctions via homotypic interactions with similar E-cadherin molecules on the cap cell. These adherens junctions orient the mitotic spindle of the germ-line stem cells such that one daughter remains attached to the cap cell and the other is displaced from the niche; similar asymmetric stem-cell divisions occur during other developmental stages in *Drosophila*, as we discuss later (see Figure 22-31b). *Armadillo (Arm)*, the fly β-catenin, connects the cytoplasmic

tails of the E-cadherin molecules to the actin cytoskeleton; like E-cadherin, Arm is important in maintaining the stem-cell niche.

Separate somatic stem cells in the germarium produce follicle cells that will make the eggshell. The somatic stem cells have a niche too, created by the inner sheath cells, which produce Wingless (Wg) protein — a fly Wnt signal — and Hh protein (Figure 22-13c). Thus two different populations of stem cells can work in close coordination to produce different parts of an egg.

The identification and characterization of *Drosophila* germ-line stem cells, as well as similar cells from *C. elegans*, were important because they convincingly demonstrated the existence of stem-cell niches and led to experiments to identify the niche-made signals that cause cells to become and remain self-renewing stem cells. Thus a stem-cell niche is a set of cells and the signals they produce, not just a location.

Intestinal Stem Cells Continuously Generate All the Cells of the Intestinal Epithelium

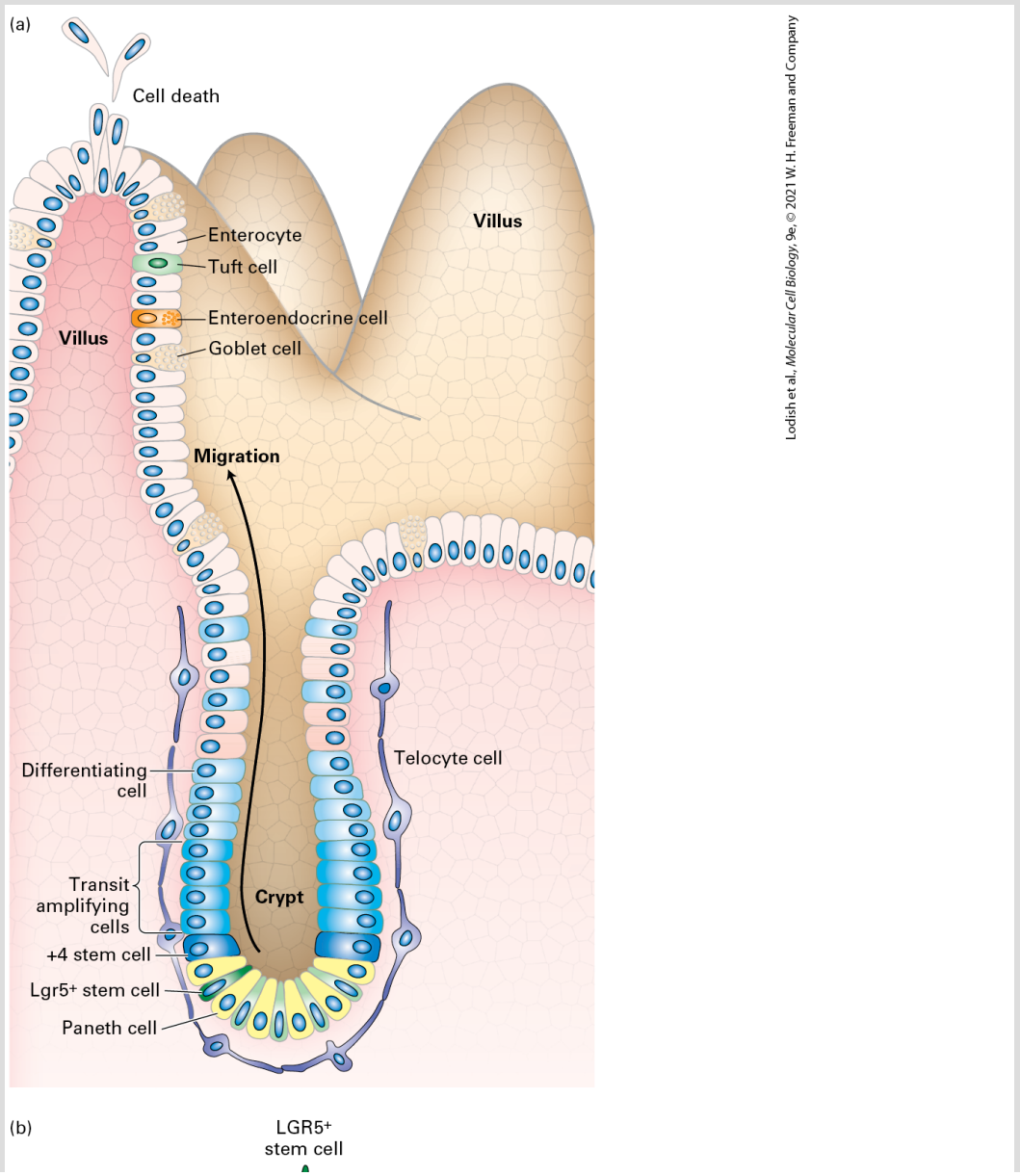
In humans, the epithelium lining the small intestine has a surface area of over 30 square meters and has two main functions: uptake of metabolites from digested food, and acting as a protective barrier

against toxins, bacteria, and other environmental stressors. It is a single cell thick (see Figure 20-11) and is composed of six types of differentiated cells: absorptive enterocytes, M cells, Tuft cells, Goblet cells, Paneth cells, and Enteroendocrine cells.

The most abundant epithelial cells, the absorptive enterocytes, transport nutrients essential for survival, such as glucose and amino acids, from the intestinal lumen into the body (see Figure 11-30). A network of capillaries underneath the epithelium transports these nutrients initially to the liver and ultimately to the rest of the body. The intestinal epithelium is the most rapidly self-renewing tissue in adult mammals, turning over every 5 days; in humans, up to 300 million intestinal epithelial cells, weighing a total of about 1 gram, are lost every day and need to be replenished.

The cells of the intestinal epithelium are continuously regenerated from a stem-cell population located deep in the intestinal wall in pits called *crypts* (Figure 22-14). Pulse-chase experiments using radiolabeled thymidine have shown that intestinal stem cells produce transit amplifying precursor cells that divide rapidly and then differentiate as they ascend the sides of crypts to form the surface layer of the fingerlike gut projections called *villi,* across which intestinal absorption occurs. The time from cell birth in the crypts to the loss of dead cells at the tips of the villi is only about 3 to 5 days (Figure 22-15). The production of new cells is precisely controlled: too little division would eliminate villi and lead to

breakdown of the intestinal surface; too much division would create an excessively large epithelium and might also be a step toward cancer. Paneth cells — a type of differentiated intestinal epithelial cells — are longer lived than the cells of the villi; they produce several antibacterial proteins, including the enzyme lysozyme, which degrades bacterial cell walls and thus protects the intestine from infections.



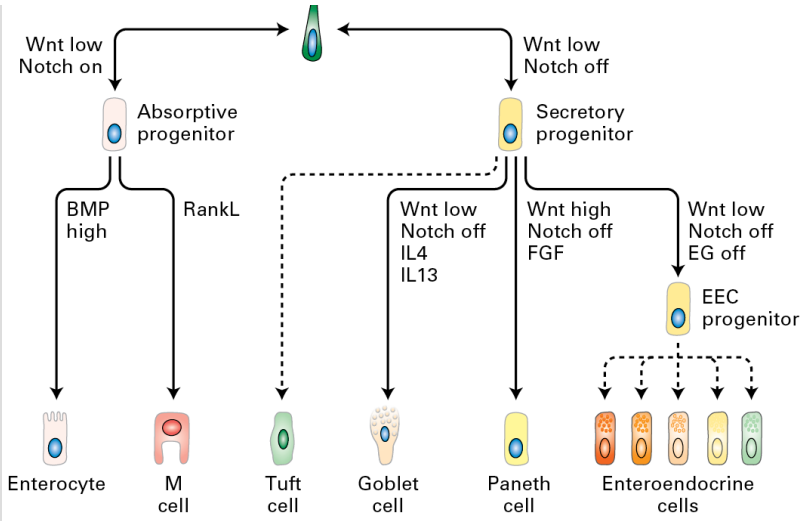
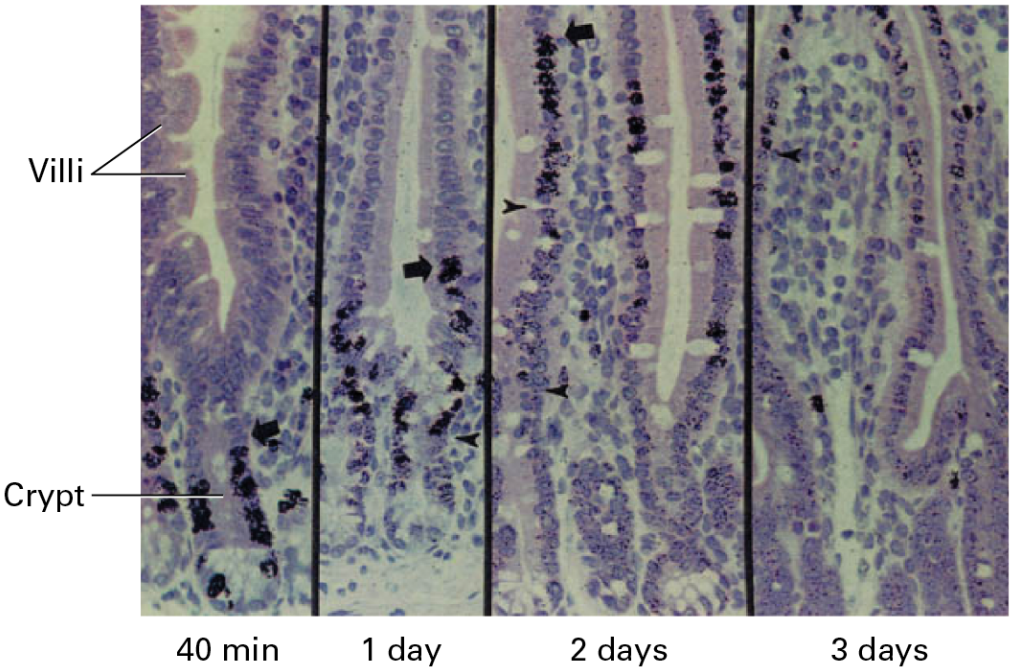


FIGURE 22-14 Intestinal stem cells and their niche. (a) Schematic drawing of an intestinal crypt and villus, showing the Lgr5-expressing ($\mathrm{Lgr5}^+$) intestinal stem cells (dark green), their mitotic progeny, the transit amplifying cells (intermediate blue), the terminal differentiating cells (light blue), and the several types of differentiated epithelial cells in the villus. The base of the crypt is the location of Paneth cells (yellow), which secrete antimicrobial defense proteins. The $+4$ cells (which occupy the fourth position from the crypt base, dark blue) can restore the $\mathrm{Lgr5}^+$ stem-cell compartment following injury. Telocytes are thin cells with long protrusions that form a three-dimensional network underlying the epithelial cells throughout the intestine, and are the major source of the Wnt and R-spondin hormones necessary for stem-cell self-renewal. (b) Lineages of cells in the small intestine and the extracellular signals that support specification of individual lineages. Epithelial turnover occurs every 3–5 days; new Paneth cells are produced from the transit amplifying cells every 3–6 weeks. $\mathrm{Lgr5}^+$ stem cells begin to differentiate as they are pushed out of the niche at the bottom of the crypt and the Wnt signal from telocytes is reduced. Depending on the extent of Notch signaling (Chapter 16) by adjacent cells they differentiate toward the absorptive (Notch on) or secretory fate (Notch off). Other hormones determine the subsequent specific fates of the absorptive or secretory transit amplifying progenitor cells into specific types of differentiated cells. The exact differentiation signals for Tuft cells are still not entirely understood; how different enteroendocrine subtypes are specified also is not yet well understood. See M. Shoshkes-Carmel et al., 2018, *Nature* **557**:242–246; and B. Degirmenci et al., 2018, *Nature* **558**:449–453.

[Part (a) information from L. C. Samuelson, 2018, *Nature* **558**:380. Part (b) information from H. Gehart and H. Clevers, 2019, *Nat. Rev. Gastroenterol. Hepatol.* **16**:19–34.]



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 22-15 Regeneration of the intestinal epithelium from stem cells can be demonstrated in pulse-chase experiments. Results from a pulse-chase experiment in which radioactively labeled thymidine (the pulse) was added to a culture of intestinal epithelial tissue. Proliferating cells incorporated the labeled thymidine into their newly synthesized DNA. After a brief period, the labeled thymidine was washed away and replaced with unlabeled thymidine (the chase); cells that underwent DNA replication after the chase did not become labeled. These micrographs show that 40 minutes after labeling, all of the label is in cells near the base of the crypt. At later times, the labeled cells are seen progressively farther away from their point of birth in the crypt. Cells at the top are shed. This process ensures constant replenishment of the gut epithelium with new cells.

[Republished with permission from John Wiley & Sons, Inc., from P. Kaur and C. S. Potten, 1986, “Cell Migration Velocities in the Crypts of the Small Intestine

After Cytotoxic Insult Are Not Dependent on Mitotic Activity," *Cell Tissue Kinet.* **6**:601–610; permission conveyed through Copyright Clearance Center, Inc.]

Experiments such as the one depicted in Figure 22-15 suggested that the intestinal stem cells were located somewhere near the bottom of the crypts, near Paneth cells, but these putative stem cells had no particular morphological characteristics that revealed their remarkable abilities. Which cells were the actual intestinal stem cells and which were the supportive cells that form the niche?

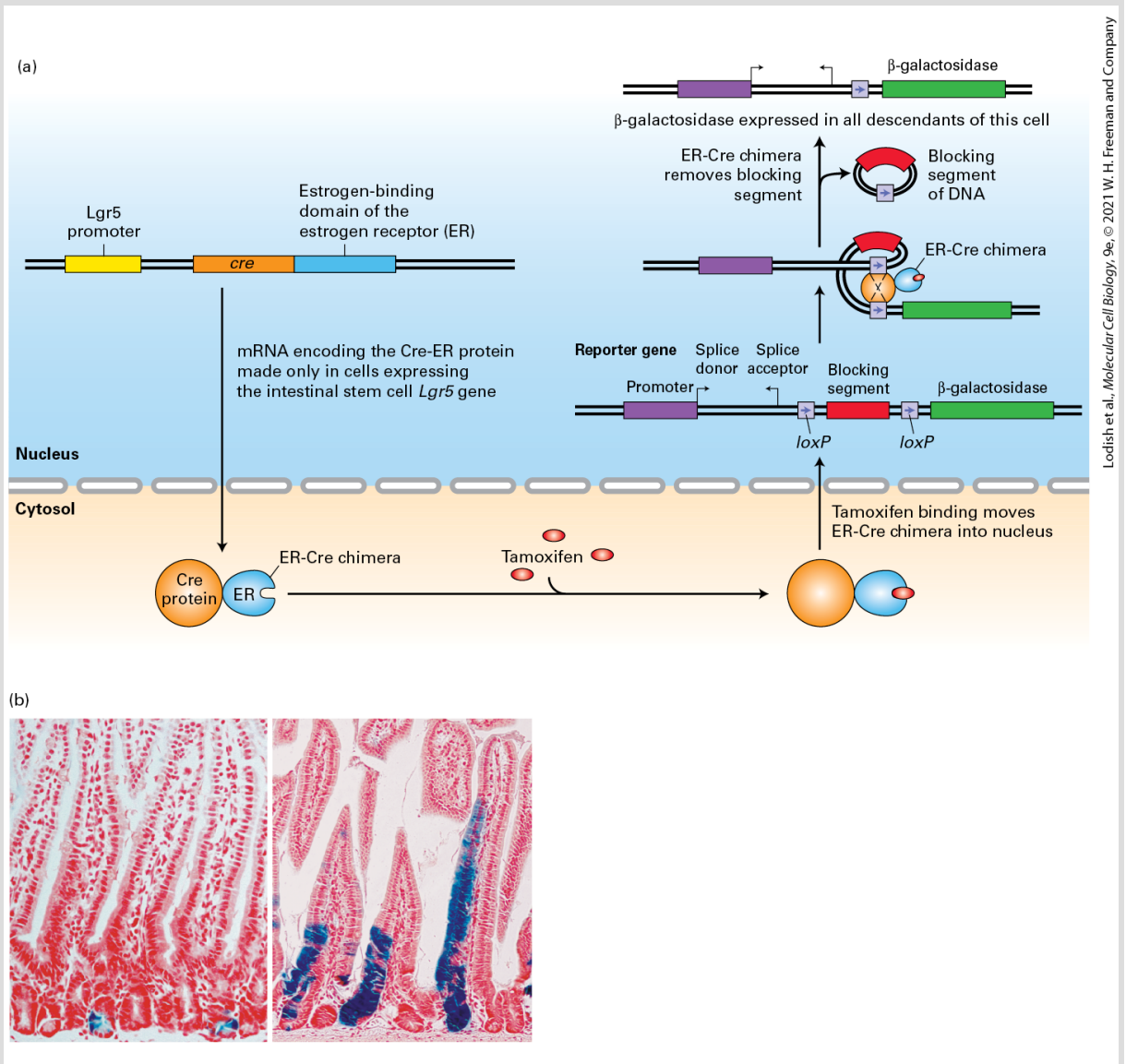
Wnt and R-Spondins Are Essential for Function of the $\mathbf{Lgr5^+}$ Intestinal Stem Cells

Prior genetic experiments had shown that Wnt signals are essential for intestinal stem cell maintenance. As evidence for the importance of these signals, overproduction of active β-catenin (normally activated by the Wnt signaling pathway; see Figure 16-26) in intestinal cells leads to excess proliferation of the intestinal epithelium. Conversely, blocking the function of β-catenin by mutating or inhibiting the Wnt-activated TCF transcription factor abolishes the stem cells in the intestine, leading to intestinal degeneration and eventual death. Thus Wnt signaling plays a critical role in the intestinal stem-cell niche, as it does in the skin, blood, and other organs. Indeed, mutations that inappropriately activate

the Wnt signaling pathway are a major contributor to the progression of colon cancer, as we will see in Chapter 25.

By analyzing a panel of genes whose expression in the intestine was induced by Wnt signaling, investigators zeroed in on *Lgr5*, a gene encoding a G protein–coupled receptor, because it was expressed only in a small set of cells at the very base of the crypts. Lgr5 binds a class of secreted hormones termed *R-spondins* and activates intracellular signaling pathways that potentiate Wnt signaling. Lineage-tracing studies showed that the descendants of these Lgr5-expressing cells indeed gave rise to all of the differentiated intestinal epithelial cells (Figure 22-16). These studies made use of genetically altered mice in which a version of the Cre recombination protein (see Figure 6-40), an estrogen receptor (ER)–Cre recombinase chimera, was placed under the control of the *Lgr5* promoter; the ER-Cre recombinase chimera was produced only in the few putative Lgr5-expressing stem cells at the bottom of the crypts. The version of Cre recombinase used in the study had been altered so that it resides inactive in the cytosol and is transferred into the nucleus only after addition of an estrogen analog (Figure 22-16a). There the Cre excises a blocking segment of DNA, activating expression of a β-galactosidase reporter gene. Importantly, all of the descendants of these cells will also express β-galactosidase. Immediately after addition of the estrogen analog, the only cells expressing β-galactosidase are the stem cells in the crypts. After a few days, all of the descendant epithelial cells also expressed β-galactosidase

(Figure 22-16b), showing that Lgr5 expression is indeed a marker of the intestinal stem cells.



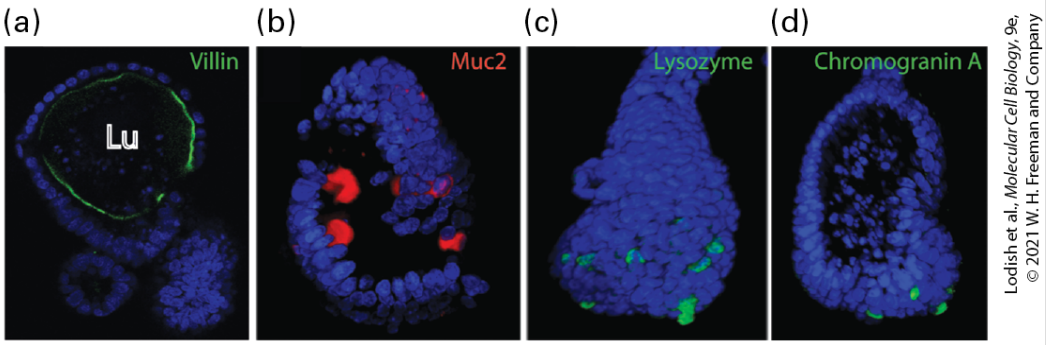
EXPERIMENTAL FIGURE 22-16 Lineage-tracing studies show that the Lgr5-expressing cells at the bases of crypts are the intestinal stem cells. (a) Using genetically altered ES cells, investigators generated one strain of mice in which a version of the gene encoding Cre recombinase (see Figure 6-40) was placed under the control of the *Lgr5* promoter, and thus Cre recombinase was produced only in cells, such as intestinal stem cells, that express the *Lgr5* gene. This version of Cre recombinase contained an additional domain from the estrogen receptor (ER) that binds the estrogen analog tamoxifen; like the ER and other

nuclear receptors (see Figure 8-44), the ER-Cre chimera is retained in the cytosol unless the estrogen analog tamoxifen is added. In the presence of tamoxifen, ER-Cre moves into the nucleus, where it can interact with *loxP* sites in the chromosomal DNA. A second reporter strain of mice contained a bacterial β-galactosidase reporter gene that was preceded by two *loxP* sites. The blocking segment of DNA in between these *loxP* sites prevented expression of the β-galactosidase gene, and the β-galactosidase gene could be expressed only in cells where an active Cre recombinase had removed the sequence in between the two *loxP* sites. The two strains of mice were mated, and offspring containing both marker transgenes were identified. In these mice, β-galactosidase was expressed only in cells in which the Lgr5-controlled ER-Cre gene was expressed and only after tamoxifen was given to the mice. Thus only Lgr5-expressing cells — and all of their descendants — would express the β-galactosidase gene. (b) One day after tamoxifen was given to these mice, the only cells expressing β-galactosidase (indicated by the blue histochemical stain) were the few Lgr5-expressing intestinal stem cells at the bases of the crypts (*left*). Five days after tamoxifen administration, additional blue cells — the epithelial descendants of the intestinal stem cells — were seen migrating up the sides of the villi. Some blue stem cells remained at the bottom of the crypt.

[Part (b) reprinted with permission from Nature Publishing Group, from N. Barker et al., 2007, "Identification of Stem Cells in Small Intestine and Colon by Marker Gene," *Lgr5 Nature* **449**:1003–1007; permission conveyed through Copyright Clearance Center, Inc.]

In subsequent studies, single Lgr5-expressing stem cells were isolated from intestinal crypts and cultured on an extracellular matrix (see Figure 20-24) containing type IV collagen and laminin, similar to the matrix that normally underlies and supports the intestinal epithelia. These cells generated small villus-like structures that contained the four principal differentiated cell types found in the mature intestinal epithelium (Figure 22-17). Such miniature three-dimensional versions of an organ, produced in culture from some type of stem cell, are termed *organoids.* Organoids are used by

researchers to study the formation of many organs, including the human brain, and to develop treatments for many diseases.



EXPERIMENTAL FIGURE 22-17 Single Lgr5-expressing intestinal stem cells build crypt-villus organoid structures in culture without niche cells. Single Lgr5-expressing cells isolated from intestinal crypts were placed in culture on a type IV extracellular matrix (see Figure 20-24). After 2 weeks, these cultures had formed epithelial sheets that resembled villi in structure. Staining of these organoids for specific marker proteins showed that they contained all four differentiated epithelial cell types: (a) villin (green) is a marker protein for the absorptive enterocytes that are localized near the apical (luminal, Lu) surface of these organoids; (b) Muc2 (red) for goblet cells; (c) lysozyme (green) for Paneth cells; and (d) chromogranin A (green) for enteroendocrine cells. The organoids were also stained with DAPI (blue) to reveal nuclei.

[Reprinted with permission from Nature Publishing Group, from T. Sato et al., 2009, "Single Lgr5 Stem Cells Build Crypt-Villus Structures In Vitro Without a Mesenchymal Niche," *Nature* **459**(7244):262–265; permission conveyed through Copyright Clearance Center, Inc.]

Taken together, these experiments established that expression of the *Lgr5* gene defines the intestinal stem cells and showed that these cells are localized at the bases of the intestinal crypts interspersed between the terminally differentiated Paneth cells (see Figure 22-

14a). Lgr5-expressing cells are also found in the stomach, colon, and pancreas — which, like the small intestine, are formed from the embryonic endoderm — and are thought to be the stem cells for these tissues. Indeed, culturing Lgr5-expressing cells from these tissues in the presence of Wnt, R-spondin, and other hormones generates organoids that contain differentiated cells characteristic of these tissues.

Telocytes — thin fibroblast-like cells with long protrusions that surround the sheet of epithelial cells throughout the intestine (see Figure 22-14a) — are the major source of both Wnt and R-spondins, hormones essential for normal function of the Lgr5-expressing stem cells. Paneth cells also secrete Wnt and were once thought to be the major Wnt-producing intestinal cells, but knocking out Wnt expression in all intestinal epithelial cells, including Paneth cells, had no effect on intestinal stem cells or renewal of intestinal epithelial cells. In contrast, conditional ablation in mouse telocytes of the porcupine gene, which is required for functional maturation of all Wnt proteins, caused a rapid cessation of Wnt signaling to intestinal crypts, followed by loss of proliferation of stem and transit amplifying cells and impaired epithelial renewal. Thus telocytes surrounding the base of the crypts are the principal source of the Wnt and R-spondin self-renewal signals to intestinal stem cells. These cells also produce EGF and TGFα, two hormones in the epidermal growth factor family that are also required for $Lgr5^{+}$ stem cell function.

Lgr5-expressing cells are not the only type of intestinal stem cells. Evidence indicates that so-called +4 cells located in the crypts (see Figure 22-14a) are reserve stem cells that can generate Lgr5-expressing stem cells following intestinal injury, such as by irradiation. Recall that transit amplifying cells have limited self-renewal potential (see Figure 22-11). During periods of intestinal injury, when many Lgr5-expressing stem cells are lost, some of the transit amplifying cells, under the influence of Wnt signals, can *dedifferentiate* — reverting to Lgr5-expressing stem cells and relocalizing to the stem cell niche at the base of the crypts! Thus the conversion of differentiated cells into stem cells, as seen experimentally during formation of iPS cells, may occur normally in the body during periods of stress or injury. As such, rather than relying on a single stem cell-to-differentiated cell lineage (see Figure 22-14b), the intestine seems capable of drawing on a pool of reserve stem-cell populations. Future work will undoubtedly clarify the role of these cells in generating intestinal epithelial cells.

Hematopoietic Stem Cells Form All Blood Cells and All Cells of the Immune System

All red and white blood cells and virtually all cells that form the immune system need to be continuously replenished. Human red blood cells survive 90–120 days before they are degraded by

macrophages, mainly in the spleen. Macrophages can survive for several months while neutrophils survive only a few days. All of these cells derive from a single type of multipotent, self-renewing *hematopoietic stem cell (HSC)* through multiple transit amplifying cell divisions leading to multipotent and then unipotent progenitors (Figure 22-18).

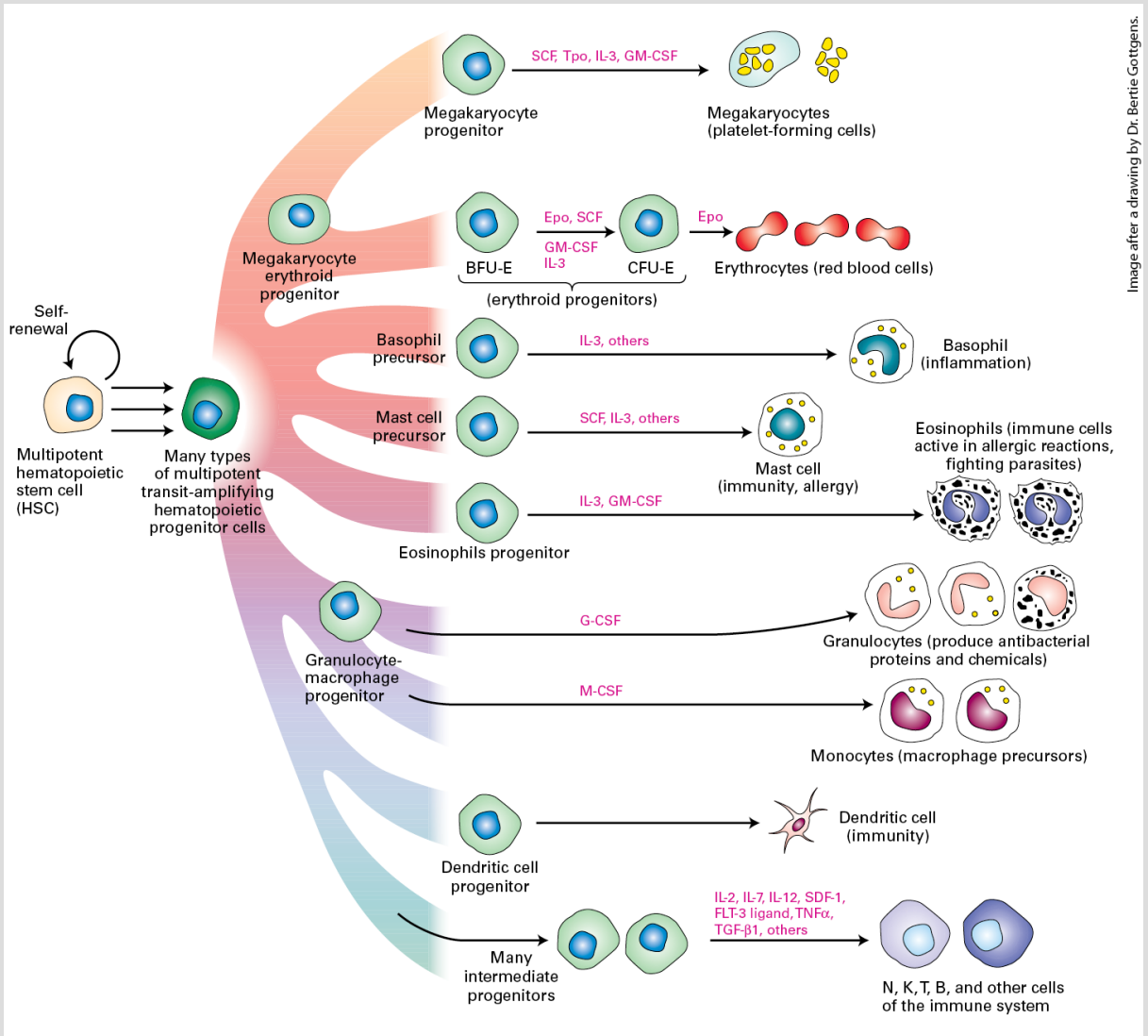


FIGURE 22-18 Formation of blood cells from hematopoietic stem cells in the bone marrow. Early in mammalian development, multipotent hematopoietic stem cells often divide symmetrically to increase the numbers of stem cells. In adults, they generally divide

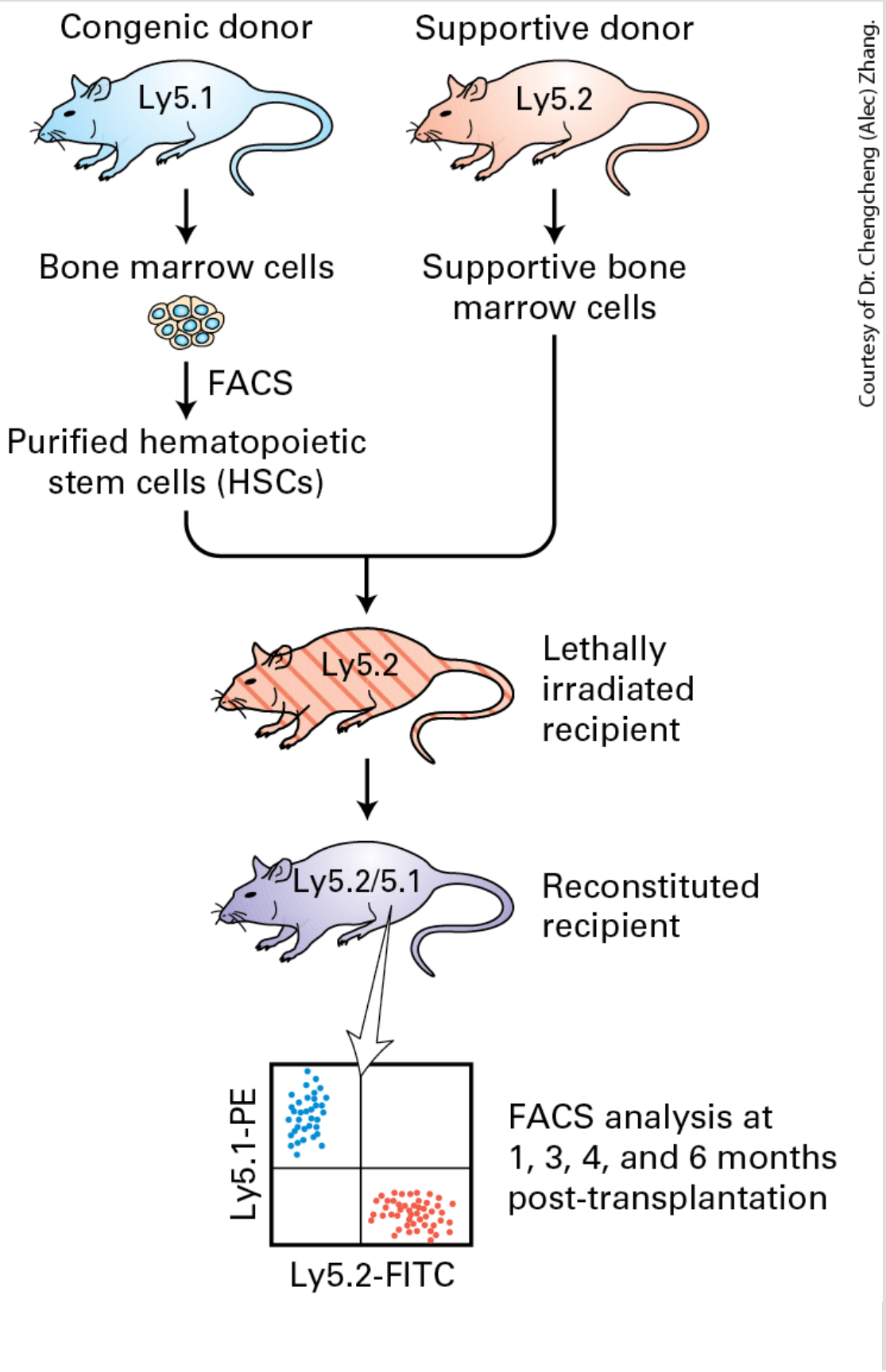
asymmetrically to form one daughter cell that is multipotent, like the parent stem cell, and another daughter cell with a more restricted fate. Recent single-cell sequencing of purified progenitor populations revealed an unanticipated heterogeneity; there are at least 18 distinct subtypes of transit-amplifying hematopoietic progenitors. These multipotent transit amplifying cells are likely capable of limited numbers of self-renewal divisions but, depending on the types and amounts of cytokines present, they undergo rapid rounds of cell division and generate different types of progenitor cells (light green). These progenitors are either multipotent (e.g., bipotential granulocyte — monocyte and erythroid — megakaryocyte progenitors) or unipotent in that they can give rise to more than one type or only a single type of differentiated blood cells, respectively; they respond to one or a few specific cytokines. The lymphoid branch has not yet been studied in such great detail with single cell transcriptomics; future work should reveal more about how the B, NK, and T cells of the immune system (Chapter 24) are specified. Some of the cytokines that support this process are indicated (pink labels). CSF = colony-stimulating factor; IL = interleukin; SCF = stem-cell factor; Epo = erythropoietin; Tpo = thrombopoietin. See S. Watcham, I. Kucinski, and B. Gottgens, 2019, *Blood* **133**:1415.

Characterizing Hematopoietic Stem Cells by Transplantation

Characterization of HSCs and the more restricted progenitors they generate was much more difficult than defining the germ or intestinal stem cells discussed earlier, as these stem cells are localized to niches that have a specific location in tissues (see Figures 22-13 and 22-14). Like all stem cells, hematopoietic stem cells are found in niches, but in embryos HSCs are dispersed among fetal liver cells and in adults HSCs and their niche cells are scattered among many other types of cells within the bone marrow. Identifying HSCs and the cells that make up the HSC niche was also

difficult because HSCs comprise only about 0.01 percent of bone marrow or fetal liver cells.

Hematopoietic stem cells and the progenitors they generate were originally detected and quantified by bone marrow transplantation experiments in mice whose hematopoietic stem and progenitor cells had been eradicated by irradiation (Figure 22-19). Purification of HSCs was difficult since no cell-surface protein is expressed only by HSCs, and thus no exclusive surface marker was available for these cells. However, much work showed that HSCs could be prospectively identified and partially purified because they do express cell-surface proteins called *CD150* and *Sca-1* (of unknown functions) and do not express any of the dozen or so *Lin (Lineage-restricted)* marker proteins found on the surface of specific types of differentiated hematopoietic cells. Further efforts took advantage of the fact that HSCs require a growth factor termed *stem-cell factor (SCF)* for their survival; this protein, which is produced as the extracellular domain of a membrane-spanning protein and is localized to the surface of an adjacent signaling cell, activates the c-Kit protein tyrosine kinase receptor on HSCs. HSCs also require CXCL12, a chemoattractant for HSCs that binds to a G protein–coupled receptor on HSCs and is required to keep HSCs in the niche. CXCR4, the receptor for CXCL12, is also found on HSCs.


Congenic donor
Ly5.1
Supportive donor
Ly5.2
Bone marrow cells
Supportive bone marrow cells
FACS
Purified hematopoietic stem cells (HSCs)
Ly5.2
Lethally irradiated recipient
Ly5.2/5.1
Reconstituted recipient
Ly5.1-PE
Ly5.2-FITC
FACS analysis at 1, 3, 4, and 6 months post-transplantation
Courtesy of Dr. Chengcheng (Alec) Zhang.

EXPERIMENTAL FIGURE 22-19 Functional analysis of mouse hematopoietic stem cells by bone marrow transplantation. The two strains of mice used in this analysis are genetically identical except for the gene encoding a protein, termed *Ly5*, found on the surfaces of all nucleated blood cells, including all T and B lymphocytes, granulocytes, and monocytes. The proteins encoded by the two alleles of the gene, Ly5.1 and Ly5.2, can be detected by specific monoclonal antibodies. A recipient Ly5.2 mouse is lethally irradiated to kill all HSCs, clearing out the stem-cell niches in the bone marrow so they can receive stem cells from a donor mouse; the mice are then injected with stem cells purified from a Ly5.1 strain. Because the stem cells take weeks or months to produce differentiated blood cells, the recipient mouse will die unless it receives bone marrow progenitor cells from a genetically identical Ly5.2 mouse (termed "supportive" cells) that will produce mature blood and immune system cells for the first few weeks after the transplant. At intervals after the transplant, blood or bone marrow is recovered and reacted with a blue-fluorescing monoclonal antibody to Ly5.1 and a red-fluorescing monoclonal antibody to Ly5.2. Mature blood cells that are descended from the donor stem cell are detected by FACS analysis, seen here as cells that fluoresce blue and not red. These cells can be sorted and stained with fluorescent antibodies specific for marker proteins found on different types of mature blood cells to show that a stem cell is indeed pluripotent, in that it can generate all types of lymphoid and myeloid cells.

Thus HSCs could be purified first by treating bone marrow cells with magnetic beads coated with a series of monoclonal antibodies each specific for a Lin protein on the surface of differentiated hematopoietic cells, and then removing all of the lineage-positive differentiated cells attached to the beads with a magnet. Next the remaining lineage-negative cells were treated with fluorescent antibodies for c-Kit, CD150, and Sca-1, and the few cells that expressed all of these marker proteins were separated in a fluorescence-activated cell sorter (FACS; see Figures 4-1 and 4-2) and used in transplantation experiments.

Remarkably, such transplantation studies revealed that a single HSC is sufficient to restore the entire blood system when transferred into a lethally irradiated mouse in which all of its HSCs had been killed. After transplantation, the HSC takes up residence in a niche in the bone marrow and divides to make more HSCs as well as progenitors of the different blood-cell lineages.

The first successful human bone marrow transplant was done in 1959, when a patient with end-stage (fatal) leukemia was irradiated to destroy her cancer cells as well as her own normal HSCs. She was transfused with bone marrow cells from her identical twin sister, thus avoiding an immune response, and was in remission for 3 months. This pioneering effort, which was awarded a Nobel Prize in Medicine in 1990, led to the present-day treatments that can often lead to a complete cure of many diseases. Because stem cells in the transplanted marrow can generate all types of functional blood cells as well as all cells of the immune system, transplants using bone marrow HSCs from normal donors are useful in patients with several hereditary anemias, including β-thalassemia (caused by mutations that lead to production of reduced amounts of the hemoglobin β chain) and sickle-cell disease (caused by a point mutation in the β chain of hemoglobin), as well as patients with diseases such as X-linked Severe Combined Immune Deficiency (SCID), in which the gene for the IL-2 receptor common gamma chain (see Figure 16-19b) is inactivated and no immune cells are produced. Transplants are also used to treat many cancer patients

who have received irradiation or chemotherapy, both of which destroy bone marrow cells as well as cancer cells.

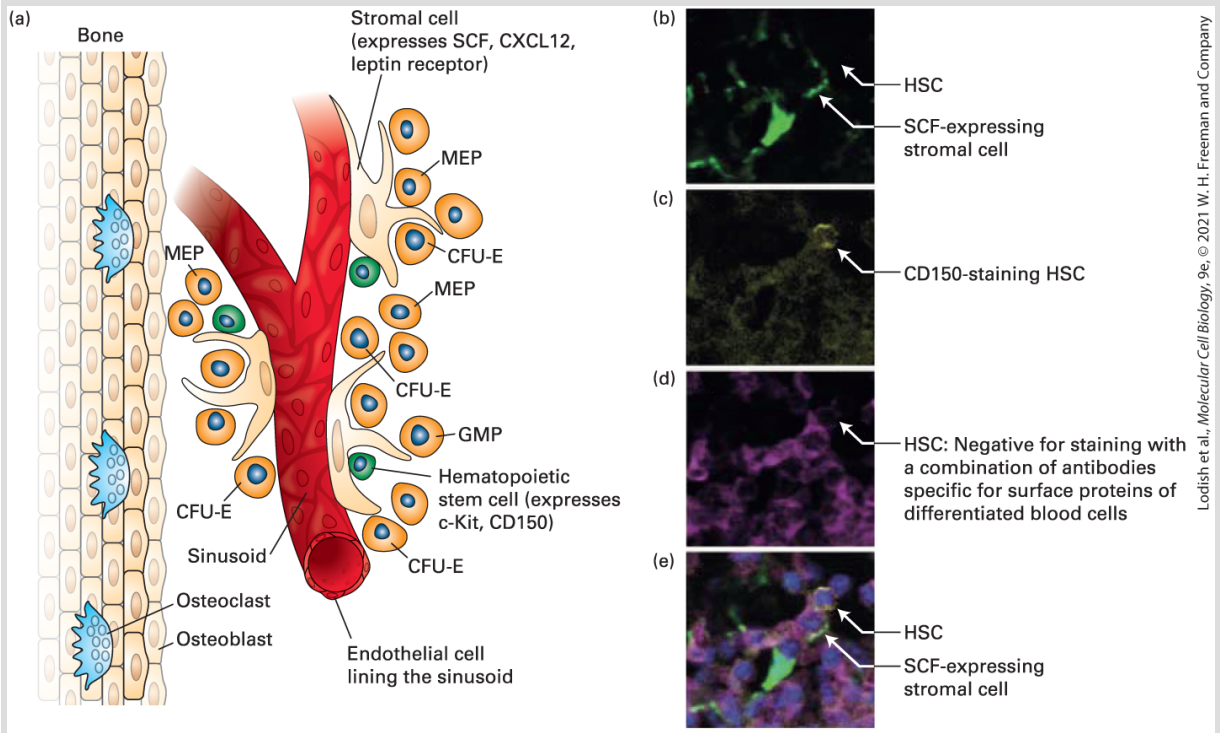
Patients with β-thalassemia, SCID, and sickle-cell disease can now be treated by a combination of gene therapy and an autologous bone marrow transplant. Cell populations containing HSCs are removed from the patient, treated in culture with a lentiviral vector (a type of retrovirus derived from HIV; see Figure 6-35) into which a normal copy of the mutant gene had been inserted, and the corrected HSCs transfused back into the patient. ■

Niches for Hematopoietic Stem Cells and Many Hematopoietic Progenitor Cells

During embryonic life, mammalian HSCs reside in the liver and often divide symmetrically, producing two daughter stem cells (see Figure 22-12b); this process allows the number of stem cells to increase over time and produce the large number of progenitor cells required to make all of the necessary blood cells before birth. Hepatocyte progenitors are the major cell that forms the HSC niche in the fetal liver. These cells express SCF and other proteins required for HSC survival. Co-culture of these hepatic progenitor cells with HSCs led to expansion of HSC numbers as well as formation of their differentiated progeny, solidifying their role as HSC niche cells.

In adult mammals, HSCs reside in the bone marrow and are largely quiescent, resting in the G_0 state in the bone marrow stem-cell niche. When more blood cells are needed, cytokines are generated locally, which signals some HSCs to divide, producing stem cells like the parent cells and rapidly proliferating transit amplifying cells that generate the progenitors depicted in Figure 22-18. Whether individual HSCs undergo symmetric or asymmetric divisions is not known.

A small number of mesenchymal cells surround the small blood vessels, termed *sinusoids*, that permeate the bone marrow. Some of these cells are termed *leptin receptor positive* as they express the receptor for the cytokine leptin on their surface. They also express large amounts of SCF, synthesize abundant CXCL12, and are thought to be the major HSC niche cells in the bone marrow (Figure 22-20a). Immunofluorescence analysis showed that about 85 percent of HSCs physically contact these cells (Figure 22-20b). Other cells in the bone marrow probably influence stem-cell maintenance or niche function by releasing other types of hormones.



EXPERIMENTAL FIGURE 22-20 The hematopoietic stem-cell niche in the bone marrow. (a) The bone marrow contains dozens of different types of cells, including osteoblasts and osteoclasts that build and degrade bone, respectively, as well as multiple types of hematopoietic stem and progenitor cells, and other cell types. The bone marrow is permeated by small blood vessels termed *sinusoids*. The predominant cells that form the HSC niche are the very rare mesenchymal stromal cells that adhere to these vessels and that express a combination of cell-surface proteins including the leptin receptor $(\mathrm{LepR} + \mathrm{cells})$ and SCF, the hormone that binds to and activates the c-Kit protein tyrosine kinase receptor on HSCs and many types of hematopoietic progenitor cells. The stromal cells also secrete CXCL12, a chemoattractant for HSCs, and the $\mathrm{LepR} + \mathrm{cells}$ also act as supportive niche cells for a number of hematopoietic progenitor cells depicted in Figure 22-18. MEP = megakaryocyte erythroid progenitor; GMP = granulocyte macrophage progenitor; CFU-E = colony forming unit erythroid–erythrocyte progenitor. See S. J. Morrison and D. T. Scadden, 2014, *Nature* **505**:327; and S. Comazzetto et al., 2019, *Cell Stem Cell* **24**:477. (b–d) Immunofluorescence detection of HSCs and niche cells in bone marrow, showing that HSCs are localized next to SCF-expressing cells. Antibodies to SCF were not available, so in order to detect SCF expression, a mouse was generated in which GFP cDNA was placed in the SCF gene locus and expressed only in cells that normally produce SCF. (b) Bone marrow sections were then examined in a fluorescence microscope to detect the SCF-expressing cells. (c) To

detect HSCs, the sections were stained with an antibody to the CD150 protein, expressed on the surface of HSCs. (d) The sections were also stained with a collection of antibodies specific for proteins expressed on the surface of specific types of differentiated blood cells, but not by HSCs. (e) When the three images are merged, an HSC (white arrows) can be seen adjacent to the SCF-expressing stromal cell.

[Parts (b–e) reprinted with permission from Nature Publishing Group, from L. Ding et al., 2012, "Endothelial and Perivascular Cells Maintain Haematopoietic Stem Cells," *Nature* **481**(7382):457–462; permission conveyed through Copyright Clearance Center, Inc.]

Like HSCs, a number of hematopoietic precursors also require SCF for their proliferation and differentiation and express c-Kit, the SCF receptor, on their surface. These cells — including many of the multipotent progenitors formed from HSCs, as well as megakaryocyte-erythrocyte progenitors, and granulocyte-macrophage progenitors — are also found in the bone marrow adjacent to the leptin receptor positive cells that express SCF. Indeed, conditional deletion of SCF from these leptin receptor positive cells depleted all of these progenitors and their descendants from the bone marrow. Thus many hematopoietic progenitors as well as HSCs require specific types of bone marrow niche cells for their survival and function.

Regulating the Production of Differentiated Hematopoietic Cells

In adults, HSCs rarely divide, but when they do, on average one of the daughters remains a stem cell like the parent (self-renewal)

while the other divides rapidly and eventually gives rise to one of several multipotent progenitors. These are more restricted in their fates and like transit amplifying cells in other lineages (see Figure 22-11) are capable of a limited number of self-renewal divisions before giving rise to lineage-specific progenitor cells, each of which in turn can divide and generate multiple numbers of a specific type of differentiated cell (see Figure 22-18).

Numerous extracellular growth factors termed **cytokines** regulate HSC self-renewal divisions as well as proliferation and differentiation of the precursor cells for various blood-cell lineages (see Chapter 16). Each branch of the blood-cell lineage tree is affected by different cytokines, allowing exquisite control of the production of specific cell types. If all blood-cell types are needed — for example, after a bleeding injury — multiple cytokines can be produced. If only one cell type is needed, specific signals control its production. For example, when a person is traveling at high altitude, *erythropoietin* is made by the kidneys and stimulates the proliferation and differentiation of CFU-E (erythroid progenitor) cells, but not other types of blood-cell precursors. The erythropoietin receptor is expressed only in erythroid precursors and activates intracellular signal transduction pathways that lead to induction of ~400 genes essential for formation of a mature red blood cell (see Figure 16-18). Similarly, G-CSF, a different cytokine, stimulates proliferation of bipotential granulocyte-macrophage progenitors and their differentiation into granulocytes, while M-CSF

stimulates production of macrophages from the same progenitor cell type (see Figure 22-18).

By transplanting specific types of hematopoietic precursors into mice and observing which blood cells were produced, researchers could infer which precursors or terminally differentiated cells (e.g., erythrocytes, monocytes) arise from a particular type of precursor. Colony-forming assays can also determine the potential fates of progenitor cells; individual progenitor cells are cultured in a semisolid matrix such as methyl cellulose, together with a group of cytokines. All of the daughter cells produced remain localized together, forming a colony of cells. From the types of differentiated cells in the colony one can infer the identity of its progenitor.

These experiments gave rise to the notion that there are discrete states of hematopoietic stem and progenitor cells, and as these cells underwent differentiation their transcriptomes would change from one defined discrete state to another. Single-cell sequencing, with its ability to analyze the transcriptomes of multiple individual cells, has modified this understanding of hematopoiesis. Because some HSCs express transcription factors required for lymphoid (immune system) differentiation and others essential for myeloid (red and white blood cell) differentiation, some HSCs are thought to be programmed to differentiate toward the lymphoid or myeloid pathways. Rather than cellular transitions between discrete progenitor states, such as between a granulocyte/monocyte progenitor and differentiated granulocytes or monocytes, multiple

studies have supported the idea of continuous differentiation pathways, with gradual changes in the transcriptomes from one differentiating cell state to another (see Figure 22-18).

🌿 Meristems Are Niches for Stem Cells in Plants

In plants, as in their multicellular animal counterparts, the production of all tissues and organs relies on small populations of stem cells. Like animal stem cells, these stem cells are defined by their ability to undergo self-renewal and to generate daughter cells that produce differentiated tissues. Also like animal stem cells, plant stem cells reside in stem-cell niches — specialized microenvironments where extracellular signals are produced that maintain the stem cells in a multipotent state. Because the last common ancestor of plants and animals was a unicellular eukaryote, it would appear that, despite common organizing principles, stem cells and their niches evolved independently and by different pathways in plants and animals — an example of *convergent evolution*.

The niches in which plant stem cells are located, called **meristems**, can persist for thousands of years in long-lived species such as bristlecone pines. The body axis of the plant is defined by two primary meristems that are established during embryogenesis, the

shoot apical meristem and the *root apical meristem*. In contrast to animal development, very few tissues or organs are specified during plant embryogenesis. Instead, organs such as leaves, flowers, and even germ cells are continuously generated as the plant grows and develops. The aboveground part of the plant is derived from the shoot apical meristem and the belowground part from the root apical meristem. Classic clonal analysis experiments have demonstrated that plant-cell fate depends on the cell's position, not its lineage. A cell's identity is reinforced by intercellular signals such as hormones, mobile signaling peptides, and miRNAs. ■

Unlike somatic stem cells in metazoan animals, somatic plant stem cells give rise to entire organs, not just specific tissues or lineages. Slowly dividing pluripotent stem cells are located at the apex of the shoot apical meristem, with more rapidly dividing multipotent transit amplifying daughter cells on the periphery. Descendants of the shoot stem cells are displaced to the periphery of the meristem and are recruited to form primordia of new organs, including leaves and stems. Division ceases as these cells acquire the characteristics of specific cell types, and most organ growth occurs by cell expansion and elongation (Figure 22-21a). New shoot stem-cell niches can form in the axils of leaf primordia, which then grow to form lateral branches. Floral meristems give rise to the four floral organs — sepals, stamens, carpels, and petals — that form flowers. Unlike shoot apical meristems, floral meristems gradually become depleted as they give rise to the floral organs.

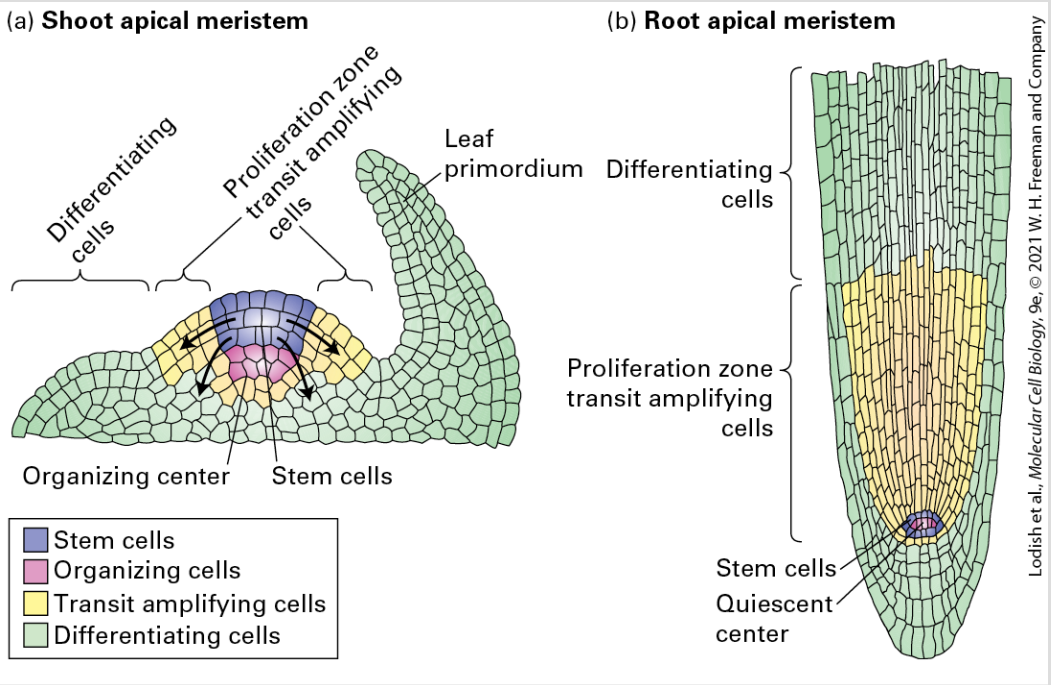


FIGURE 22-21 Structures of the *Arabidopsis thaliana* shoot and root meristems. (a) Transverse section through the apex of the shoot apical meristem. The organizing center cells signal to maintain the overlying stem cells. The stem cells produce daughters by division in the direction of the black arrows, generating rapidly dividing transit amplifying cells that will eventually differentiate and give rise to entire organs, such as a leaf. (b) Transverse section through the root meristem. Stem cells surround the mitotically less active quiescent center, four cells that send signals to prevent stem-cell differentiation. Each stem cell divides asymmetrically: one daughter remains adjacent to the quiescent center and becomes a stem cell (self-renewal); the other daughter becomes a transit amplifying cell that divides a number of times before exiting the cell cycle, elongating, and assuming a specific differentiated state.

[Information from R. Heidstra and S. Sabatini, 2014, *Nat. Rev. Mol. Cell Biol.* **15**:301.]

A Negative Feedback Loop Maintains the Size of the Shoot Apical Stem-Cell Population

Genes required for stem-cell identity, maintenance, and cell differentiation have been defined by genetic screens in the mustard-family weed *Arabidopsis thaliana* (see Figure 1-23h) for mutants exhibiting larger, smaller, or non-replenishing meristems, as well as by more recent gene-expression profiling studies of isolated meristem-cell populations. One shoot apical meristem determinant is the gene called *WUSCHEL (WUS)*, which encodes a homeodomain transcription factor (see Chapter 8). *WUS* is required for maintenance of the stem-cell population but is expressed in the supportive cells underlying the stem cells. These cells, collectively termed the *organizing center*, are analogous to the niche cells in metazoans (see Figure 22-21a). While *WUS* mRNA and protein are synthesized in the cells of the organizing center, WUS moves from the organizing-center cells into the stem cells, presumably through the interconnecting plasmodesmata (see Figure 20-43). In one study, a WUS-GFP fusion protein, when expressed in *WUS*-negative *Arabidopsis* plants, was able to rescue the mutant phenotype. Subsequent microscopic analysis showed that this WUS-GFP chimera accumulated in the stem cells, indicating it had moved there from the organizing-center cells.

Once in the stem cells, WUS binds to many sites in the genome; it represses a large number of genes that are expressed in differentiating cells, including a group of differentiation-promoting transcription factors required for leaf development. WUS also directly activates the expression of *CLAVATA3 (CLV3)* in stem cells. *CLV3* encodes a small, secreted peptide that binds to the CLV1 receptor on the surface of organizing-center cells and generates an intracellular signal that negatively regulates WUS expression. Overexpression of WUS causes a large expansion of the meristem stem-cell population at the expense of production of differentiated cells. Thus the negative feedback loop between a transcription factor, WUS, and a signaling peptide, CLV3, maintains the size of the stem-cell population and the number of their dividing daughter cells over the lifetime of the plant (Figure 22-22). ■

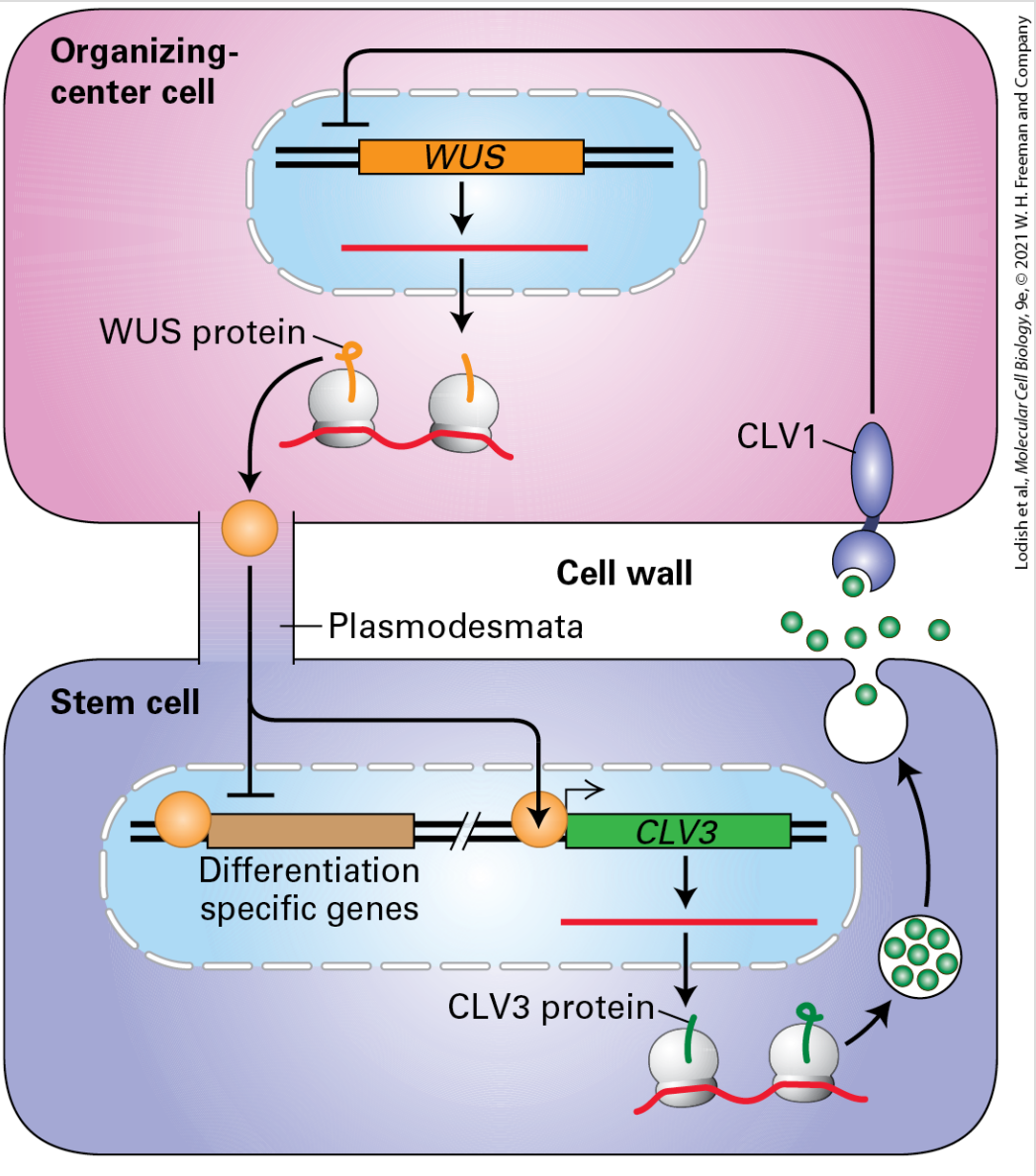


FIGURE 22-22 Regulatory network in the *Arabidopsis* shoot meristem stem-cell niche. The transcription factor WUS (orange circles) is synthesized in the organizing-center cells and moves via plasmodesmata into stem cells, where one of its functions is to induce expression of the CLV3 hormone (green circles). Secreted CLV3 protein binds to CLV1, the CLV3 receptor protein kinase on the surface of organizing-center cells; there it activates a signal that represses *WUS* transcription. See E. Aichinger et al., 2012, *Annu. Rev. Plant Biol*; **63**:615.

[Data from R. Heidstra and S. Sabatini, 2014, *Nat. Rev. Mol. Cell Biol.* **15**:301.]

Several other transcriptional regulatory proteins are essential for the normal function of both shoot and root meristem cells, including the plant homolog of the human retinoblastoma (Rb) tumor suppressor protein (see Chapter 25), called *RBR*. As in animal cells, RBR binds to and inhibits the function of an E2F transcription factor; release of RBR from E2F or genetic loss of RBR allows the E2F factor to promote transcription of multiple genes that promote entry into the cell cycle and cell division (see Figure 19-16). Reduced levels of RBR result in an increase in stem-cell numbers, and increased RBR levels lead to stem-cell differentiation; both of these observations indicate a prominent role for RBR in stem-cell maintenance.

The Root Meristem Resembles the Shoot Meristem in Structure and Function

Unlike the shoot meristem, the root meristem consists of lineage-restricted stem cells. These cells are organized around the *quiescent center*, four very slowly dividing cells that serve as the stem-cell niche (Figure 22-21b). Stem-cell division is asymmetric (also unlike that in the shoot), and the daughter cell that loses contact with the quiescent center divides several more times and then differentiates. A WUS homolog, WOX5, is expressed in the quiescent center and is required for stem-cell maintenance, although other

transcription factors are also important. The plant hormone auxin (indole-3-acetic acid) coordinates many processes involved in plant growth and differentiation; in particular, it is essential for formation of the root meristem niche. If the quiescent center is ablated, a new niche is formed in an area of high auxin concentration. However, the effect of auxin on stem cells depends on the specific cell type. For example, in the stem cells that give rise to the root cap, auxin promotes cell differentiation by repressing *WOX5* via auxin-responsive transcription factors.

Plants have an amazing capacity for regeneration. The home gardener will be familiar with the ability of some leaf or stem cuttings to form roots with little inducement beyond a glass of water and a sunny windowsill. Experiments performed in the mid-twentieth century demonstrated that single cells isolated from carrot roots could regenerate entire plants when placed on media containing the appropriate mix of nutrients and hormones. Recent studies have utilized detailed molecular analyses, including sequencing the transcriptomes of single cells and lineage tracing, to the process of regeneration. In some instances, regeneration appears to follow a program similar to the embryonic processes that initially establish meristems. For example, excision of the root tip removes the quiescent center and disrupts the auxin gradient. An auxin maximum is then established at the new distal tip via polar auxin transport. *WOX5* initially exhibits a broad expression domain but gradually becomes restricted to a newly formed quiescent center

as hormone gradients are reestablished and the meristem is reconstructed. ■

KEY CONCEPTS OF SECTION 22.2

Stem Cells and Niches in Multicellular Organisms

- Planaria contain pluripotent stem cells termed *cNeoblasts* that are important for regeneration of body parts removed by amputation.
- Most stem cells in animals are multipotent, except for germ-line stem cells that are unipotent.
- Stem cells are undifferentiated; they can undergo symmetric or asymmetric self-renewal divisions such that their number remains constant or increases over the organism's lifetime (see Figure 22-12).
- Stem cells are formed in niches that provide signals to maintain a population of undifferentiated stem cells. The niche must maintain stem cells without allowing their excess proliferation and must block differentiation.
- Stem cells are prevented from differentiating by specific controls that operate in the niche. A high level of β-catenin, a component of the Wnt signaling pathway, has been implicated in preserving stem cells in the germ line and intestine by directing cells toward self-renewal division rather than differentiation states.
- In the *Drosophila* germarium, a few cells form the germ stem-cell niche, sending signals directly to the adjacent stem cells. Daughter cells that are displaced from the niche cells undergo proliferation and differentiation into germ cells (see Figure 22-13).
- Populations of stem cells associated with the intestinal epithelium and many other tissues regenerate differentiated tissue cells that are damaged, sloughed, or aged (see Figure 22-14).
- Intestinal stem cells reside in the bases of intestinal crypts, adjacent to Paneth cells, and are marked by expression of the Lgr5 receptor (see Figure 22-14).
- In the blood-cell lineage, different precursor types form and proliferate under the control of distinct cytokines (see Figure 22-18). This system allows the body to specifically induce the replenishment of some or all of the necessary blood-cell types.
- Hematopoietic stem cells can be detected and quantified by bone marrow transplant experiments (see Figure 22-19) and their niche cells detected using a combination of marker surface proteins (see Figure 22-20).
- Plant stem cells persist for the life of the plant in the meristem. Meristem cells can give rise to a broad spectrum of cell types and structures (see Figure 22-21).

- A negative feedback loop involving the WUS transcription factor maintains the size of the shoot apical stem-cell population.

22.3 Mechanisms of Cell Polarity and Asymmetric Cell Division

We have discussed the importance of asymmetric division in generating cell diversity during development and in maintaining the number of stem cells in a population. What mechanisms underlie the ability of cells to become asymmetric before cell division to give rise to cells with different fates? Cell asymmetry is a concept we have encountered before, under the name of *cell polarity*, so let us first review what it means for a cell to be polarized.

Cell polarity — the ability of cells to organize their internal structure, resulting in changes in cell shape and the generation of regions of the plasma membrane with different morphologies and protein and lipid compositions — has been introduced in several chapters. For example, we have seen that polarized intestinal epithelial cells have an apical domain with abundant microvilli separated from the basolateral domain by tight junctions (see Figures 17-1 and 20-11). Epithelial transport requires these cells to have different transport proteins in the apical and basolateral membranes (see Figure 11-30). As we discuss later in this section, these epithelial cells are responding to extracellular signals that instruct them how to polarize. These cells represent just one example of cell polarity — essentially all cells in animals are polarized, and we discuss several examples in which the underlying mechanisms have been defined. What emerges are three core principles of cell polarity. First, cells

have an *intrinsic polarity program*, as revealed by their remarkable ability to polarize in the absence of external cues. As we will see in our examples, a master and common regulator of this program is the small GTPase Cdc42. Second, this intrinsic polarity program can be directed by external or internal *cues*. Third, the polarity of individual cells is often maintained by intracellular *mutually antagonistic complexes*.

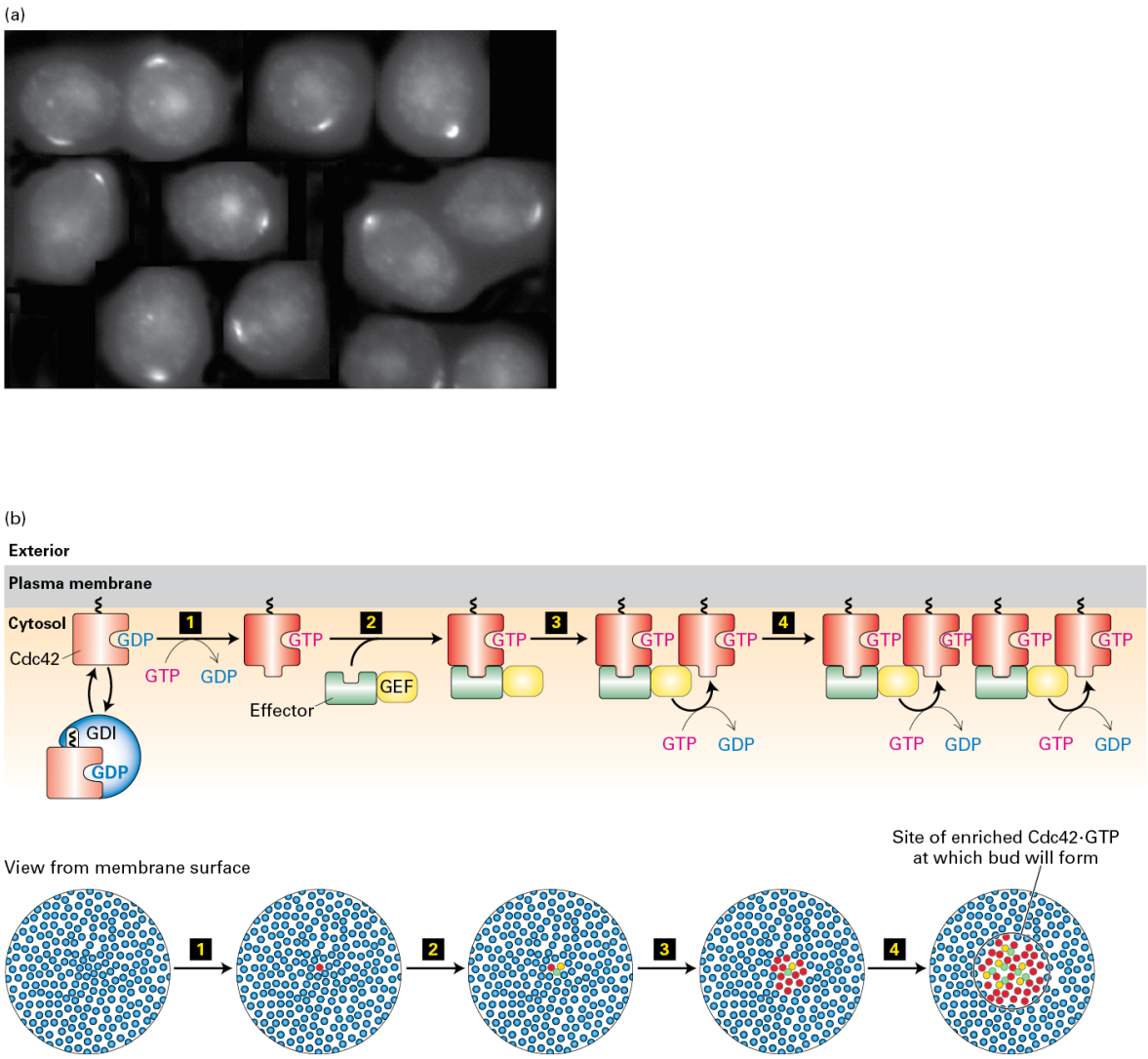
We first discuss the intrinsic polarity program in budding yeast because, given that all the components of the mechanism are shared with animals, the principles uncovered in yeast are likely to be conserved. We then turn our discussion to examples in which cells respond to external cues to establish cell polarity depending on antagonistic interactions. Finally, we discuss an example of asymmetric cell division that gives rise to a daughter stem cell and a differentiated cell.

The Intrinsic Polarity Program Depends on a Positive Feedback Loop Involving Cdc42

Budding yeast grows by selecting a single site on its surface at which to assemble a new bud (see Figure 19-4). Importantly, it must reliably select just one site. If a cell grew two buds simultaneously, during mitosis the duplicated chromosomes might be segregated

between the parent cell and one bud, leaving the other bud without chromosomes and therefore inviable. In haploid yeast, this so-called singularity of budding is guided by a signal, or cue, left at the cell surface, which directs the next budding event to a site adjacent to the former budding site. Remarkably, if the genes that specify the nonessential components of this cue are deleted, yeast cells grow just as well, but each assembles a single bud at a random site. Thus yeast has an intrinsic polarity program that, even in the absence of cues from the previous budding cycle, can guide selection of a single site for bud growth. This program requires the concentration of Cdc42 at the site from which a bud will emerge.

Surprisingly, Cdc42 concentration at the site for a new bud does not depend on either actin filaments or microtubules, as this small GTPase localizes to a single spot even when both filament systems are disrupted (Figure 22-23a). Long before biologists had thought about how this might occur, the brilliant mathematician and computer pioneer Alan Turing considered what mechanism might shift a uniform distribution of a factor to a concentration at a single site. In 1952, Turing suggested that such a shift could be achieved if a positive feedback reaction amplified a random increase in the concentration of the polarity factor — and he was right!



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 22-23 The intrinsic polarity program of budding yeast involves a positive feedback loop for activation of the GTPase Cdc42. (a) Diploid yeast lacking polarity cues show polarized Cdc42, visualized here by immunofluorescence microscopy, when they are about to assemble a bud. The cells were treated with drugs to disassemble both actin filaments and microtubules to show that polarization of Cdc42 is not dependent on these cytoskeletal filaments. (b) Positive feedback loop for activation of Cdc42. Inactive Cdc42·GDP is in equilibrium between a cytosolic pool of complexes with the guanine nucleotide dissociation inhibitor (GDI) and a membrane-bound pool. Step 1: One of the membrane-associated Cdc42·GDP proteins may spontaneously become an activated Cdc42·GTP. Step 2: Active Cdc42·GTP recruits a complex containing the guanine nucleotide exchange factor (GEF). Step 3: The recruited GEF now locally converts more Cdc42·GDP to Cdc42·GTP. Step 4: This active Cdc42·GTP recruits more GEF, driving a

positive feedback loop that results in the local accumulation of Cdc42·GTP (visualized as bright red puncta). See J. G. Chiou, M. K. Balasubramanian, and D. J. Lew, 2017, *Annu. Rev. Cell Dev. Bi.* **33**:77.

[Part (a) reprinted with permission from Nature Publishing Group, from J. E. Irazoqui, A. S. Gladfelter, and D. J. Lew, 2003, "Scaffold-Mediated Symmetry Breaking by Cdc42p," *Nat. Cell Biol.* **5**(12):1062–1070; permission conveyed through Copyright Clearance Center, Inc.]

Recall that Cdc42 is a member of the Rho family of small GTP-binding proteins (see Figure 17-40). It acts as a molecular switch, existing in an inactive (Cdc42·GDP) and an active membrane-bound (Cdc42·GTP) state. Binding to its specific guanine nucleotide exchange factor (Cdc42-GEF) causes Cdc42 to release GDP and bind GTP. The active Cdc42·GTP binds many effectors and thereby activates downstream signaling events. In its inactive state, Cdc42·GDP exists in equilibrium between a cytosolic pool, which is bound to a guanine nucleotide dissociation inhibitor (GDI), and a membrane-bound Cdc42·GDP pool (Figure 22-23b). Occasionally and randomly, the membrane-bound Cdc42·GDP will release its GDP and bind GTP, which converts it to the active Cdc42·GTP state (Figure 22-23b, step **1**). One of the effectors that is recruited to Cdc42·GTP is a protein complex that contains the Cdc42-GEF (Figure 22-23b, step **2**). Thus when an active Cdc42 arises in the plasma membrane, it recruits Cdc42-GEF, which locally activates more Cdc42, which recruits more Cdc42-GEF, and this simple positive feedback loop generates a site on the cell surface that is highly and locally enriched for Cdc42·GTP (steps **3** and **4**). Computational modeling — also pioneered by Turing — shows that this system can result in a

winner-takes-all scenario to yield just one site of polarization. This positive feedback loop is the core mechanism by which one budding site is generated in yeast.

As we discussed in Chapter 18, Cdc42 is the master regulator that guides the polarization of migrating cells (see Figure 18-54) and is also involved in regulating many additional examples of cell polarity. It is important to note that in cases in which polarization needs to be flexible, negative feedback loops ensure that the single site of polarization is not too strong, so that it can be redirected to another site on the cell surface upon receiving appropriate signals. For example, in addition to its fast-acting GEF, Cdc42·GTP might also recruit a slow-acting negative regulator that modulates the degree of positive feedback. In fact, yeast Cdc42 recruits a kinase that phosphorylates and inhibits the recruited Cdc42-GEF, thereby introducing a negative feedback loop. Thus the local concentration of Cdc42·GTP builds up fast, then levels off or disappears as the slower negative regulator comes into play. As we will see, specific cues normally guide these intrinsic polarity programs, which in turn lead to the physical polarization of the cell.

Cell Polarization Before Cell Division Follows a Common Hierarchy of Steps

The polarization of a cell, with or without cell division, follows the general pattern diagrammed in Figure 22-24a. In order to know in which direction to polarize, or become asymmetric, a cell generally senses specific cues that provide it with spatial information (step **1**). Such cues can be soluble signals from other cells or from the extracellular matrix. To be receptive to these cues, a cell must have appropriate receptors or other machinery on its surface (step **2**). Once the cues have been detected, the cell responds by feeding the incoming signal into its polarity program to define the orientation of polarity (step **3**). Generally the next step involves the local reorganization of cytoskeletal elements, notably microfilaments and microtubules (step **4**). Once the cell has structural asymmetry, molecular motors direct the trafficking of polarity factors — which, depending on the system, may be cytoplasmic proteins or membrane proteins synthesized by the secretory pathway, or both — to their appropriate locations (step **5**). The polarity can be reinforced or maintained by moving the polarity determinants from sites of lower concentration back to the polarization site to maintain the highest concentration there (step **6**).

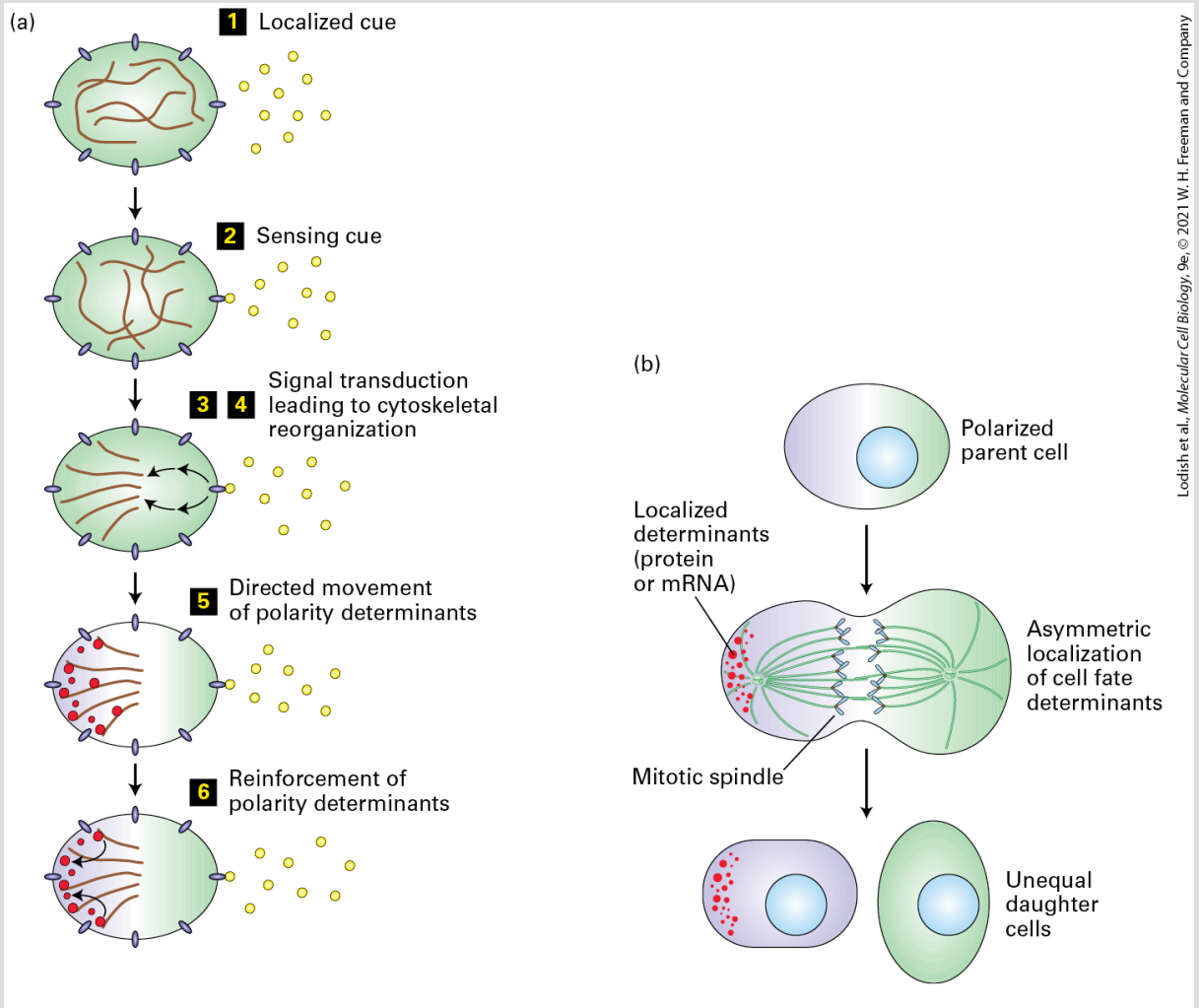


FIGURE 22-24 General features of cell polarity and asymmetric cell division. (a) General hierarchy of the steps in generating a polarized cell. See text for details. (b) To generate cells with different fates, specific determinants, including mRNAs, proteins, and lipids, must be asymmetrically distributed in a cell. If the mitotic spindle is positioned so that these determinants are segregated during cell division, the two daughter cells will have different fates. However, if the mitotic spindle is not oriented appropriately, the determinants will not be segregated properly, and the daughter cells could have the same fate (not shown).

If a cell becomes polarized and cell division then occurs in a plane perpendicular to the direction of polarization, the cell has undergone asymmetric cell division. In this way, fate determinants,

such as specific proteins or mRNAs, can be differentially segregated between the cells (Figure 22-24b).

Cell polarization can be a very dynamic process. Consider a macrophage chasing a bacterium in order to catch and destroy it by phagocytosis: the macrophage must continually sense the bacterium, which it does by following a gradient of a peptide left by the bacterium (see Figure 17-45). This signal orients — or polarizes — the macrophage to move in the correct direction. This example highlights an important and common aspect of cell polarity: in many cases, it must be dynamic so that it can quickly change direction. Although we have illustrated the dynamics of cell polarity in terms of a macrophage, the polarity of epithelial and other cells that appear very static is probably quite dynamic when those cells are moved to different environments.

In the next section, we discuss a simple cell that shows asymmetry: a yeast cell responding to a soluble cue during mating. In later sections, we turn to animal cells, in which conserved polarity proteins are instrumental in interpreting polarity cues and generating cell asymmetry prior to cell division. We then consider how these same polarity proteins are used to polarize epithelial cells. Finally, we discuss aspects of asymmetric division in stem cells.

Polarized Membrane Traffic Allows Yeast to Grow Asymmetrically During Mating

One of the simplest and best studied forms of cell asymmetry occurs when budding yeast cells mate. As we have seen, yeast can exist in a haploid state (with a single copy of each chromosome) or a diploid state (with two copies of each chromosome). The haploid state can exist in either of two mating types ("sexes"), **a** or α. The preferred state of yeast in nature is the diploid state, so **a** cells are always looking to mate with α cells to restore the diploid state (see Figure 1-24b). Each mating type secretes a specific mating pheromone — **a** cells secrete **a** factor and α cells secrete α factor — and each cell expresses on its surface a receptor that senses the pheromone of the opposite mating type. Thus **a** cells have a receptor for α factor and α cells have a receptor for **a** factor (see Figure 16-15). When cells of opposite mating types are placed near each other, the receptors on each cell bind and detect the pheromone cue of the other cell and determine its spatially highest concentration in order to know in which direction to mate. When the cells detect the opposite mating factor, two important processes occur. First, they synchronize their cell cycles by arresting at G_0 so that when they mate, the two haploid genomes will be at the same stage of the cell cycle. Second, they target cell growth in the direction of the pheromone to assemble a mating projection called a *shmoo*. If shmooing cells of opposite

mating types touch, they fuse at the shmoo tips, and the haploid nuclei come together to restore the diploid state.

By looking for mutants in yeast haploids that cannot shmoo in response to the opposite mating pheromone, researchers have discovered how the asymmetric growth necessary for shmoo formation occurs (Figure 22-25). This mechanism initially involves a signal transduction pathway that establishes a polarized cytoskeleton, which in turn guides membrane traffic to the appropriate location for asymmetric growth. Activation of the mating-factor receptor — a typical G protein–coupled receptor (see Figures 15-12 and 16-15) — results in activation of the intrinsic polarity program, which in turn results in the localized accumulation and activation of Cdc42 in the region of the cell cortex closest to the pheromone source (Figure 22-25, step **1**). This active Cdc42·GTP leads to the local activation of a formin protein (step **2**). As we saw in Chapter 17, formin proteins nucleate the assembly of polarized actin filaments, whose (+) ends remain bound to the formin (see Figure 17-13). This process provides the tracks for the transport of secretory vesicles by a myosin V motor to the (+) ends of the filaments for localized growth and hence shmoo formation (step **3**). Notice that this mechanism requires polarity proteins, which include Cdc42·GTP, to remain concentrated at the growing shmoo tip. To ensure that polarity is maintained during shmoo growth, a directed endocytic cycle is believed to exist. In this cycle, polarity factors that have diffused away from the site of

concentration may be internalized by endocytosis and transported back to the shmoo tip, thereby reinforcing polarity (step 4).

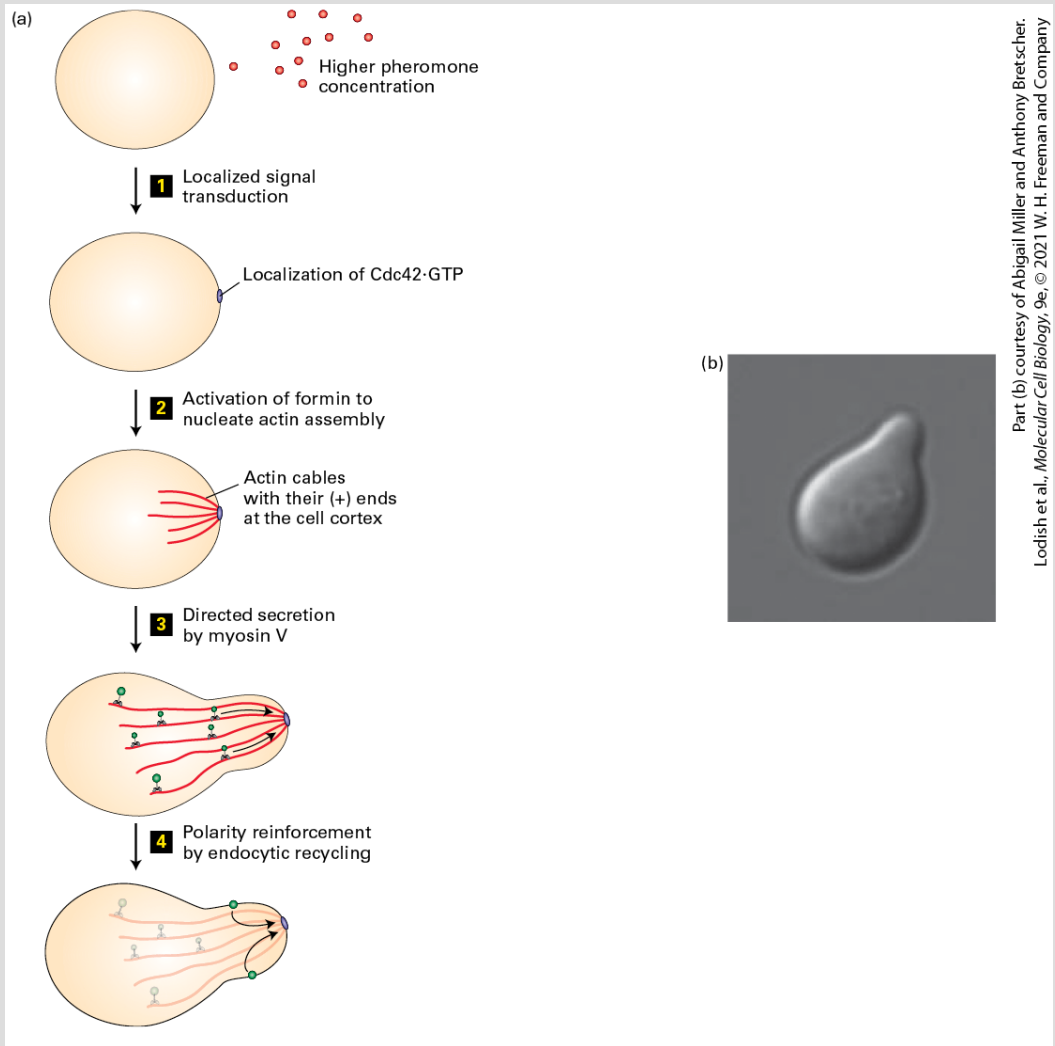
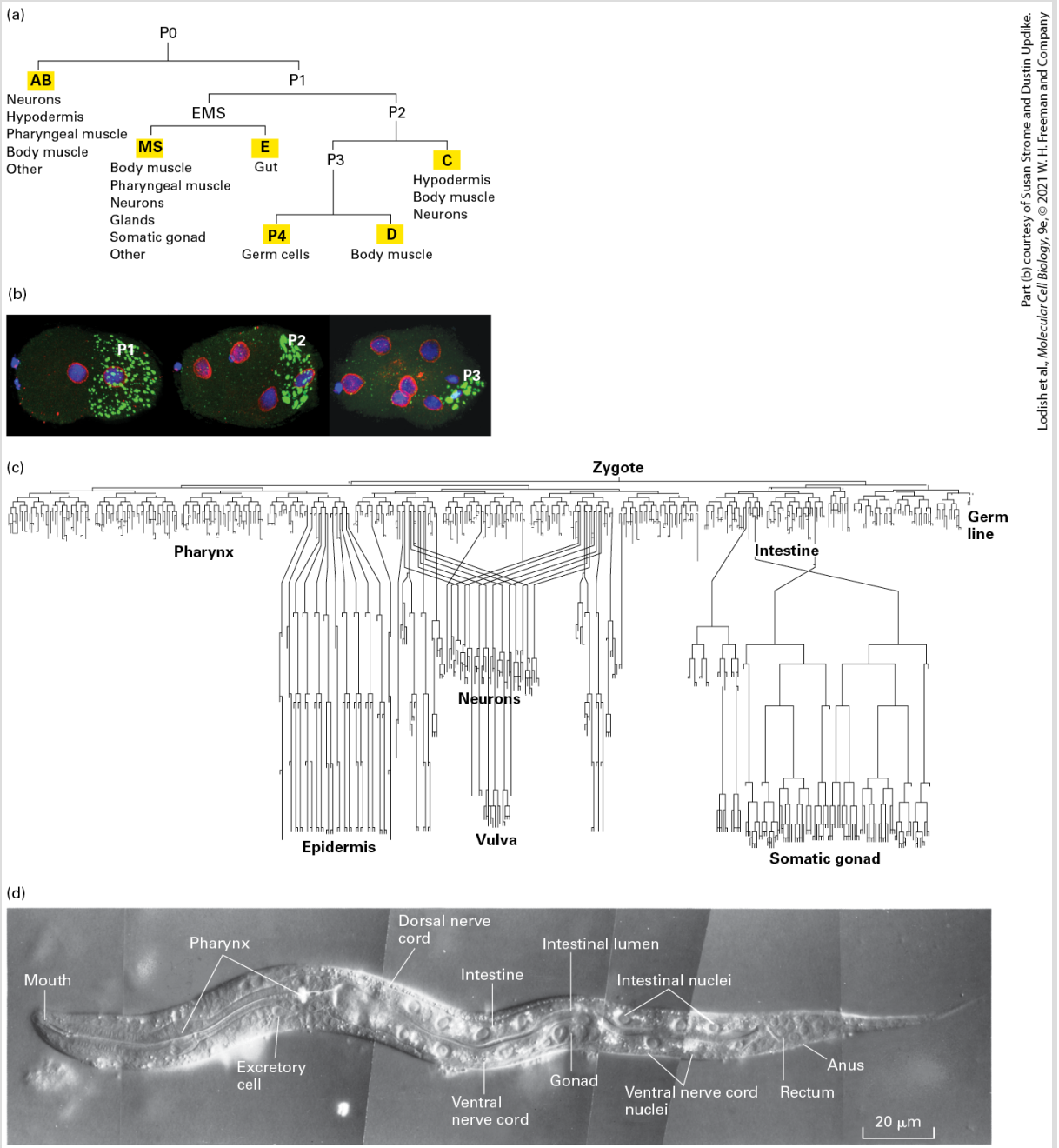


FIGURE 22-25 Mechanism of shmoo formation in yeast. (a) The haploid yeast cell must grow toward the highest concentration of mating factor of the opposite mating type: a receptor on its surface signals the location of the highest concentration. This signal induces the localization and activation of Cdc42 to generate a higher concentration of Cdc42·GTP at this site (step 1). The Cdc42·GTP locally activates a formin, which nucleates and elongates actin filaments from this site (step 2). Because formins bind to the (+) ends of actin filaments, the (+) ends are oriented toward Cdc42·GTP and thus the highest concentration of the mating factor. A myosin V motor transports

secretory vesicles along the actin filaments, resulting in the growth of the shmoo (step 3). The polarity of the shmoo can be reinforced by an endocytic cycle that returns diffusing polarity factors back along the actin filaments to the signal site (step 4). (b) Differential interference contrast (DIC) light-microscope image of a shmooing yeast cell.

The Par Proteins Direct Cell Asymmetry in the Nematode Embryo

The nematode roundworm *Caenorhabditis elegans* has provided a powerful model system for understanding cell fate decisions (see Figure 1-23c). It was selected for study because the animal is transparent and has a rapid life cycle; it is easy to generate and characterize mutants; and the lineage of cells from the one-cell embryo to the adult is invariant (Figure 22-26a, c, d). A critical aspect of this lineage is the first cell division, in which the P0 cell — the fertilized egg, or zygote — gives rise to the AB and P1 cells by an asymmetric cell division; each of these two cells then gives rise to different lineages. Much is known about this first asymmetric division, which is where we focus our attention.



Part (b) courtesy of Susan Strome and Dustin Updike.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

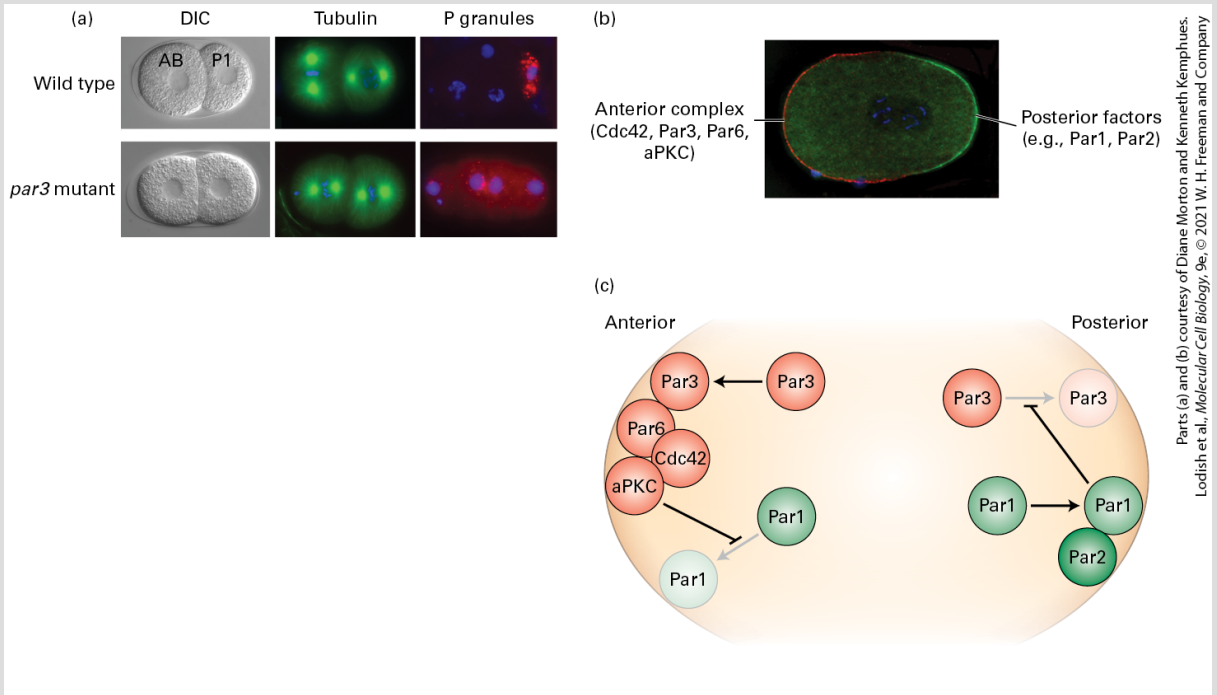
FIGURE 22-26 Cell lineage in the nematode worm *C. elegans*. (a) Pattern of the first few divisions, starting with P0 (the zygote) and leading to formation of the six founder cells (yellow highlights). The first division is asymmetric, giving rise to the AB and P1 cells. The EMS cell is so named because it gives rise to most of the endoderm and mesoderm. The P4 lineage gives rise to the cells of the germ line. (b) Micrographs of two-, four-, and eight-cell embryos. DNA is stained blue, the nuclear envelope red, and P granules green. The P1, P2, and P3 cells, which give rise to the germ line, are indicated. (c) The full lineage of the entire

body of the worm, showing some of the tissues formed. Cell division is indicated by the splitting of a line, and the time of cell division is indicated in the vertical direction. (d) Newly hatched larva. Some of the 959 somatic-cell nuclei found in the adult hermaphrodite form can be seen in this micrograph obtained by DIC microscopy.

[Part (d) republished with permission from Elsevier, from J. E. Sulston and H. R. Horvitz, 1977, "Post-Embryonic Cell Lineages of the Nematode, *Caenorhabditis elegans,*" *Devel. Biol.* **56**(1):110–156; permission conveyed through Copyright Clearance Center, Inc.]

Before the first cell division, the zygote is visibly asymmetric: cytoplasmic complexes called *P granules* are concentrated in the region of the cell that will give rise to the posterior end of the embryo (Figure 22-26b). It turns out that during further cell divisions, these P granules always segregate to cells that will give rise to the germ line, where they ultimately play an important role in germ-line development. The first asymmetric division of the P0 cell gives rise to the P1 cell, containing the P granules, and the larger AB cell. Following that, at the two-cell stage, the mitotic spindles are arranged at right angles to one another so that the ensuing cell divisions are also at right angles to one another (Figure 22-27a). To begin to understand how this first essential asymmetric division occurs, mutations in six different genes were identified that resulted in a symmetric first division. Since the P granules were not partitioned correctly in these mutants, the genes identified in this study were called *partition defective*, or *par*, genes. In these mutants, P granules did not properly localize to the posterior end of the zygote, and the mitotic spindles were not oriented correctly in preparation for the second division (see Figure 22-27a). A key insight

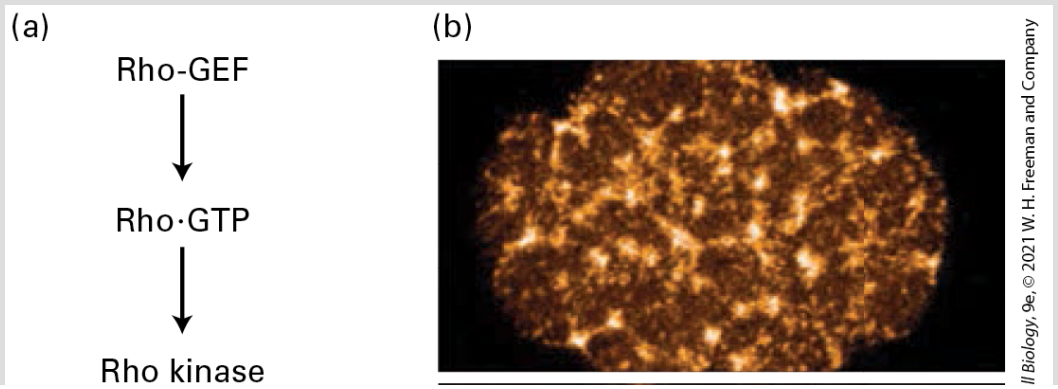
came when the products of the *par* genes were localized. In wild-type zygotes, many of the Par proteins localize either at the cortex of the anterior half of the cell or at the cortex of the posterior half. For example, Par3 localizes anteriorly and recruits the other components of the anterior complex comprising Par6 and aPKC (atypical protein kinase C), while Par1 and Par2 localize posteriorly (Figure 22-27b). Subsequent work has shown that mutually *antagonistic interactions* exist between these protein complexes; that is, if the Par3-Par6-aPKC complex is localized to one region, it excludes Par2, and vice versa. This is shown by the finding that the Par3-Par6-aPKC complex spreads over the whole cortex in *par2* mutants and Par2 spreads over the whole cortex in *par3* or *par6* mutants. The molecular nature of this antagonism to generate mutual exclusion is partly mediated by the protein kinase aPKC, which phosphorylates any local Par1, inhibiting its ability to bind to the anterior cortex. In a reciprocal manner, the Par1 kinase phosphorylates any local Par3 to inhibit its association with the cortex (Figure 22-26c). Additionally, the anteriorly localized Par6 now recruits Cdc42·GTP to maintain the functional polarity of the zygote, and a GTPase activating protein Cdc42-GAP that deactivates Cdc42. GTP is recruited to the posterior end to ensure that any active Cdc42·GTP is confined to the anterior end.



EXPERIMENTAL FIGURE 22-27 Par proteins are asymmetrically localized by antagonistic interactions in the one-cell worm embryo. (a) DIC images of wild-type and *par3* mutant embryos. In wild-type cells, the AB cell is larger than the P1 cell, whereas they are the same size in the *par3* mutant. The *par3* mutant also has a defect in spindle orientation (as seen by microtubule staining in green) and P-granule (red) segregation. DNA is stained blue. (b) Complementary localization of the anterior Par complex (Par3-Par6-aPKC) (red) and posterior determinants (green) in the one-cell embryo. (c) Antagonistic interactions between the anterior and posterior Par complexes. When Par3 becomes localized to the anterior it recruits the other members of the complex, Par6 and aPKC. The kinase activity of aPKC can phosphorylate local Par1, which inhibits its association with the anterior cortex. Par1 and Par2 associate with the posterior complex, and Par1 kinase can phosphorylate any Par3 to locally inhibit its association with the posterior cortex.

The unfertilized egg is symmetric, so what breaks this symmetry to generate a polarized zygote? It turns out that the position of the sperm that fertilizes the egg determines the posterior end. Prior to sperm entry, the entire egg cortex is under tension provided by an actin meshwork containing active myosin II. As we saw in Chapter

17, myosin II can form bipolar filaments that pull on actin filaments to generate tension, as is also seen in muscle and the contractile ring. Myosin II activity is regulated by a signal transduction pathway involving the small GTPase Rho (see Figure 17-42). In the unfertilized egg, Rho is maintained in its active Rho·GTP state by the uniform distribution of its activator, the guanine nucleotide exchange factor Rho-GEF. Rho·GTP activates Rho kinase, which phosphorylates the myosin light chain of myosin II to activate it (Figure 22-28a). Shortly after fertilization, the Aurora-A kinase activity associated with the sperm centrosome results in the local depletion of Rho-GEF activity. Thus the asymmetric position of the sperm centrosome defines the posterior region by local depletion of Rho·GTP, thereby lowering the activity of myosin II. With this local reduction in contractile activity, the actin-myosin network contracts toward the anterior (Figure 22-28b), and as it does so, it transports Par3 to the anterior region, where Par3 then recruits the additional components of the anterior complex (Par6 and aPKC) (Figure 22-28c). With the removal of the anterior complex, Par2 can now occupy the posterior cortex, the antagonistic interactions described above come into play, and cell asymmetry is established.



ll Biology, 9e, © 2021 W. H. Freeman and Company

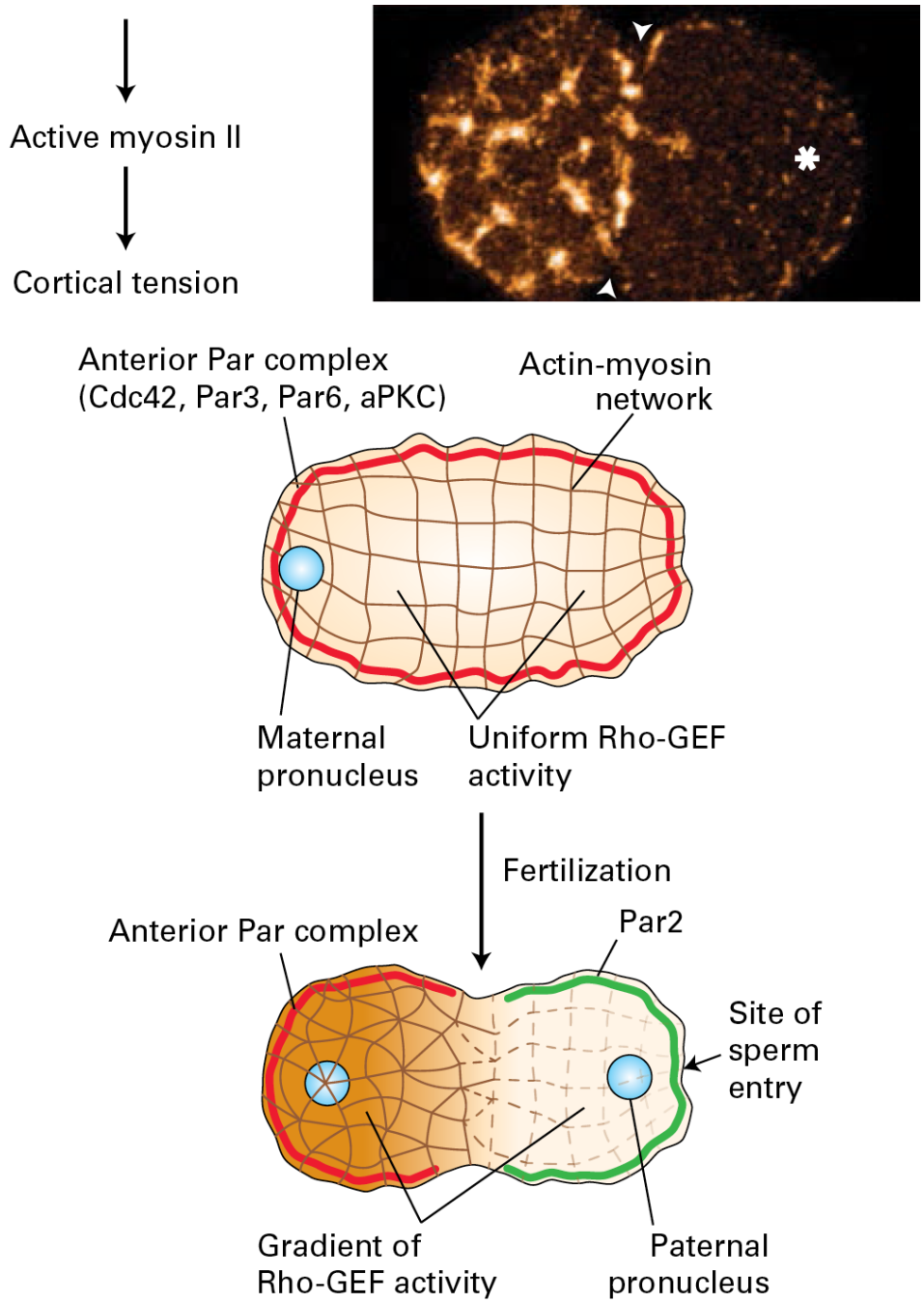


FIGURE 22-28 Mechanism of segregation of the anterior Par complex in the one-cell worm embryo. (a) Before fertilization, the cell cortex is under tension due to the activity of Rho-GEF, the guanine nucleotide exchange factor for the small GTPase Rho. Rho·GTP activates Rho kinase, which phosphorylates the regulatory light chain of myosin II to activate it. Together with actin filaments, the active myosin II maintains tension in the cell cortex. (b) Localization of myosin II before (*top*) and after (*bottom*) fertilization. The asterisk marks the

region of sperm entry. (c) Before fertilization, Rho-GEF is uniformly active, the cortex is under tension from active myosin II, and the anterior Par complex is uniformly distributed around the cortex. Upon fertilization, Rho-GEF becomes locally reduced, resulting in local deactivation of myosin II. This deactivation generates unequal tension, so the actin-myosin network contracts toward the future anterior end, moving Par3 with it, which then associates with Par6 and aPKC. Once the anterior complex is localized, factors such as Par2 associate with the posterior cell cortex.

[Part (b) republished with permission from Elsevier, from E. Munro, J. Nance, and J. R. Priess, 2004, "Cortical Flows Powered by Asymmetrical Contraction Transport PAR Proteins to Establish and Maintain Anterior-Posterior Polarity in the Early *C. elegans* Embryo," *Dev. Cell*, 2004, **7**(3):413–424; permission conveyed through Copyright Clearance Center, Inc. Part (c) Data from D. St. Johnston and J. Ahringer, 2010, *Cell* **141**:757.]

It turns out that the master regulator Cdc42 is not needed for the initial asymmetry induced by actin-myosin network contraction. However, active Cdc42·GTP binds Par6 and is necessary for maintaining the anterior complex at the anterior end. Recent work has also implicated an endocytic reinforcement cycle, as we discussed for yeast shmoo formation, to maintain polarity. Thus the steps of responding to a spatial cue, establishing asymmetry, and maintaining asymmetry are conserved features of both systems.

The Par Proteins and Other Polarity Complexes Are Involved in Epithelial-Cell Polarity

In vertebrates, polarized epithelial cells use cues from adjacent cells and the extracellular matrix to orient their axis of polarization. The process of polarization in epithelial cells of vertebrates is quite similar to that in *Drosophila*. Much of our knowledge has come from the fly system because of the ease with which mutants can be isolated and analyzed.

Genetic screens in the fly have uncovered multiple genes necessary for the generation of epithelial-cell polarity. Analyses of the proteins encoded by these genes and of the phenotypes of mutants have identified three major groups of proteins: the complex made up of Cdc42, Par3, Par6, and aPKC (in this system known as the *apical Par complex*, or simply as the *Par complex*), the Crumbs complex, and the Scribble complex. By extensive analyses of the effects of these complexes on one another when individual components are missing, a general understanding of their contributions to epithelial-cell polarization has been achieved, although a detailed molecular understanding is still emerging (Figure 22-29a).

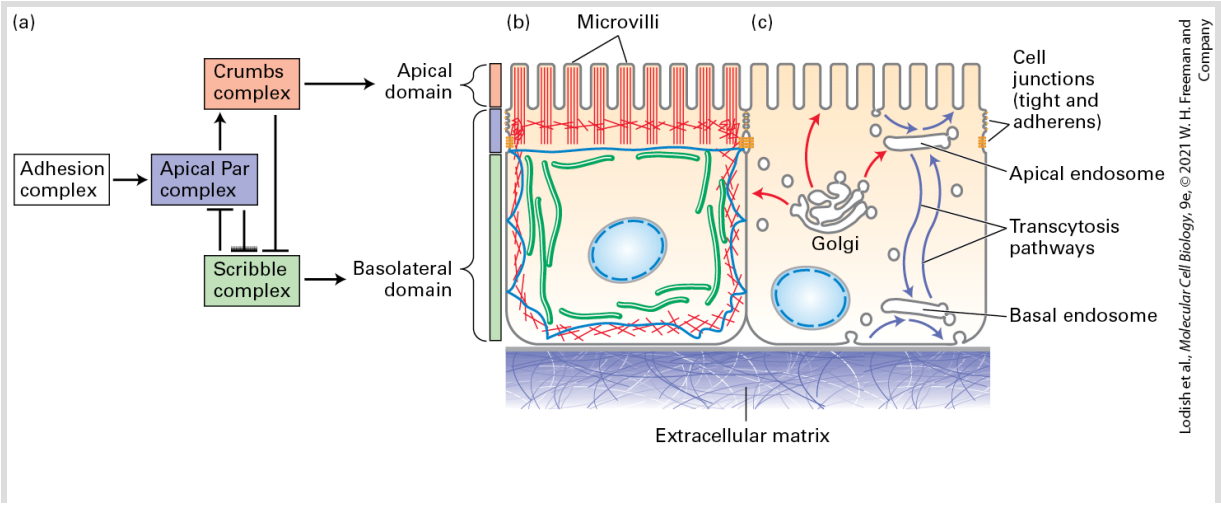


FIGURE 22-29 Establishment of polarity in epithelial cells. (a) Polarity determination in epithelial cells is driven by an apical Par complex. The formation of a cell-cell adhesion complex induces the recruitment of the Par complex. Then, intricate and antagonistic interactions of the Par complex with both the basolateral Scribble complex and the apical Crumbs complex lead to the establishment and maintenance of epithelial-cell polarity. The localization of the different complexes to membrane domains is indicated by colored bars: the Scribble complex associates with the lateral membrane, the Par complex associates with the region at the cell junctions, and the Crumbs complex is immediately apical to the Par complex. Functional epithelial polarity is maintained by both (b) a polarized cytoskeleton and (c) membrane trafficking pathways. In the biosynthetic pathway, proteins and lipids destined for the apical and basolateral domains are sorted in the Golgi complex and transported to their respective surfaces (red arrows). Endocytic pathways (blue arrows) regulate the abundances of proteins and lipids on each surface and sort them between surfaces by transcytosis.

The first known steps in epithelial-cell polarization are interactions with the extracellular matrix and adjacent cells. The transmembrane protein b1-integrin binds to collagen in the extracellular matrix, signaling that this is the basal end of the cell. Interaction with adjacent cells in vertebrates occurs through nectin, a cell-adhesion molecule in the Ig superfamily, and a junctional protein called *JAM-A*. This induces the formation of adherens and tight junctions (see Figure 20-1) and then the apical recruitment of Par3. By Par3 interacting with other members of the apical complex — Par6 and aPKC — the Crumbs complex is recruited more apically than the Par complex, and the Scribble complex defines the basolateral surface. In the absence of the Par complex, cells cannot polarize, and, as in the nematode embryo, the Par complex is the master regulator of cell polarity. In the absence of the Scribble

complex, the apical domain is greatly expanded, whereas in the absence of Crumbs, the apical domain is greatly reduced. These observations have led to the realization that there are mutually antagonistic relationships between these complexes, in which, for example, the apical Par complex kinase aPKC antagonizes the basolateral Scribble complex by phosphorylation, and the basolaterally localized Par1 kinase antagonizes the apical Crumbs complex (see Figure 22-29a). Thus, as is the case in the nematode embryo, asymmetry is mediated by polarity complexes working antagonistically against each other.

In a manner that is only partially understood, this arrangement of polarity proteins reorganizes the cytoskeleton so that distinct arrangements of microfilaments become associated with the apical and basolateral membranes. The distribution of microtubules in epithelial cells is rather unusual, as they do not all associate with a centrosome; instead, lateral microtubules orient their (−) ends toward the apical domain and other microtubules run perpendicular to the lateral microtubules below the microvilli and also along the base of the cell (Figure 22-29b); how these arrangements are established is not known. Membrane traffic is also polarized (Figure 22-29c). Newly made membrane proteins destined for the apical and basolateral membranes are sorted and packaged into distinct transport vesicles at the *trans*-Golgi network and then transported to the appropriate surface. In addition, endocytic pathways from both the apical and basolateral surfaces regulate the abundance of

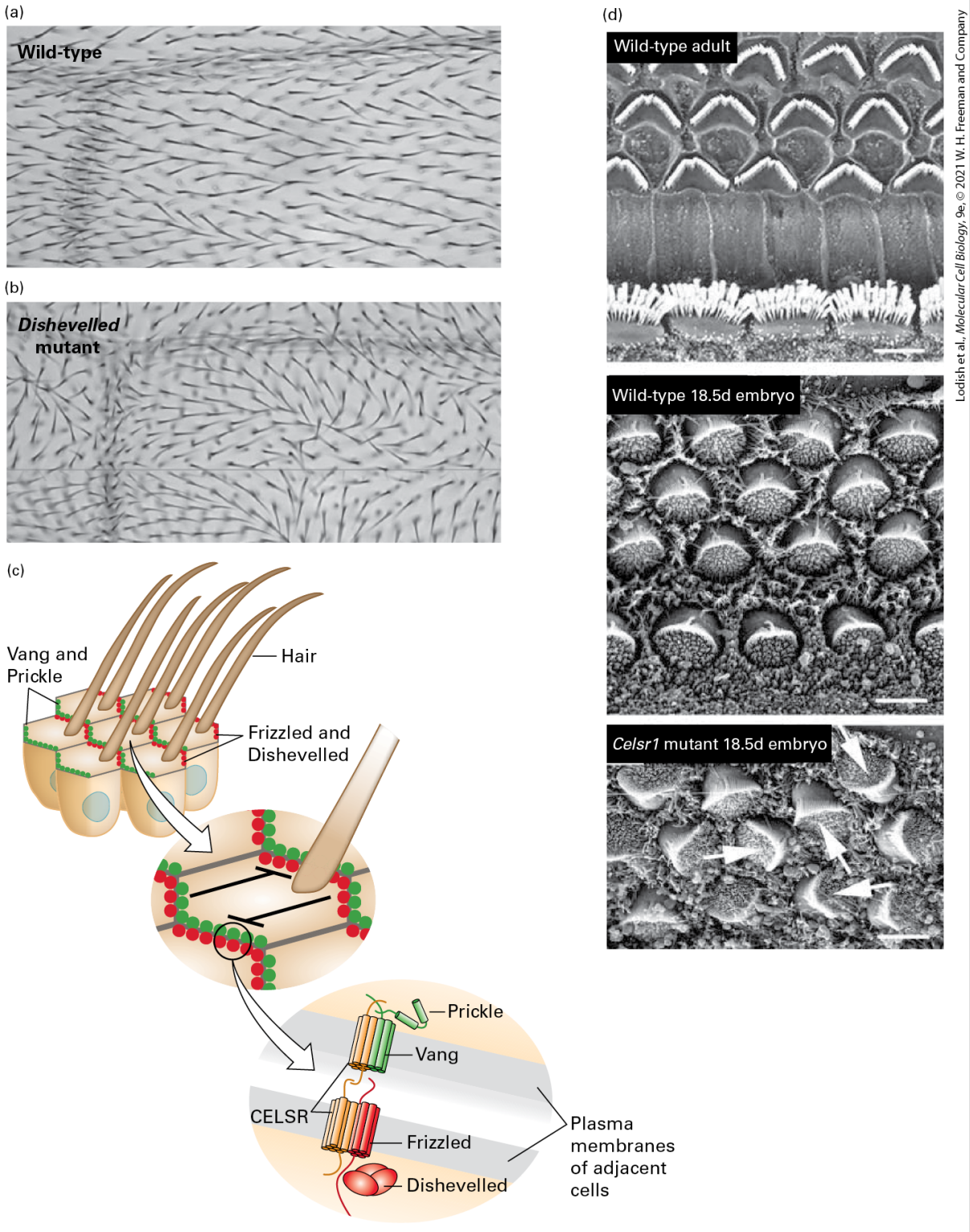
membrane proteins and transport missorted proteins using a complex set of sorting endosomes in a process known as *transcytosis*.

In genetic screens for additional components important for epithelial-cell polarity in the fly, components of endocytic trafficking were found. For example, one such mutant affects the trafficking of the apical transmembrane protein Crumbs, so that when endocytosis is compromised, the level of Crumbs on the surface goes up and the apical domain expands. Thus epithelial-cell polarity involves responses to spatial cues and reorganization of the cytoskeleton that provides a framework for both secretory and endocytic membrane traffic pathways for establishment and maintenance of the polarized state.

The Planar Cell Polarity Pathway Orients Cells Within an Epithelium

We have so far discussed asymmetry in only one dimension, but in many cases cells in multicellular organisms are polarized in at least two dimensions — top to bottom, and along a body axis. Just looking at the scales of fish, the feathers of birds, or the hairs on your arm makes it clear that the groups of cells that give rise to these structures must be organized not only in a top-to-bottom (apical/basal) manner but also in a head-to-tail, or proximal/distal,

manner. This type of polarity is called **planar cell polarity (PCP)**. A well-studied example from the fly is the single hair that points backward on each cell of the wing (Figure 22-30a). As we have seen, the fly is a model system that is particularly amenable to genetic dissection. Genetic analysis has shown that each wing cell responds to the planar direction of its neighbor, and components that specifically affect PCP have been identified (Figure 22-30b). The overall planar polarity of an epithelium is probably determined by a gradient of some ligand, such as a form of Wnt, or of mechanical tension across the tissue. This gradient polarizes all the cells in the epithelium in the same manner, causing proteins encoded by the *Frizzled* and *Dishevelled* genes to localize on one side of each cell and the proteins encoded by *Vang* and *Prickle* on the other (Figure 22-30c). Both Frizzled and Vang are transmembrane proteins, and they both associate — on opposite sides of the cell — with another membrane protein called *CELSR* (see Figure 22-30c). This asymmetric distribution of PCP proteins leads to the growth of the hair with the appropriate orientation. We have met Frizzled as a transmembrane receptor and Dishevelled as an adapter protein in the context of the Wnt pathway (see Figure 16-26), and their role in PCP may involve a form of Wnt and some other ligand. When components of the PCP pathway are disrupted — for example, in a *Dishevelled* mutant — the epithelium is intact, but the hairs are misoriented (see Figure 22-30b).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 22-30 Planar cell polarity (PCP) determines the orientation of cells. (a) The hairs on each cell of the fly wing all point in the same direction in a wild-type fly. (b) In a fly defective in PCP, as in this *Dishevelled* mutant, the orientation of the hairs

becomes disorganized, although the cells in the epithelium are still well organized. (c) The directionality of the hair is determined by the asymmetric localization of components of the PCP pathway, as indicated for Frizzled and Dishevelled, and Vang and Prickle, all of which are needed for orienting the hair appropriately. Planar cell polarity is propagated across a tissue due to two mechanisms. First, Frizzled associated with CELSR on one cell binds to CELSR associated with Vang on the adjacent cell. Second, within each cell, the distribution of Frizzled and Vang is mutually exclusive due to antagonism between their protein complexes. (d) The sensory hair cells of the vertebrate inner ear have V-shaped arrangements of stereocilia on their surface. In the adult and 18.5-day embryo (top and center images), all the cells are oriented in precisely the same way. In a mouse Celsr1 mutant defective in PCP, the cells in the 18.5-day embryo appear normal, but their relative orientations are disrupted (arrows in bottom panel). Scale bars = 2.5 μm.

[Parts (a) and (b) reprinted with permission from John Wiley & Sons, Inc., from J. D. Axelrod and C. J. Tomlin, 2011, "Modeling the Control of Planar Cell Polarity," *Wiley Interdiscip. Rev.: Syst. Biol. Med.* **3**(5):588–605; permission conveyed through Copyright Clearance Center, Inc. Part (d) republished with permission from The Company from Biologists Ltd., from M. Fanto and H. McNeill, 2004, "Planar Polarity from Flies to Vertebrates," *J. Cell Sci*. **117**(Pt4):527–533; permission conveyed through Copyright Clearance Center, Inc.]

The complementary arrangement of PCP components means that the membrane protein Vang on the side of one cell will be adjacent to the Frizzled protein on the adjacent cell. As indicated above, each protein associates with CELSR and it is believed that CELSR can adopt two distinct states, one with Frizzled and one with Vang. Like the polarity complexes in nematodes and flies, these protein complexes show intracellular mutual antagonism (see Figure 22-30c). Thus when Frizzled on one cell is next to Vang on the adjacent cell, that adjacent cell will enrich Frizzled on its opposite side, where it will associate with Vang on the next cell, and this pattern will repeat across the tissue. In this way, complementary

interactions between complexes containing Frizzled and Vang between cells and their mutual antagonism within a cell propagate PCP across a whole tissue.

Another clear example of PCP is the sensory hair cells of the inner ear that allow vertebrates to perceive sounds. Each of these cells has an ordered array of stereocilia (actin-based structures on sensory cells in the ear, which are required for hearing — see Figure 17-20b) arranged in a V-shaped pattern, and each cell is oriented precisely like its neighbor. In a mouse with a defect in the PCP gene *Celsr1*, the orderly arrangement of stereocilia within any given cell is preserved, but the relative orientations of cells to one another are disrupted (Figure 22-30d). These types of defects can result in deafness.

The Par Proteins Are Involved in Asymmetric Division of Stem Cells

We have seen that stem cells often give rise to a daughter stem cell and a differentiated daughter cell (see Figure 22-12). What are the cues that set up these asymmetric cell divisions? Two types of mechanisms have been found (Figure 22-31). In one mechanism, cell fate determinants are segregated to one end of the cell before cell division in response to external cues. This mechanism involves the Par proteins, which, as we have already seen, are instrumental

in the first asymmetric division of the nematode embryo and in establishing epithelial-cell polarity. In the second mechanism, the stem cell divides with a reproducible orientation so that it remains associated with the stem-cell niche, whereas the daughter cell is displaced away from the niche and can then differentiate. This is the situation we have already encountered in the *Drosophila* ovary, where the cap cells form a niche for the germ-line stem cells (see Figure 22-13).

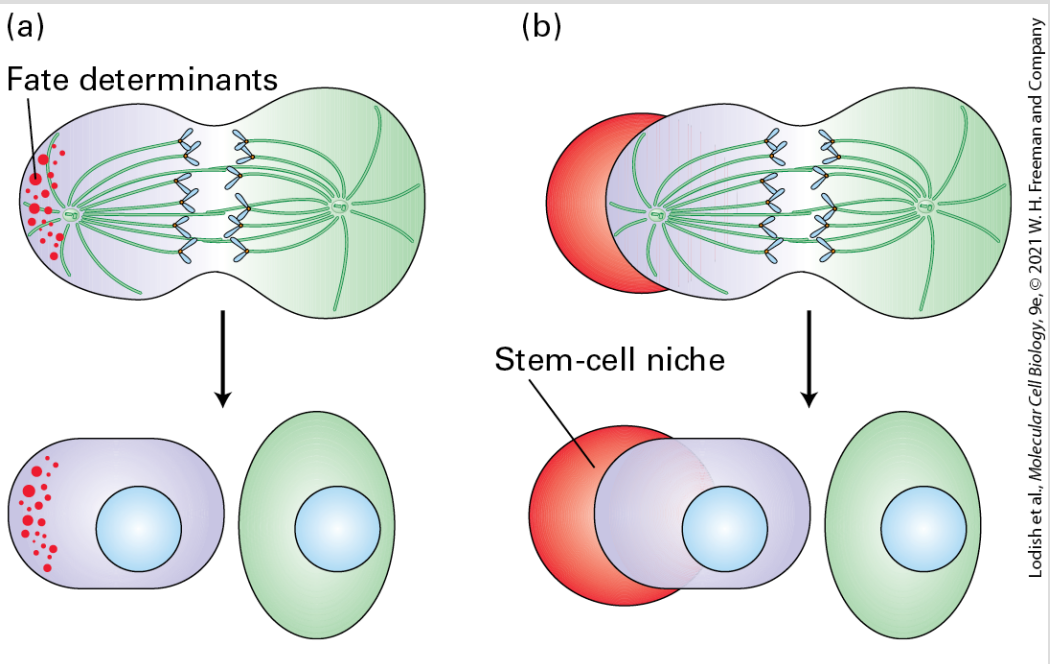


FIGURE 22-31 Two ways that stem cells can be induced to divide asymmetrically. (a) In response to an external cue, the cell polarizes and fate determinants become segregated before cell division. Division then produces one stem cell and one differentiating cell. (b) Stem cells interacting with a stem-cell niche orient their mitotic spindle to give rise to a stem cell associated with the niche and a differentiating cell displaced from it. See S. J. Morrison and J. Kimble, 2006, *Nature* **441**:1068.

A particularly well-understood example of the asymmetric division of stem cells is the formation of neurons and glial cells in the central nervous system of the fly (Figure 22-32). In this system, a neuroblast stem cell arises from the neurogenic ectoderm, which is a typical epithelial layer with apical and basal surfaces. The neuroblast enlarges (step **1**) and moves basally into the interior of the embryo but remains in contact with the neurogenic ectoderm epithelium (step **2**). It then divides asymmetrically (step **3**) to give rise to a new neuroblast and a ganglion mother cell (GMC) (step **4**). The GMC can divide only once, giving rise to two cells, either neurons or glial cells. The neuroblast, which remains a stem cell by maintaining an association with the neurogenic ectoderm niche, can divide repeatedly, giving rise to many ganglion mother cells and hence neurons and glial cells (step **5**) and thus populates the central nervous system. Thus the key event is the ability of the neuroblast to divide asymmetrically (Figure 22-32b). Once again, this process involves the asymmetric accumulation of the apical Par complex — Par3-Par6-aPKC — and its positioning at the side of the cell closest to the epithelium in an antagonistic relationship with Scribble. Par3 recruits factors that interact with the astral microtubules emanating from the spindle pole to orient the spindle — and hence cell division (Figure 22-32b). Other polarity-determining factors are then positioned at the basal side of the cell by an adaptor protein called *Miranda*, so that cell division segregates these factors (Figure 22-32c). Miranda associates with factors that

control proliferation and differentiation so they are segregated away from the neuroblast and into the GMC following cell division.

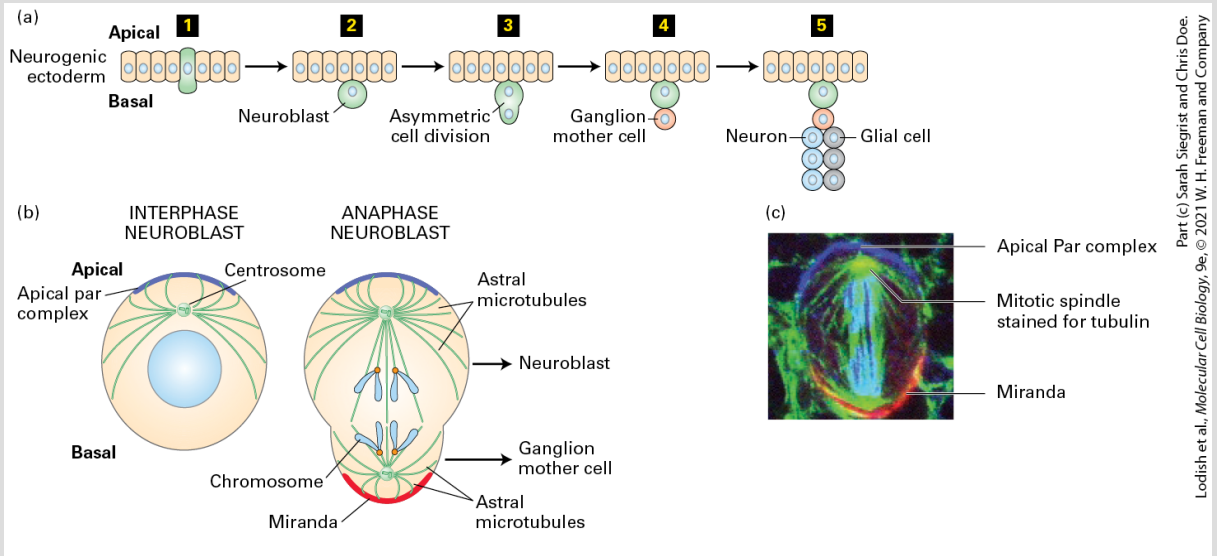


FIGURE 22-32 Neuroblasts divide asymmetrically to generate neurons and glial cells in the central nervous system. (a) Neuroblasts, which are stem cells, originate from the neurogenic ectoderm by means of signals that induce them to enlarge (step 1). They then move basally out of the ectoderm but remain in contact with it (step 2). Neuroblasts then undergo an asymmetric division (step 3) that produces a neuroblast and a ganglion mother cell (GMC) (step 4). The GMC then divides once to give two neurons or glial cells (step 5). Meanwhile, the neuroblast can divide many times to produce more GMCs, populating the neural tissue. (b) The asymmetric division of the neuroblast requires the correct orientation of the mitotic spindle, mediated by the apical Par complex and distal Miranda complex, to give rise to a larger neuroblast and a smaller GMC. (c) A neuroblast at anaphase, showing the segregation of the apical Par proteins (blue) and the basal Miranda protein (red).

The first three sections of this chapter illustrated that all metazoan organisms have evolved complex genetic proteins to ensure that, during development, the correct number of each type of differentiated cells is formed and that these cells are positioned and

polarized appropriately with respect to their neighbors. Perhaps surprisingly, many types of cells are born only to quickly die. During nematode development (see Figure 22-26a), about 15 percent of the cells that are formed are programmed to die quickly. Formation of different types of antibodies against pathogens is an essential function of the B cells of the immune system (Chapter 24), and B cells that fail to produce an antibody that targets a foreign protein also are programmed to die. In the last section of this chapter, we discuss evolutionarily conserved genetic programs that regulate cell death, and learn how these are also absolutely crucial for the formation and maintenance of many body tissues (see Figure 22-1c).

KEY CONCEPTS OF SECTION 22.3

Mechanisms of Cell Polarity and Asymmetric Cell Division

- Cell polarity involves the asymmetric distribution of proteins, lipids, and other macromolecules in the cell.
- Cells have an intrinsic program that can generate polarity using feedback loops.
- A key regulator of the polarity program in many systems is the small GTP-binding protein Cdc42.
- When a yeast cell buds, the intrinsic polarity program exploits feedback loops to concentrate Cdc42·GTP at a single site.
- Asymmetry requires cells to sense a cue, respond to it by assembling a polarized cytoskeleton, and then use this polarity to distribute polarity factors appropriately.
- Mating in haploid yeast involves assembly of a mating projection (shmoo) by polarization of the cytoskeleton in the direction of highest concentration of mating pheromone and targeting of secretion of cellular components for cell expansion there.
- Anterior/posterior asymmetry in the first division of the *C. elegans* embryo involves asymmetric contraction of the actin-myosin network to localize the anterior Par3-Par6-aPKC complex to the anterior cortex, followed by the association of posterior

factors such as Par2 with the posterior cortex. The asymmetry of the anterior and posterior complexes is maintained by mutually antagonistic pathways.

- Apical/basal epithelial-cell polarity is also driven by an apical Par3-Par6-aPKC complex, which functions in antagonistic relationships with the apical Crumbs complex and the basal Scribble complex.
- Planar cell polarity (PCP) regulates the orientation of epithelial cells across a sheet using a different set of antagonistic relationships.
- Asymmetric cell division requires that cells first become polarized, then divide to segregate fate determinants asymmetrically.
- Asymmetric division of stem cells often involves association of the stem cell with a niche, in which case the stem cell gives rise to another stem cell and a differentiating cell.
- Asymmetric stem-cell division also involves the asymmetric distribution of the Par complex, which is retained in the stem cell during division, whereas fate determinants are localized away from the Par complex to end up in the differentiating cell.

22.4 Cell Death and Its Regulation

Regulated cell death is an essential process in metazoan organisms for the removal of cells that are harmful or no longer needed. Our focus in this section will be on one principal type of programmed cell death, termed **apoptosis** — a common molecular pathway that largely is conserved in invertebrates and vertebrates. During embryogenesis, programmed apoptotic death of specific cells keeps chicken feet as well as our hands from being webbed (Figure 22-33); it also prevents our embryonic tails from persisting and our brains from being filled with unneeded nerve connections. In fact, up to 50 percent of the nerve cells generated during brain development subsequently die of apoptosis because they do not develop appropriate connections with other cells. Many kinds of muscle, epithelial, and white blood cells constantly wear out and need to be removed by apoptosis and replaced. We also discuss necroptosis, one of several types of programmed cell death that vertebrate cells employ as a response to infection or other stresses. We will see that necroptosis can lead to an immune response, which in turn can either lead to amelioration of the infection or cause inflammation and extensive tissue damage.

Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 22-33 A web-footed chicken. During the development of many vertebrate limbs, cells in the soft tissue between the embryonic digits undergo programmed cell death. In the chicken foot, this process leads to the formation of four separate toes (*left*). During chicken foot development, bone morphogenetic proteins (BMPs) (members of the TGF-β superfamily of hormones; see Figure 16-22a) are expressed by interdigital cells and induce apoptosis. In this experiment, a dominant-negative type I BMP receptor was expressed in a developing chicken foot, blocking BMP signaling and preventing the programmed cell death that normally occurs. This manipulation allowed the survival of cells that then divided and differentiated into a web (*right*). The similarity of this webbing to webbed duck feet led to studies showing that BMPs are not expressed in duck interdigital cells. These results indicate that BMP signaling actively mediates cell death in the embryonic limb.

[Republished with permission from the AAAS, from H. Zou and L. Niswander, 1996, "Requirement for BMP Signaling in Interdigital Apoptosis and Scale Formation," *Science* **272**(5262):738–741; permission conveyed through Copyright Clearance Center, Inc.]

Cell-cell interactions regulate cell death in two fundamentally different ways. First, most, if not all, cells in multicellular organisms require specific protein hormone signals to stay alive. In the absence of such survival signals, frequently referred to as **trophic factors**, cells activate a *suicide program*. Second, in some developmental contexts including the immune system, and in adult life, other specific hormone signals (often on the surface of a cell) can induce the activation of a suicide program in another cell and thus *murder* the target cell. Whether cells commit suicide for lack of survival signals or are murdered by killing signals from other cells, cell death is most often mediated by apoptosis. The cell corpses are ingested by neighboring cells, and their contents are broken down into small molecules and reused to build other cells. Because there is no rupture of the plasma membrane, contents of the cell do not leak into the surrounding medium, which otherwise would lead to an inflammatory immune response.

A different form of cell death, **necrosis**, occurs when cells are subjected to injury or excessive stresses, such as heat, absence of oxygen, or infection by pathogens. Necrosis involves swelling of the cell followed by formation of holes in the plasma membrane, causing leakage of intracellular contents, which in turn attracts white blood cells, including macrophages and other phagocytic cells. In turn this leads to an inflammatory response — heat, pain, swelling — whose function is to clear out the necrotic cells and

initiate repair of the tissue. Perhaps surprisingly, necrosis can be the consequence of activating another mechanism of cellular suicide, termed **necroptosis**, which is only found in higher vertebrates. Necroptosis is often triggered by extracellular cytokines such as tumor necrosis factor alpha (TNFα). Activation of this cell-death pathway frequently causes inflammation and contributes to the development of many human diseases, including nerve degeneration and atherosclerosis.

Excess cells may also be extruded from epithelial layers, such as in gut and skin, by a mechanically regulated process that can cause detachment and the loss of trophic survival signals; extruded cells then die by yet another pathway termed *anoikis*. This maintains the epithelial barrier function as cells die and matches the rates of cell proliferation and death.

In this section, we first distinguish programmed cell death from passive cell death and then describe how genetic studies in the nematode worm *C. elegans* led to the elucidation of an evolutionarily conserved effector pathway that leads to apoptosis. We then turn to vertebrates, in which cell death is actively suppressed by trophic factors, as exemplified by their importance in programmed cell death in neuronal development, and which may be activated by pro-inflammatory cytokines and cell stresses such as DNA damage in disease conditions. We illustrate the key roles of mitochondria in initiating vertebrate cell-death pathways and describe how apoptotic

cells generate an “eat me” signal on their surface membranes. Finally, we discuss necrosis and necroptosis and how our understanding of these processes has paved the way for treating certain human diseases.

Most Programmed Cell Death Occurs Through Apoptosis

The demise of cells by programmed cell death is marked by a well-defined sequence of morphological changes, collectively referred to as *apoptosis*, a Greek word that means “dropping off” or “falling off,” like leaves from a tree. Dying cells shrink, condense, and then fragment, releasing small membrane-bound apoptotic bodies, which are then engulfed by other cells (Figure 22-34). Within these apoptotic cells, mitochondria undergo fission and lose their membrane potential, nuclei condense, and the DNA is fragmented by nuclease cleavage. Importantly, the intracellular constituents are not released into the extracellular milieu, where they could have deleterious effects by eliciting immune responses, but instead are phagocytosed by neighboring cells. The stereotypical changes that occur during apoptosis, such as condensation of the nucleus and phagocytosis by surrounding cells, suggested to early scientists that this type of cell death was under the control of a strict gene expression program. This program is critical during both embryonic and adult life to maintain normal cell number and composition.

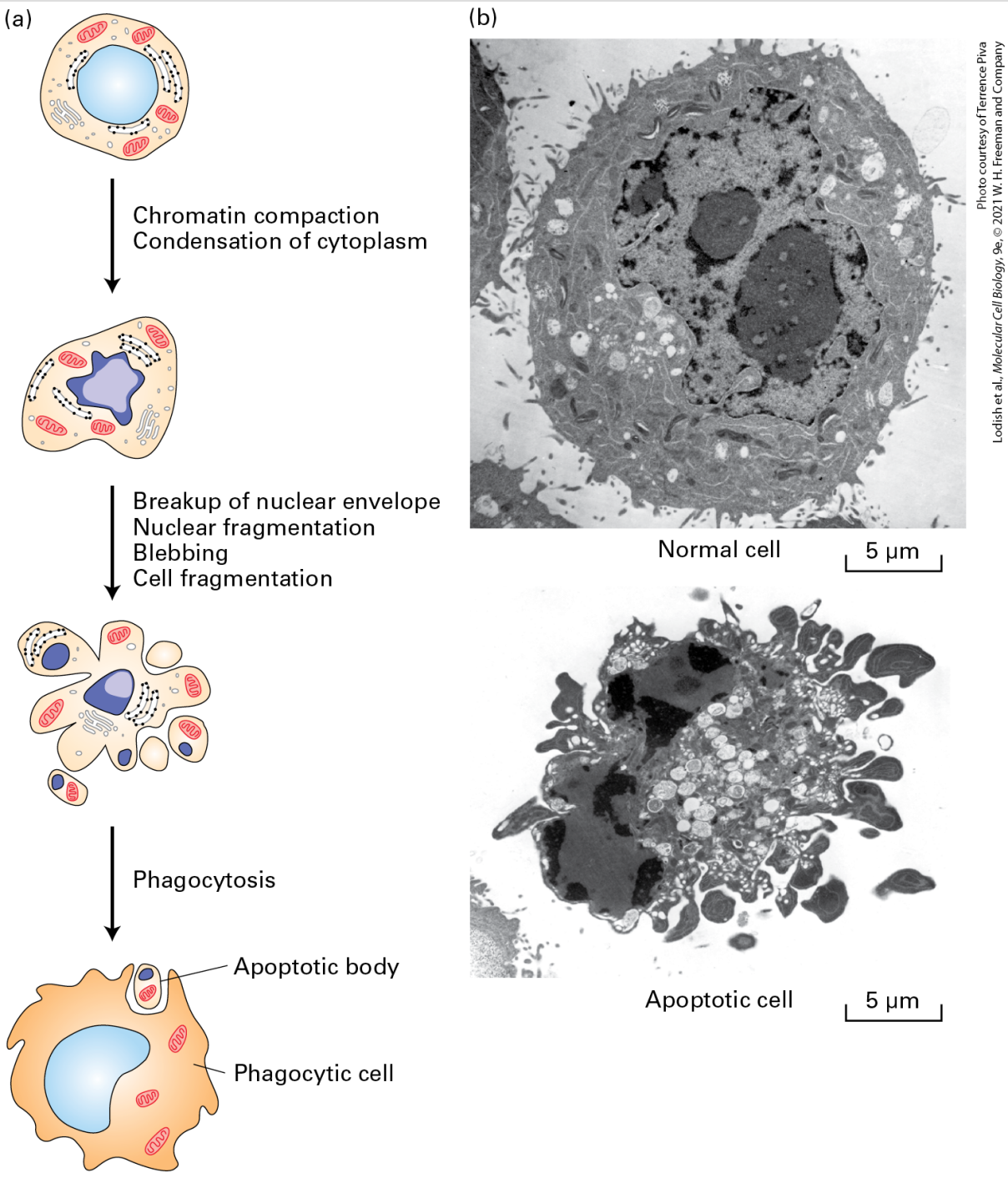


FIGURE 22-34 Ultrastructural features of cell death by apoptosis. (a) Schematic drawings illustrating the progression of morphological changes observed in apoptotic cells. Early in apoptosis, dense chromosome condensation occurs along the nuclear periphery. The cell body also shrinks, although most organelles remain intact. Later, both the nucleus and the cytoplasm fragment, forming membrane-enveloped apoptotic bodies, which are

phagocytosed by surrounding cells without leakage of the cellular contents into the surrounding medium. (b) Photomicrographs comparing a normal HeLa cell and an apoptotic cell. Clearly visible in the latter are dense spheres of compacted chromatin as the nucleus begins to fragment.

[Part (b) T. J. Piva et al., 2012, "Increased Activity of Cell Surface Peptidases in HeLa Cells Undergoing UV-Induced Apoptosis Is Not Mediated by Caspase 3," *Int. J. Mol. Sci*. **13**(3):2650–2675; photo courtesy of Terrence Piva.]

The genes involved in controlling cell death encode proteins with three distinct functions:

- *Killer* or *initiation* proteins are required for a cell to begin the apoptotic process.
- *Destruction* proteins perform functions such as digesting proteins and DNA in a dying cell.
- *Engulfment* proteins are required for phagocytosis of the dying cell by another cell.

At first glance, engulfment seems to be simply an after-death cleanup process, but evidence indicates that it is part of the final death process. For example, mutations in killer genes always prevent cells from initiating apoptosis, whereas mutations that block engulfment genes sometimes allow cells to persist for a while before dying. Engulfment involves the assembly of a halo of actin in the engulfing cell around the dying cell or cell fragment, triggered by activation of Rac, a monomeric G protein that helps regulate actin polymerization (see Figure 17-44).

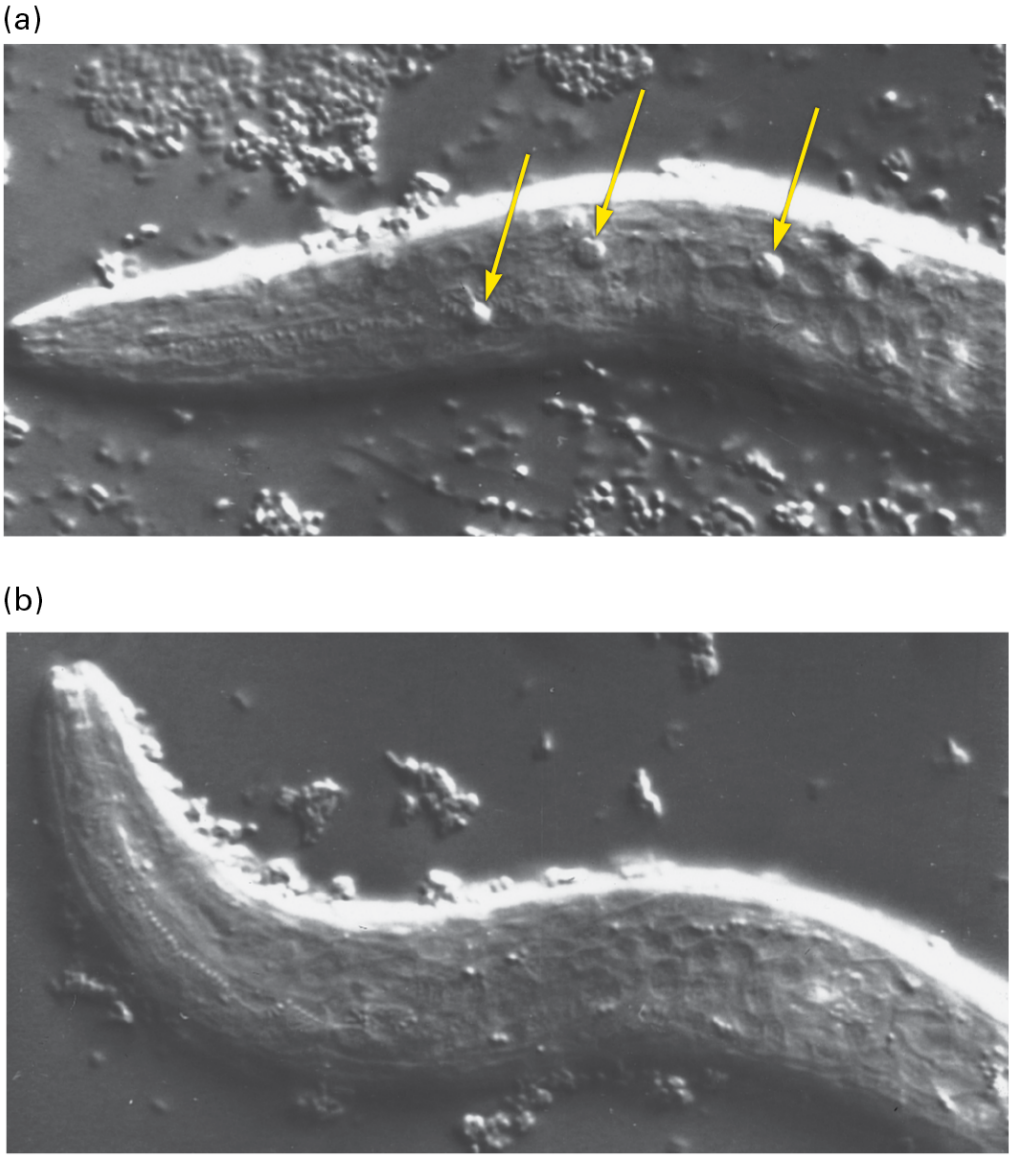
In contrast to apoptosis, cells that die by necrosis or necroptosis exhibit very different morphological changes. Cells that undergo this process swell and burst, releasing their intracellular contents, which can damage surrounding cells and frequently cause inflammation — triggering the clearance of necrotic cells and tissues damaged from the original insult and the initiation of tissue repair.

Evolutionarily Conserved Proteins Participate in the Apoptotic Pathway

The confluence of genetic studies in *C. elegans* and studies on human cancer cells suggested that an evolutionarily conserved pathway mediates apoptosis. In *C. elegans*, cell lineages are under tight genetic control and are identical in all individuals of the species. About 10 rounds of cell division or fewer create the adult worm, which is about 1 mm long and 70 μm in diameter (see Figure 22-26c). The adult worm may be a hermaphrodite (having both male and female organs) or a male. The hermaphrodite form has 959 somatic-cell nuclei, whereas the male has 1031. Scientists have traced the lineage of each somatic cell in *C. elegans* from the fertilized egg to the mature worm by following the development of live worms using DIC microscopy (see Figure 22-26d).

Of the 947 nongonadal cells generated during development of the adult hermaphrodite, 131 undergo programmed cell death. Specific

mutations revealed four genes whose encoded proteins play an essential role in controlling programmed cell death during *C. elegans* development: *ced-3, ced-4, ced-9*, and *egl-1*. In *ced-3* or *ced-4* mutants, for example, the 131 doomed cells survive (Figure 22-35). These mutants provided the first pieces of evidence that apoptosis was under a genetic program and led to the 2002 Nobel Prize for H. Robert Horvitz. The mammalian proteins that correspond most closely to the worm CED-3, CED-4, CED-9, and EGL-1 proteins are indicated in Figure 22-36. In discussing the worm proteins, we will occasionally include the mammalian names in parentheses to make it easier to keep the relationships clear.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

EXPERIMENTAL FIGURE 22-35 Mutations in the *ced-3* gene block programmed cell death in *C. elegans*. (a) Newly hatched mutant larva carrying a mutation in the *ced-1* gene. Because mutations in this gene prevent engulfment of dead cells, highly refractile (and thus easily visualized) dead cells accumulate (arrows). (b) Newly hatched larva with mutations in both the *ced-1* and *ced-3* genes. The absence of refractile dead cells in these double mutants indicates that no cell deaths occurred. Thus the CED-3 protein is required for programmed cell death.

[Republished with permission from Elsevier, from H. M. Ellis and H. R. Horvitz, 1986, "Genetic Control of Programmed Cell Death in the Nematode *C. elegans*,"

Cell **44**(6):817–829; permission conveyed through Copyright Clearance Center, Inc.]

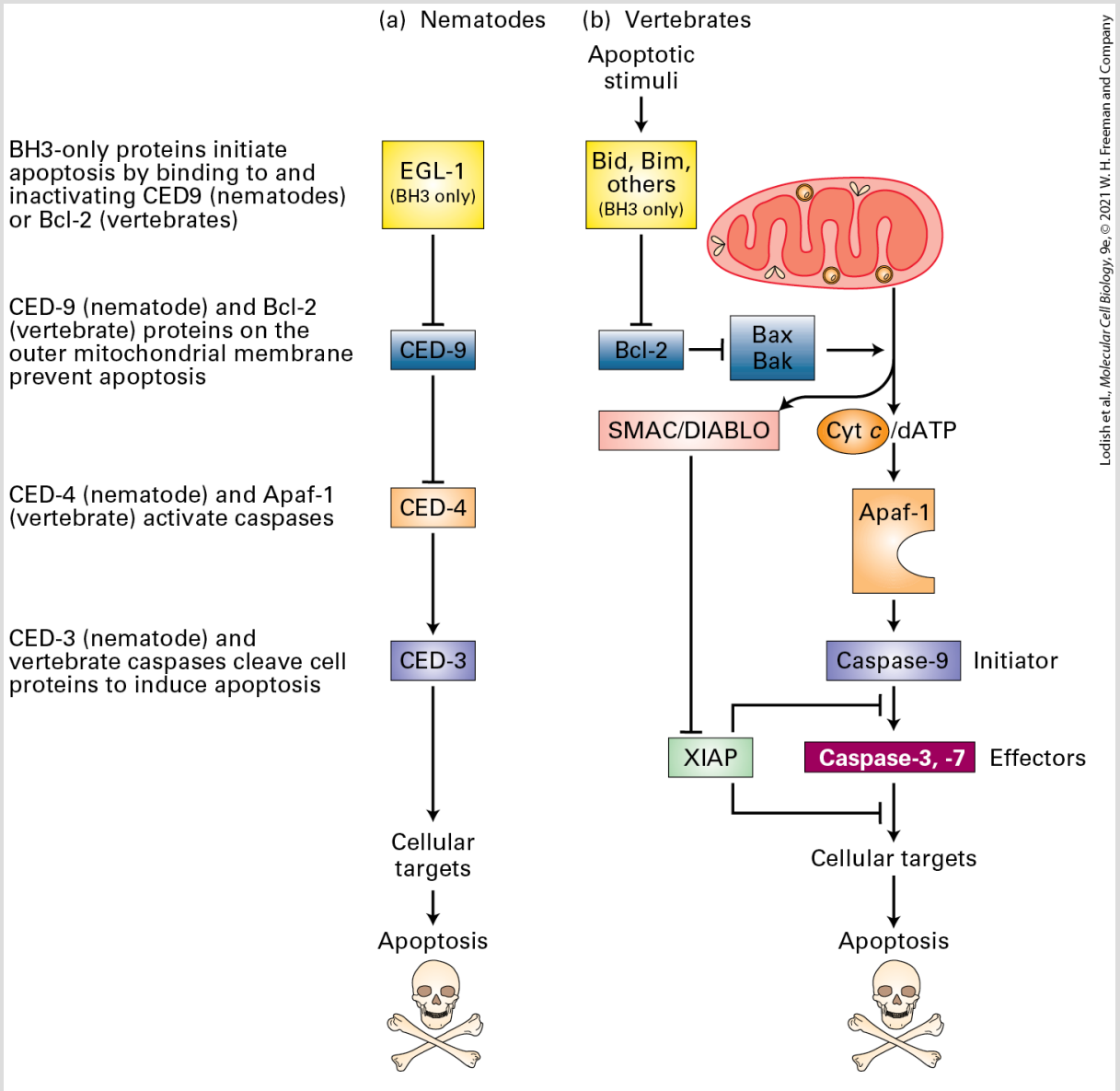


FIGURE 22-36 Evolutionary conservation of apoptosis pathways. Similar proteins, shown in identical colors, play corresponding roles in nematodes and in vertebrates. (a) In nematodes, the BH3-only protein called *EGL-1* binds to CED-9 on the outer mitochondrial membrane; this interaction releases CED-4 from the CED-4/CED-9 complex. Free CED-4 then binds to and activates by autoproteolytic cleavage, the caspase CED-3, which cleaves cell proteins to drive apoptosis. These relationships are shown as a genetic pathway, with EGL-1

inhibiting CED-9, which in turn inhibits CED-4. Active CED-4 activates CED-3. (b) In vertebrates, homologs of the nematode proteins, as well as many other proteins not found in the nematode, regulate apoptosis. The Bcl-2 protein is similar to CED-9 in promoting cell survival in part by preventing activation of Apaf-1, which is similar to CED-4. Bcl-2 also prevents apoptosis by other mechanisms depicted in Figures 22-42 and 22-44. Several types of BH3-only proteins, detailed in Figures 22-41 and 22-42, inhibit Bcl-2 and thus allow apoptosis to proceed. Many apoptotic stimuli lead to damage of the outer mitochondrial membrane, causing release into the cytosol of several proteins that stimulate apoptosis. In particular, cytochrome *c* released from mitochondria activates Apaf-1, which in turn activates caspase-9, the homolog of the nematode CED-3. This initiator caspase then cleaves and thus activates effector caspase-3 and caspase-7, eventually leading to apoptosis. See text for discussion of other mammalian proteins (SMAC/DIABLO and XIAP) that have no nematode homologs.

[Information from S. J. Riedl and Y. Shi, 2004, *Nat. Rev. Mol. Cell Biol.* **5**:897.]

The first mammalian apoptotic gene to be cloned, *bcl-2*, was isolated as an oncogene from human follicular lymphomas, tumors of the antibody-producing B cells of the immune system. A mutant form of this gene was formed in a patient's lymphoma cells; a chromosomal rearrangement joined the protein-coding region of the *bcl-2* gene to an immunoglobulin-gene enhancer. The combination results in overproduction of the Bcl-2 protein, which keeps these cancer cells alive when they would otherwise become programmed to die by apoptosis. The human Bcl-2 protein and worm CED-9 protein are 23 percent identical in amino acid sequence, and the expression of a *bcl-2* transgene can block the extensive cell death observed in *ced-9* mutant worms. Thus both proteins act as regulators that suppress the apoptotic pathway (see Figure 22-36). In addition, both proteins contain a single transmembrane domain and are localized mainly to

the outer mitochondrial membrane, where they serve as sensors that control the apoptotic pathway in response to external stimuli. As we discuss next, other regulators promote apoptosis.

In the worm apoptotic pathway, CED-3 (caspase-9 in mammals) is a protease required to cleave key cell proteins during apoptosis. CED-4 (Apaf-1) is a protease-activating factor that causes autoproteolytic cleavage of the CED-3 protease to promote its activation and to initiate cell death, as CED-3 is synthesized by and present in all cells as an inactive precursor (a zymogen) (see Figures 22-36 and 22-37). Cell death does not occur in *ced-3* and *ced-4* single loss-of-function mutants or in *ced-9/ced-3* double mutants. In contrast, in *ced-9* loss-of-function mutants, all cells die by apoptosis during embryonic life, so the adult form never develops. These genetic studies indicate that CED-3 and CED-4 are killer proteins required for cell death, and that CED-9 (Bcl-2) suppresses apoptosis. The observation that all cells die in *ced-9* loss-of-function mutants shows that the apoptotic pathway is present in and can be activated in all cells. Moreover, the absence of cell death in *ced-9/ced-3* double mutants suggests that CED-9 acts upstream of CED-3 to suppress the apoptotic pathway.

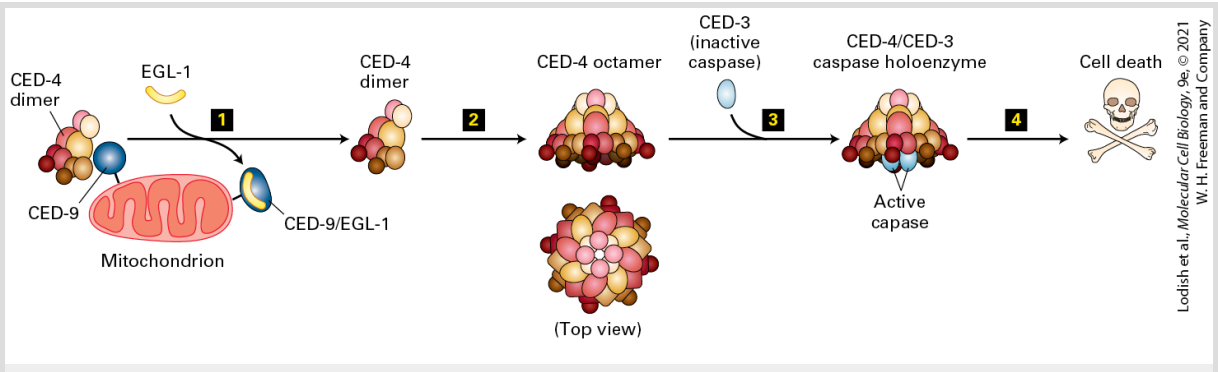


FIGURE 22-37 Activation of CED-3 protease in *C. elegans*. EGL-1, a BH3-only protein that is produced in response to developmental signals that trigger cell death, displaces a CED-4 dimer from its association with CED-9 (step **1**). Four freed CED-4 dimers combine to form a CED-4 octamer (step **2**), which binds two molecules of the CED-3 zymogen (an enzymatically inactive precursor of a caspase protease). Bringing the two CED-3 zymogens together triggers the self-cleavage of the CED-3 zymogens into active CED-3 proteases (step **3**). This effector caspase then begins to cleave multiple cell proteins, leading to cell death (step **4**). See N. Yan et al., 2005, *Nature* **437**:831; and S. Qi et al., 2010, *Cell* 141:446.

The mechanism by which CED-9 (Bcl-2) controls CED-3 (caspase-9) in the nematode is known and is somewhat different from the mechanism, discussed later (see Figure 22-43), in mammalian cells. In all nematode cells the CED-9 protein, localized to the outer mitochondrial membrane, forms a complex with a dimer of the CED-4 (Apaf-1) protein, thereby preventing the CED-4 protein from activating the CED-3 protease (Figure 22-37). As a result, the cell survives. This mechanism fits with the genetics, which shows that the absence of CED-9 causes all cells in the embryo to die, but that the absence of CED-9 has no effect if CED-3 is also missing (*ced-3/ced-9* double mutants have no cell death). The three-dimensional structure of the trimeric CED-4/CED-9 complex reveals a huge contact surface between each of the two CED-4 molecules and the single CED-9 molecule; the large contact surface makes the association highly specific, but in such a way that the dissociation can be regulated.

Transcription of *egl-1*, the fourth genetically defined apoptosis regulator gene, occurs only in *C. elegans* cells that are programmed

to die. Newly produced EGL-1 protein binds to CED-9, alters its conformation, and catalyzes the release of CED-4 from CED-9 (see Figure 22-37). Both EGL-1 and CED-9 contain a 12-amino-acid sequence termed the *BH3 domain*. Because EGL-1 lacks most of the other domains of CED-9, it is called a **BH3-only protein**. The mammalian BH3-only proteins closest in sequence and function to EGL-1 are the pro-apoptotic proteins Bid and Bim, discussed later. Once EGL-1 causes dissociation of the CED-4/CED-9 complex, four released CED-4 dimers join to make an octamer, which then binds to and activates the CED-3 protease by a mechanism we will discuss shortly. Cell death soon follows (see Figure 22-37).

Evidence that the steps described here are sufficient to induce apoptosis comes from experiments in which these steps were reconstituted in vitro with purified proteins. CED-3, CED-4, a segment of the CED-9 protein that lacked its mitochondrial membrane anchor, and EGL-1 were purified, as was a CED-4/CED-9 complex. Purified CED-4 (Apaf-1) was able to bind to and accelerate the autoproteolytic cleavage and activation of purified CED-3 (caspase-9) protease, but addition of CED-9 (Bcl-2) to the reaction mixture inhibited the autocleavage. When the CED-4/CED-9 complex was mixed with CED-3, autocleavage did not occur, but addition of EGL-1 to the reaction mixture restored CED-3 autocleavage by releasing CED-4 from its complex with CED-9.

To see the importance of regulated EGL-1 expression in apoptosis, consider a class of neurons in *C. elegans* found in hermaphrodites, but not in males. These hermaphrodite-specific neurons are generated embryonically in both hermaphrodites and males but undergo programmed cell death in males. In hermaphrodites, expression of the *egl-1* gene in these neurons is repressed by the transcription factor TRA-1A, and deletion of TRA-1A in hermaphrodites causes these neurons to undergo apoptosis. This finding reinforces a point made earlier: all metazoan cells can potentially undergo apoptosis, so this process needs to be carefully regulated!

Caspases Amplify the Initial Apoptotic Signal and Destroy Key Cellular Proteins

The key enzymes that mediate apoptosis, the **caspases**, are so named because they contain a key *c*ysteine residue in the catalytic site and selectively cleave proteins at sites just C-terminal to *asp*artate residues. Caspases function as homodimers, with one domain of each stabilizing the active site of the other; they mediate apoptosis by proteolytic cleavage of selected substrates that contain their preferred cleavage sites. Their specific intracellular targets include proteins that form the nuclear lamina and cytoskeleton, whose cleavage leads to the demise of a cell. In vertebrates, targets

include other pro-apoptotic proteins such as Bid, which in turn promote mitochondrial damage. While apoptosis in *C. elegans* is executed by a single caspase CED-3, humans have 14 different caspases. In mammalian cells, apoptosis is mediated by a cascade of caspases which sequentially amplifies the initial death signal. All caspases are initially made as *procaspases* that require either binding to specific protein complexes or a proteolytic cleavage to become active.

Apoptosis of mammalian cells is activated by either intrinsic or extrinsic signaling pathways. The intrinsic apoptosis pathway is mediated by an activated *initiator caspase*, such as caspase-9, which in turn mediates the activation of the effector proteases, such as caspase-3 and caspase-7; in this way, the proteolytic activity of a few activated initiator caspases becomes rapidly and hugely amplified, leading to a massive increase in the total caspase activity level in the cell (see Figure 22-36) and consequently cell death. The extrinsic apoptosis pathway, discussed later, is initiated by ligands of death receptor family, such as TNFα (see Figure 22-44) and FasL by binding to their cognate receptors. This leads to the activation of caspase-8, an initiator caspase, which in turn activates the effector caspases, such as caspase-3 and caspase-7, to execute cell death.

Phosphatidylserine: an “Eat Me” Signal on the Surface of Apoptotic

Cells

As we learned in Section 1 of Chapter 10, the phospholipid phosphatidylserine is normally found in the inner, cytosolic leaflet of the plasma membrane. During apoptosis, increasing amounts of phosphatidylserine are found in the exoplasmic leaflet, where it acts as an "eat me" signal: it binds to several types of phosphatidylserine receptors on the surface of neighboring phagocytic cells that initiate engulfment of the cell or cell fragment. Experimentally, the presence of phosphatidylserine on the surface of a cell is indicative that it is undergoing apoptosis and can be detected by the binding of the protein annexin V (which is visualized by conjugation with a fluorescent molecule). While the normal cellular function of annexin V is obscure, it does bind tightly and specifically to phosphatidylserine.

In normal vertebrate cells, ATP-powered pumps called *ATP11A* and *ATP11C*, members of the ABC class of transmembrane proteins (see Figure 11-16), continuously transport phosphatidylserine from the outer to the inner leaflet of the plasma membrane. The phospholipid scramblases XKR8 and TMEM16F, which otherwise would translocate some of the phosphatidylserine back to the outer membrane leaflet, are inactive (Figure 22-38). During apoptosis, caspase-3 cleaves both ATP11A and ATP11C, inactivating them. Additionally, caspase-3 cleaves XKR8, and in so doing activates its function: the truncated XKR8 binds to two other proteins, dimerizes,

and becomes an active translocase, shuttling all phospholipids, including phosphatidylserine, from one leaflet of the plasma membrane to the other. The end result is appearance of the macrophage, “eat me” signal phosphatidylserine on the outer surface of the plasma membrane, where it binds to a set of receptors on the surface of macrophages and other immune system cells, triggering phagocytosis of the cell or apoptotic body (see Figure 22-34).

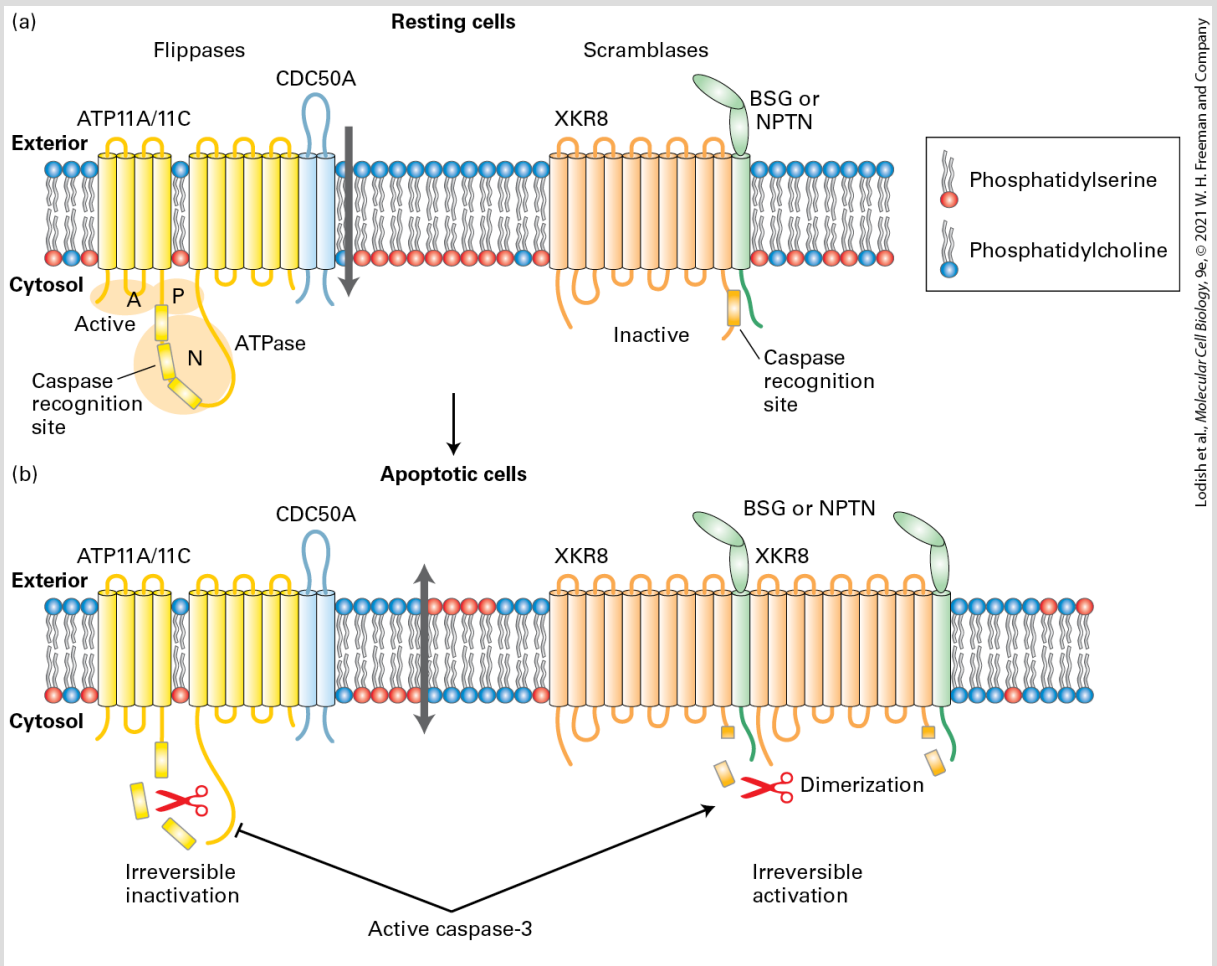


FIGURE 22-38 Movement of phosphatidylserine to the outer plasma membrane leaflet during apoptosis. (a) In normal vertebrate cells, ATP-powered phospholipid flippases ATP11A and ATP11C continuously translocate phosphatidylserine from the outer to the

inner leaflet of the plasma membrane (unidirectional gray arrow), and the phospholipid scramblase XKR8 which otherwise would equilibrate phosphatidylserine in both plasma membrane leaflets, is inactive. (b) In cells undergoing apoptosis, caspase-3 cleavage irreversibly inactivates ATP11A and ATP11C. In addition, caspase-3 cleaves the inactive phospholipid scramblase XKR8. The truncated XKR8, together with either basigin (BSG) or neuroplastin (NPTN), undergoes dimerization and activation, and translocates phospholipids in both directions. The end result is appearance of the "macrophage, eat me" signal phosphatidylserine on the outer surface of the plasma membrane.

[Information from S. Cory, 2018, *Proc. Nat'l. Acad. Sci. USA* **115**:12092–12094.]

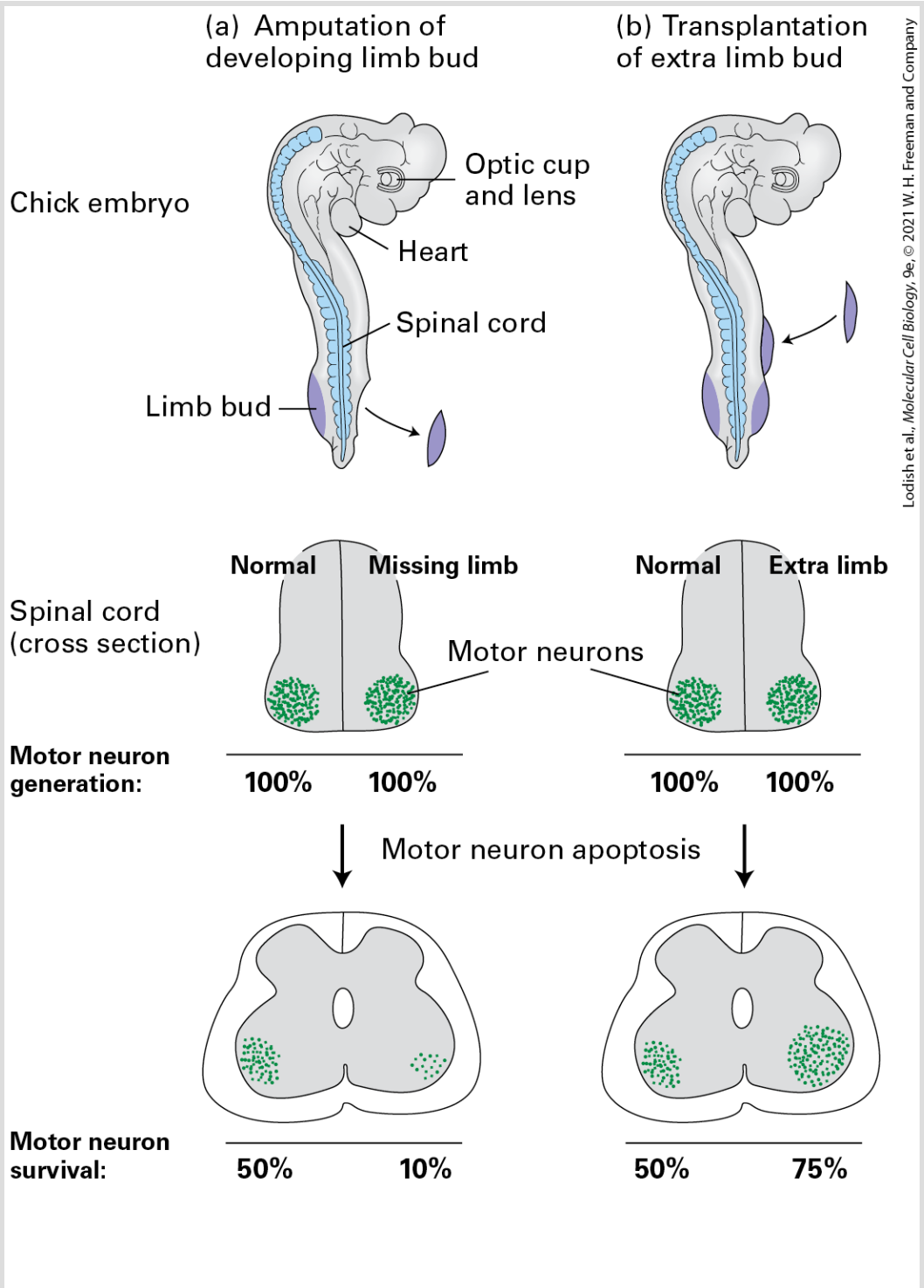
Neurotrophins Promote Survival of Neurons

We noted that, in the nematode *C. elegans*, all cell lineages are under tight genetic control and are identical in all individuals of the species; during development of the adult hermaphrodite form, precisely 131 cells undergo programmed cell death. During development of vertebrates, too, many cells are born only to die, but apoptosis of most vertebrate cells is not under programmed genetic control. Rather, apoptosis is regulated by intracellular signals generated from many secreted and cell-surface protein hormones, as well as by many environmental stresses, such as ultraviolet irradiation and DNA damage. While the core apoptosis machinery in *C. elegans* is conserved in vertebrates, many other intracellular proteins also regulate apoptosis (see Figure 22-36, *right*).

Before plunging into these molecular details, we'll illustrate the importance of trophic factors in apoptosis with a brief analysis of the developing nervous system. When neurons grow to make connections to other neurons or to muscles, sometimes over considerable distances, more neurons grow than will eventually survive. As an example, the cell bodies of many sensory and motor neurons are located in the spinal cord and adjacent ganglia, while their axons grow outward from the cell body toward their target cells — often muscle cells that they will innervate (see Figure 18-55). Those neurons that make signaling connections, termed *synapses* (see Figure 23-3), with their intended target cells prevail and survive; those that fail to connect will die.

In the early 1900s, the number of neurons innervating peripheral cells was shown to depend on the size of the tissue to which they connect, the so-called *target field*. For instance, removal of limb buds from a developing chick embryo leads to a reduction in the number of both sensory and motor neurons innervating muscles in the bud (Figure 22-39). Conversely, grafting additional limb tissue to a limb bud leads to an increased number of neurons in the corresponding regions of the spinal cord and sensory ganglia. Indeed, incremental increases in target-field size are accompanied by commensurate incremental increases in the number of neurons innervating the target field. This relationship was found to result from the selective survival of neurons, rather than changes in their differentiation or proliferation. The observation that many sensory and motor

neurons die after reaching their peripheral target field suggested that these neurons compete for survival factors produced by the target tissue.



EXPERIMENTAL FIGURE 22-39 In vertebrates, the survival of motor neurons depends on the size of the muscle target field they innervate. (a) Removal of a limb bud from one side of a chick embryo at about 2.5 days of development results in a marked decrease in the number of motor neurons on the affected side. In an amputated embryo (*top*), normal numbers of motor neurons are generated on both sides of the spinal cord (*middle*) and grow outward to innervate limb muscles. Later in development, however, many fewer motor neurons remain on the side of the spinal cord with the missing limb than on the normal side (*bottom*). Note that normally only about 50 percent of the motor neurons that are generated survive. (b) Transplantation of an extra limb bud into an early chick embryo produces the opposite effect, more motor neurons on the side with additional target tissue than on the normal side. See D. Purves, 1988, *Body and Brain: A Trophic Theory of Neural Connections*, Harvard University Press.

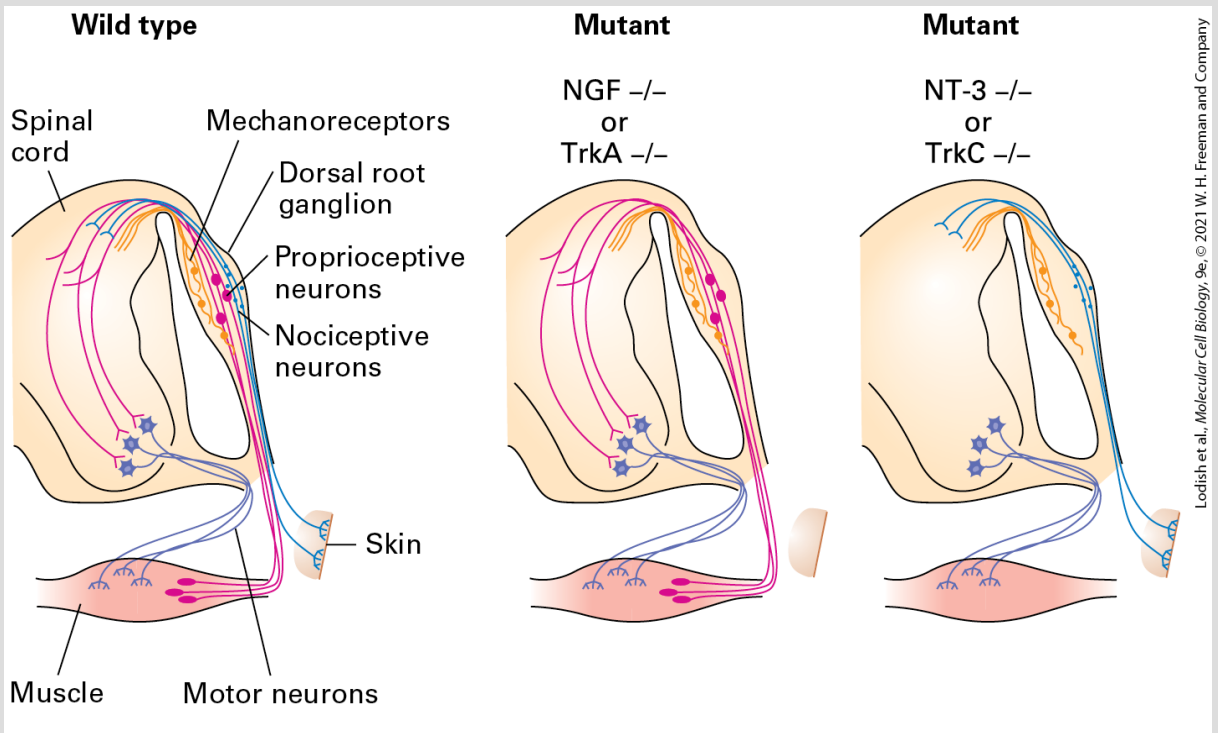
[Data from E. R. Kandel, J. H. Schwartz, and T. M. Jessell, 2000, *Principles of Neural Science*, 4th ed., McGraw-Hill, p. 1054, Fig. 53-11.]

Subsequent to these early observations, scientists discovered that transplantation of a mouse sarcoma (muscle tumor) into a chick led to a marked increase in the local numbers of certain types of neurons. This finding implicated the tumor as a rich source of a presumed trophic factor. To isolate and purify this factor, known simply as *nerve growth factor* (NGF), scientists used cell culture assays in which outgrowth of neurites from sensory ganglia was measured. *Neurites* are extensions of the neuronal cytoplasm that can grow to become the long processes of the nervous system, the axons and dendrites (see Figure 23-1). The later discovery that the submaxillary gland of the mouse also produces large quantities of NGF enabled Rita Levi-Montalcini to purify and sequence it; she was

rewarded with a Nobel Prize. A homodimer of two 118-residue polypeptides, NGF belongs to a family of structurally and functionally related trophic factors collectively referred to as **neurotrophins**. Brain-derived neurotrophic factor (BDNF) and Neurotrophin-3 (NT-3) are also members of this protein family.

Neurotrophins bind to and activate a family of receptor tyrosine kinases called *Trks* (pronounced "tracks"). (The general structure of receptor tyrosine kinases and the intracellular signaling pathways they activate are covered in Chapter 16.) Each neurotrophin binds with high affinity to one type of Trk receptor: NGF binds to TrkA; BDNF to TrkB; and NT-3 to TrkC. NT-3 can also bind with lower affinity to both TrkA and TrkB. All neurotrophins also bind to a distinct type of receptor called $p75^{NTR}$ (also called *NTR 5 neurotrophin receptor*), but with lower affinity. $p75^{NTR}$ forms heteromeric complexes with the different Trk receptors and enhances Trk receptor signaling, much like the way the Type III TGF-β receptor enhances the binding of TGF-β to the two signaling TGF-β receptors (see Figure 16-23). These binding relationships between trophic factors and their receptors provide survival signals for different classes of neurons. As nerve axons extend outward from the spinal cord to the periphery, neurotrophins produced by their target tissues bind to Trk receptors on the growth cones (see Figure 18-55) at the tips of the extending axons, promoting survival of those neurons that successfully reach their targets.

To investigate the role of neurotrophins in development, scientists produced mice with knockout mutations in each of the neurotrophins or their receptors. These studies revealed that different neurotrophins and their corresponding receptors are required for the survival of different classes of sensory neurons (Figure 22-40). For instance, pain-sensing (nociceptive) neurons, which express TrkA, are selectively lost from the dorsal root ganglion of knockout mice lacking NGF or TrkA, whereas TrkB- and TrkC-expressing neurons are unaffected in such knockouts. In contrast, TrkC-expressing proprioceptive sensory neurons, which detect the position of the limbs, are missing from the dorsal root ganglion in *TrkC* and *NT-3* mutants.



EXPERIMENTAL FIGURE 22-40 Different classes of sensory neurons are lost in knockout mice lacking different trophic factors or their receptors. In animals lacking nerve growth

factor (NGF) or its receptor TrkA, small nociceptive (pain-sensing) neurons (light blue) that innervate the skin are missing. These neurons express the TrkA receptor and innervate NGF-producing target tissues. In animals lacking either neurotrophin-3 (NT-3) or its receptor TrkC, large proprioceptive neurons (red) innervating muscle spindles are missing. Muscle tissue produces NT-3, and the proprioceptive neurons express TrkC. Mechanoreceptors (orange; see Figure 23-33), another class of sensory neurons in the dorsal root ganglion, are unaffected in these mutants.

[Data from W. D. Snider, 1994, *Cell* **77**:627.]

Mitochondria Play a Central Role in Regulation of Apoptosis in Vertebrate Cells

As discussed previously, *C. elegans* CED-9 and its mammalian homolog Bcl-2 play central roles in regulating apoptosis. In nematodes, CED-9 does so by binding to and blocking the activation of CED-4. In vertebrates, **Bcl-2 family** members have been expanded to include Bcl-2 and Bcl-x_L that function to block apoptosis, as well as Bax and Bak that function to promote apoptosis (Figure 22-41). Bcl-2 family members either constitutively reside in the outer mitochondrial membrane or translocate to the mitochondria when induced. Bcl-2 family members primarily function to maintain (anti-apoptotic) or disrupt (pro-apoptotic) the structural integrity and low permeability of the outer mitochondrial membrane, and to prevent or induce the release of mitochondrial proteins such as cytochrome

c (see Figure 12-22) from the intermembrane space into the cytosol that in turn activate apoptotic caspases.

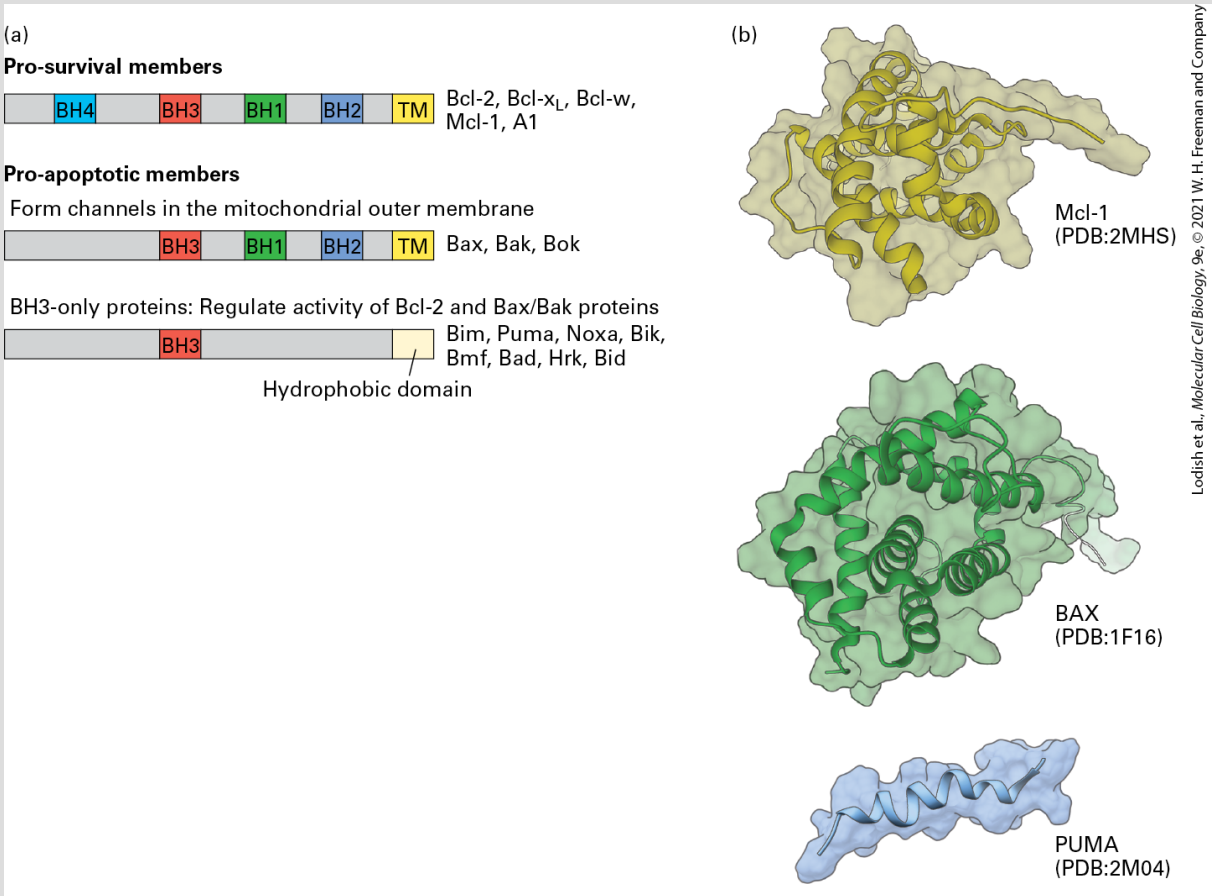


FIGURE 22-41 Structures of members of the Bcl-2 family of proteins. (a) The Bcl-2 family, which comprises proteins that contain functional Bcl-2 homology domains (BH1–4), can be divided into three classes. All of the pro-apoptotic and anti-apoptotic proteins, but only some of the BH3-only proteins, contain a hydrophobic and presumably transmembrane (TM) domain that may function to anchor the protein in the outer mitochondrial membrane. (b) Corresponding three-dimensional structures of representative members of each of these three classes of Bcl-2 homology domains: myeloid cell leukemia (MCL)-1; BCL-2-associated X (BAX); and p53 up-regulated modulator of apoptosis (PUMA) BH3 domain.

[Part (a) Data from Giam et al., 2009, *Oncogene* **27**:S128. Part (b) Data from M. P. A. Luna-Vargas and J. E. Chipuk, 2016, *Trends Cell Biol.* **26**(12):906.]

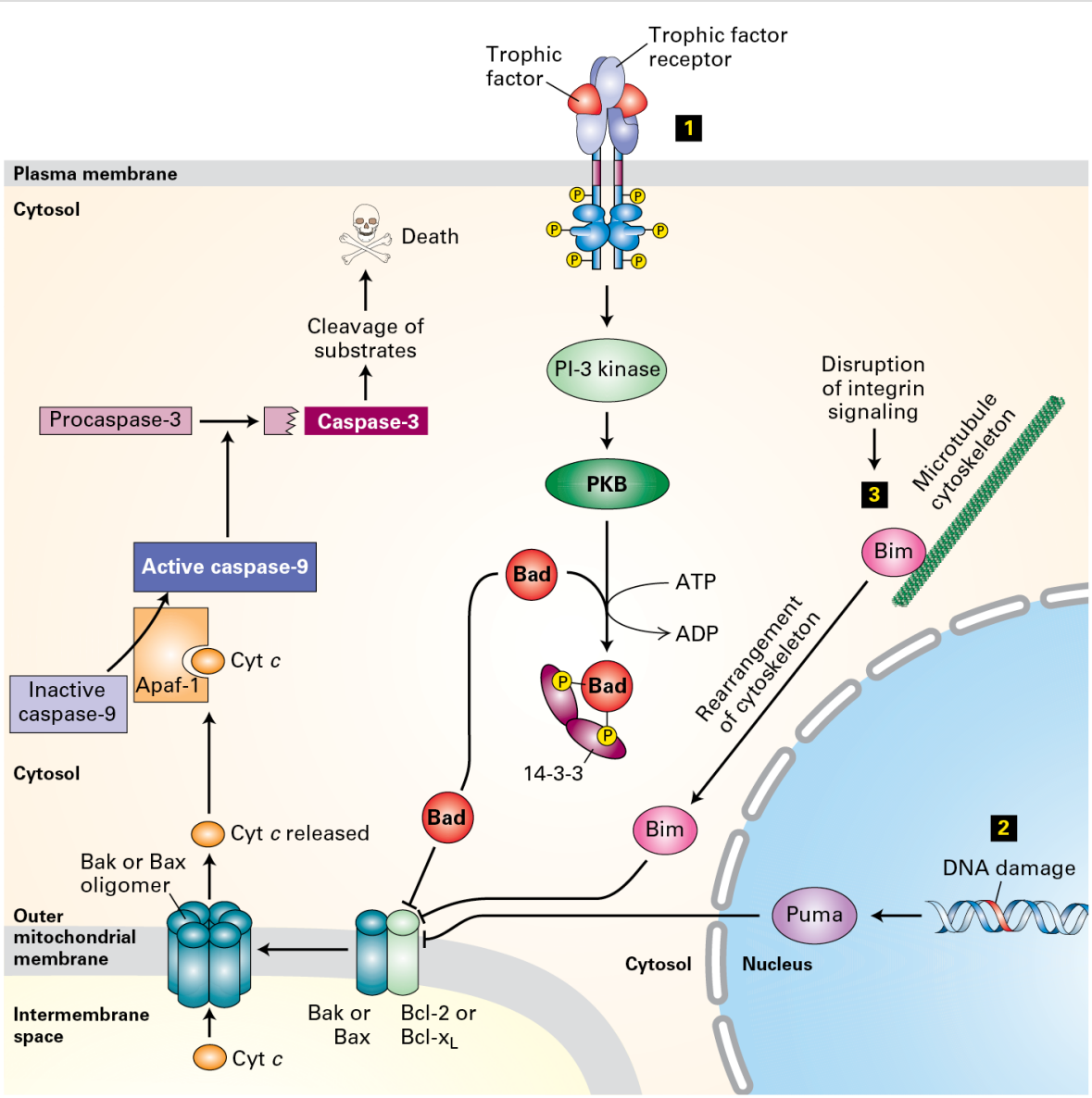
Before we can explain how the activity of Bcl-2 family members is regulated by trophic factors as well as by many environmental stimuli, we need to introduce several other important members of this family of proteins. All members of the Bcl-2 family share a close homology in up to four characteristic regions, termed the *B*cl-2 *h*omology domains (BH1–4; see Figure 22-41). Each of these proteins has either an anti-apoptotic or a pro-apoptotic function. All members of this family participate in oligomeric interactions; many have hydrophobic sequences at their C-termini that can anchor the proteins in the outer mitochondrial membrane.

The Pro-Apoptotic Proteins Bax and Bak Form Pores and Holes in the Outer Mitochondrial Membrane

In vertebrate cells, Bax or Bak is required for mitochondrial damage and induction of intrinsic apoptosis pathways. These two similar pro-apoptotic proteins contain three of the BH1–4 domains (see Figure 22-41) and have three-dimensional structures very similar to that of the anti-apoptotic members of the family. As evidence for the role of these proteins in promoting apoptosis, most mice lacking both Bax and Bak die in utero, and those that survive show significant developmental defects, including the persistence of interdigital webs and accumulation of extra cells in the central nervous and hematopoietic systems. Cells isolated from these mice

are resistant to virtually all apoptotic stimuli. Conversely, overproduction of Bax in cultured cells induces apoptotic death.

Bak resides in the outer mitochondrial membrane, normally tightly bound to the anti-apoptotic protein Bcl-2 or the related protein Bcl-x_L (Figure 22-42). When released from Bcl-2 — either by being present in excess, or by being displaced by the binding to Bcl-2 of certain BH3-only proteins — Bak forms large oligomers that generate pores and holes in the outer mitochondrial membrane. In healthy cells Bax is mainly cytosolic, with a small fraction shuttling to and transiently attached to the mitochondrial surface. Binding of certain pro-apoptotic proteins, discussed later, causes Bax, like Bak, to oligomerize and insert into the outer mitochondrial membrane, forming pores and holes that can be several hundreds of nanometers in diameter. Both pores and holes in the outer mitochondrial membrane allow release into the cytosol of mitochondrial proteins such as cytochrome *c* that, in normal healthy cells, are localized to the intermembrane space.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 22-42 Integration of multiple signaling pathways in vertebrate cells that regulate outer mitochondrial membrane permeability and apoptosis. In healthy cells, the anti-apoptotic protein Bcl-2, or its homolog Bcl-x_L, binds to Bak and Bax pro-apoptotic proteins, blocking the ability of Bak or Bax to oligomerize and form pores in the outer mitochondrial membrane. Binding of any of several BH3-only proteins, including Bad, Bim, and Puma, directly to Bcl-2 causes Bak or Bax to dissociate from Bcl-2 and form oligomeric pores and holes in the outer mitochondrial membrane. These holes allow cytochrome *c* to enter the cytosol, where it binds to the adapter protein Apaf-1, promoting activation of caspase-9 that then initiates the apoptotic cascade and leads to cell death. Several stimuli trigger or repress this apoptotic pathway. Step 1: The presence of specific trophic factors (e.g., NGF)

leads to activation of their cognate receptor tyrosine kinases (e.g., TrkA) and activation of the PI-3 kinase–PKB (protein kinase B) pathway (see Figure 16-17). PKB phosphorylates Bad, and phosphorylated Bad then forms a complex with a cytosolic 14-3-3 protein. This Bad is sequestered in the cytosol and is unable to bind to Bcl-2. In the absence of trophic factors, nonphosphorylated Bad binds to Bcl-2, releasing Bax and Bak and allowing them to form oligomeric membrane pores and holes, activating the apoptosis pathway. Step **2**; DNA damage or ultraviolet irradiation leads to induction of synthesis of the BH3-only Puma protein. Puma binds to Bcl-2, allowing Bak and Bax to form oligomeric pores; Puma may also bind to Bak and Bax, directly activating their ability to form oligomers and permeabilize the outer mitochondrial membrane. Step **3**: Removal of a cell from its substratum disrupts integrin signaling, leading to release of the BH3-only Bim protein from the cytoskeleton. Bim also binds to Bcl-2 to promote pore formation.

[Data from D. Ren et al., 2010, *Science* **330**:1390; P. E. Czabootar et al., 2014, *Nat. Rev. Mol. Cell Biol.* **15**:49; and M. P. A. Luna-Vargas and J. E. Chipuk, 2016, *Trends Cell Biol.* **26**(12):906.]

As depicted in Figure 22-43, released cytochrome *c* activates caspase-9, predominantly by binding to and activating Apaf-1, the mammalian homolog of CED-4 (see Figures 22-36, *right* and 22-42). Binding triggers formation of an apoptosome containing seven molecules each of Apaf and cytochrome *c* — a disk-shaped 1.4-MDa heptameric wheel of death. Two of the Apaf proteins bind to domains in two procaspase-9 molecules, stimulating procaspase-9 binding to adjacent sites in the apoptosome. The conformation changes that occur upon binding trigger activation of this initiator protease; caspase-9 does not require cleavage to become activated. Caspase-9 then cleaves multiple molecules of effector caspases, such as caspase-3 (see Figures 22-36 and 22-42), leading to their activation and subsequent destruction of cell proteins and cell death. As evidence for this pathway, overproduction of Bcl-2 in cultured cells

blocks release of cytochrome *c* and apoptosis; conversely, overproduction of Bax promotes release of cytochrome *c* into the cytosol and promotes apoptosis.

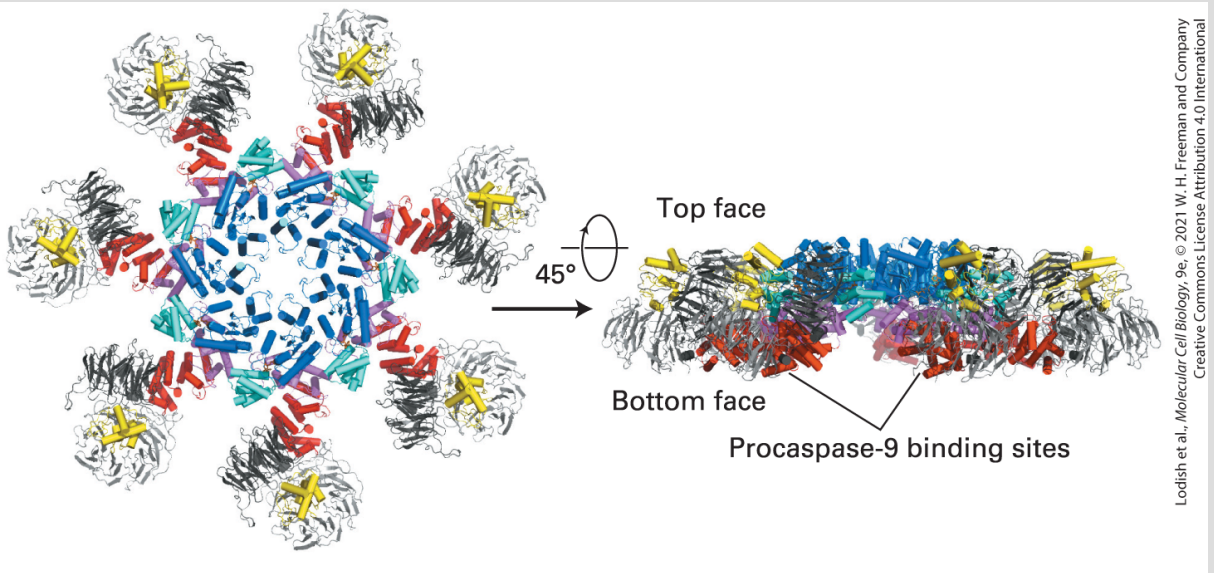


FIGURE 22-43 Structure of the mammalian Apaf-1 apoptosome. Binding of cytochrome *c* to the cytosolic protein Apaf triggers formation of an apoptosome containing seven molecules each of Apaf and cytochrome *c* and having a sevenfold symmetry. The top face refers to the cytochrome *c*–exposed side of the apoptosome disk. Cytochrome *c* is colored yellow, and the several domains within each of the seven Apaf-1 protomers are color-coded. Each Apaf-1 monomer contains a caspase recruitment domain. In the heptameric apoptosome, two of these domains bind to domains in two procaspase-9 molecules, stimulating procaspase-9 binding and dimerization, and thus triggering activation of this initiator protease.

[© 2015 Zhou et al.; Published by Cold Spring Harbor Laboratory Press]

Release of SMAC/DIABLO Proteins from Mitochondria Also Promotes Caspase

Activation

In vertebrates and flies, but not in nematodes, apoptosis is regulated by several other proteins (see Figure 22-36, *right*). XIAP, one member of a family of *i*nhibitor of *a*poptosis *p*roteins (IAPs), provides another way to restrain both initiator and effector caspases. XIAP has three N-terminal BIR domains; the one termed *BIR2* binds to and inhibits two effector caspases, caspase-3 and caspase-7, while BIR3 binds to and inhibits initiator caspase-9. (Other members of the IAP family inhibit apoptosis induced by TNFα; see Figure 22-44.) The inhibition of caspases by IAPs, however, creates a problem when a cell needs to undergo apoptosis. Mitochondria enter the picture once again, since they are the source of a family of proteins, called *SMAC/DIABLOs*, that inhibit IAPs. Assembly of Bax or Bak oligomers (see Figure 22-42) leads to the release of SMAC/DIABLOs, as well as cytochrome *c*, from the mitochondrial intermembrane space. SMAC/DIABLOs then bind to XIAP in the cytosol, thereby blocking XIAP from binding to caspases. By relieving XIAP-mediated inhibition, SMAC/DIABLOs promote caspase activity and cell death.

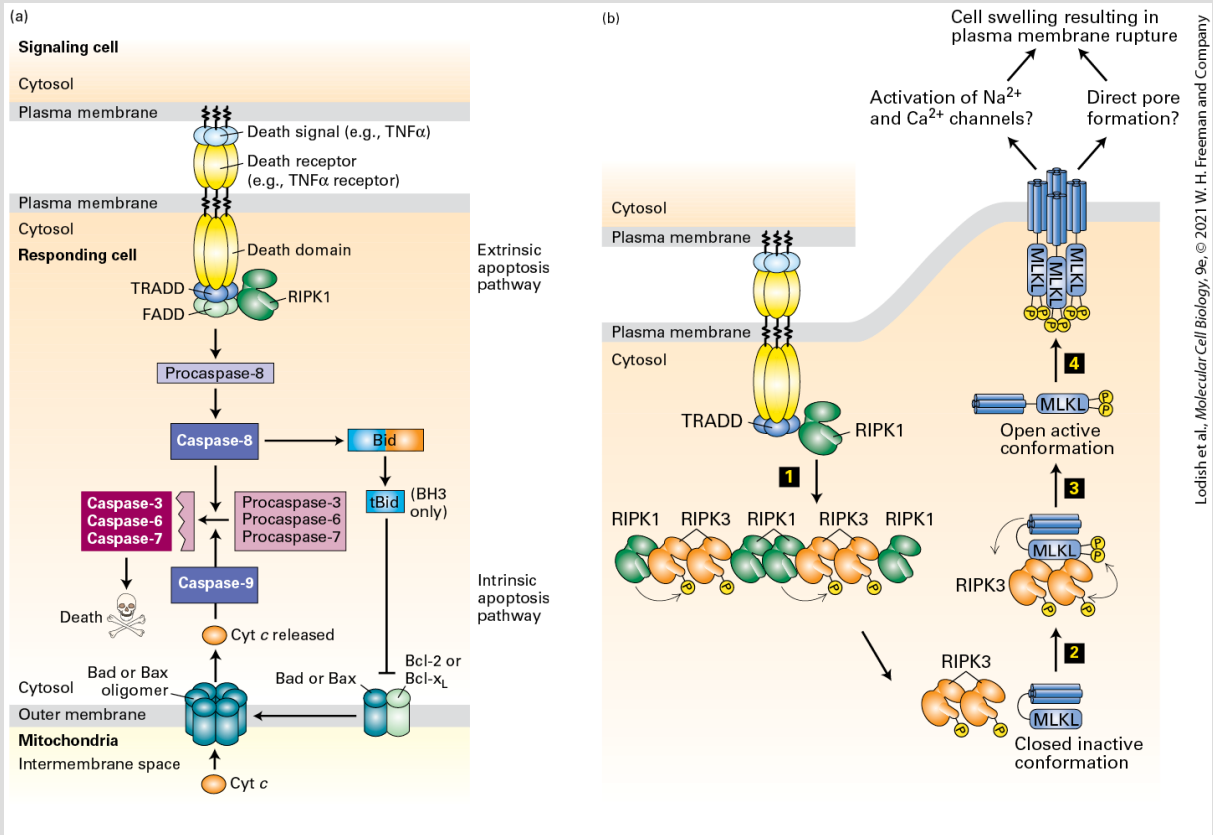


FIGURE 22-44 Cell murder: the extrinsic apoptosis and necroptosis pathways. (a) Extrinsic (or death receptor–regulated) apoptosis pathways are found in many types of cells, which may be activated by Fas, TNFα, and other members of the TNFα family of proteins. For example, binding of TNFα on the surface of one cell to the TNFα death receptor on an adjacent cell leads to the binding of two cytosolic death domain (DD)-containing proteins, the adapter protein TRADD (TNF receptor-associated DD protein) and receptor-interacting-protein 1 (RIPK1), to the DD on the intracellular segment of TNFα receptor 1, leading to the formation of a large intracellular signaling complex. FADD (Fas-associated DD protein) can then bind, leading to the binding, dimerization, and proteolytic activation of the initiator caspase-8. Active caspase-8 then cleaves and activates effector caspases 3, 6, and 7, which cleave vital cellular proteins and induce cell death. Cleavage of the BH3-only protein Bid (BH3-interacting-domain death agonist) by caspase-8 generates the tBid fragment that binds to Bcl-2 on the outer mitochondrial membrane, leading to release of Bad or Bax followed by their oligomerization and pore formation in the outer mitochondrial membrane, release of cytochrome *c* into the cytosol, and activation of the intrinsic apoptosis pathway (see Figure 22-41) as well. RIPK1 is an essential kinase and adaptor in the apoptosis and necroptosis pathway; it contains an N-terminal kinase domain and a RHIM domain. When the decision is to die by apoptosis, RIPK1 binds to FADDs, as above leading to the binding,

dimerization, and proteolytic activation of the initiator caspase-8. (b) When apoptosis fails to be activated, for example in the absence of caspase-8 or FADD, the necroptosis pathway can be activated, triggered by activation of RIPK1 kinase. Step **1**: Activated RIPK1 forms multiprotein complexes via its RHIM domain with the kinase RIPK3, and RIPK3 becomes activated by phosphorylation of a residue in its activation loop. Step **2**: Active RIPK3 then binds to the kinase-like domain of the MLKL protein and phosphorylates it. Step **3**: This induces a conformational change in MLKL that (step **4**) enables it to form oligomers that insert into the plasma membrane. Either directly by forming ion channels or indirectly by interacting with other plasma membrane proteins this destabilizes the integrity of the plasma membrane, inducing cell swelling and rupture. See P. Bouillet and L. A. O'Reilly, 2009, *Nat. Rev. Immunol.* **9**:514; A. Ashkenazi and G. Salvesen, 2014, *Annu. Rev. Cell Dev. Bi.* **30**:337; Y. Dondelinger et al., 2016, *Trends Cell Biol.* **26**:721, https://doi.org/10.1016/j.tcb.2016.06.004; and B. Shan et al., 2018, *Genes Dev.* **32**:327, https://doi.org/10.1101/gad.312561.118.

Trophic Factors Induce Inactivation of Bad, a Pro-apoptotic BH3-Only Protein

We saw earlier that neurotrophins such as NGF protect neurons from cell death; this effect is mediated by inactivation of a pro-apoptotic BH3-only protein called *Bad*. In the absence of trophic factors, Bad is nonphosphorylated and binds to Bcl-2, or to the closely related anti-apoptotic protein Bcl-x_L, at the outer mitochondrial membrane (see Figure 22-42). This binding inhibits the ability of Bcl-2 and Bcl-x_L to bind Bax and Bak, thereby allowing Bak and Bax to oligomerize in the plane of the outer mitochondrial membrane and form pores and holes in it, triggering release of cytochrome *c* and induction of apoptosis.

A number of trophic factors, including NGF, induce the PI-3 kinase signaling pathway, leading to the activation of protein kinase B (see Figure 16-17). Activated protein kinase B phosphorylates Bad; phosphorylated Bad cannot bind to Bcl-2 or Bcl-x_L and is found in the cytosol complexed to the phosphoserine-binding protein 14-3-3 (see Figure 16-13). As evidence for this pathway, a constitutively active form of protein kinase B can rescue cultured neurotrophin-deprived neurons, which would otherwise undergo apoptosis and die. These findings support the mechanism for the survival action of trophic factors depicted in Figure 22-42. In other cell types, different trophic factors promote cell survival through post-translational modification of other components of the cell-death machinery.

Apoptosis in Vertebrates Is Induced by BH3-Only Pro-apoptotic Proteins That Are Activated by Environmental Stresses

Whereas nematodes contain a single BH3-only protein, EGL-1, mammals express at least eight, including Bad, in a cell type- and stress-specific manner. The pro-apoptotic activities of these proteins are tightly regulated by diverse transcriptional and post-transcriptional mechanisms. Two BH3-only proteins, Puma and the related protein Noxa (see Figure 22-42), are transcriptionally induced by the p53 protein (see Figures 19-32 and 19-34); they

induce apoptosis primarily by binding to and inactivating the anti-apoptotic Bcl-2 and Bcl-x_L proteins, thus promoting the polymerization and pore formation of pro-apoptotic Bcl-2 family members Bax and Bak. This interaction is part of the checkpoint pathway by which unrepaired damage to DNA can induce apoptosis; thus the loss of p53 seen in many cancers allows cells to live with severe DNA damage (see Figure 25-23). Another BH3-only protein, Bim, is normally sequestered by the microtubule cytoskeleton by binding to a dynein light chain (see Chapter 18). Detachment of cells from their substratum disrupts integrin signaling, rearranges the cytoskeleton, and leads to release of Bim. Like Puma, Bim binds directly to Bcl-2, releasing Bak and Bax from Bcl-2 and allowing formation of pores and holes in the outer mitochondrial membrane (see Figure 22-42). Thus apoptosis of mammalian cells is regulated by a careful balance of activities of anti-apoptotic proteins such as Bcl-2 and Bcl-x_L and multiple pro-apoptotic BH3-only proteins.

Apoptosis and Necroptosis Can Be Triggered by Tumor Necrosis Factor, Fas Ligand, and Related Death Proteins

Although cell death can arise as a default in the absence of survival factors, apoptosis can also be stimulated by positively acting death signals that activate the extrinsic apoptosis pathway; the extrinsic

pathway activated by death ligands can also activate the alternative programmed cell death pathway termed *necroptosis*. Murder of infected cells is an effective means of stemming the spread of a pathogen throughout an organism. To that end macrophages and other immune system cells are recruited to sites of infection, where they become activated and produce cytokines such as tumor necrosis factor alpha (TNFα) that can remain attached to the cell surface or be cleaved off and function as a secreted hormone. TNFα, in turn, triggers the death of adjacent cells that might have been infected. However, overproduction of TNFα can lead to the tissue destruction seen in certain chronic inflammatory diseases (see Chapter 24). Another important death-inducing signal, the Fas ligand, is a cell-surface protein produced by activated natural killer cells and cytotoxic T lymphocytes. By binding to the Fas protein on target cells, this signal can trigger death of virus-infected cells, some tumor cells, and foreign graft cells. Depending on the type of cell and infection, death can be by apoptosis or by necroptosis (discussed later in this section).

Both TNFα, depicted in Figure 22-44, and the Fas ligand (also called *CD95 ligand*) are trimeric proteins present on the surface of one cell that bind to *death receptors* on an adjacent cell. Each of these death receptors has an extracellular domain that binds to its cognate ligand, a single transmembrane domain, and an intracellular protein-protein interaction domain known as a *death domain (DD),* which is why they are called *death receptors*. These receptors are

activated when binding of a trimeric ligand brings the receptor molecules into close proximity to allow the formation of an intracellular complex that recruits downstream signaling molecules, similar to the formation of the interleukin 1 signalsome (see Figure 16-31). We will use TNFα, a highly pro-inflammatory cytokine involved in mediating a myriad of human diseases, as an example.

Binding of TNFα to the TNF receptor 1 (TNFR1), a DD-containing death receptor, leads to the formation of an intracellular signaling complex which is initiated by the recruitment of another two DD-containing signaling molecules, *T*NF *r*eceptor-*a*ssociated *d*eath *d*omain protein (TRADD) and *r*eceptor-*i*nteracting *p*rotein *k*inase 1 (RIPK1) (Figure 22-44a). TRADD and RIPK1 in turn recruit a large signaling complex, not depicted in the figure, including E3 ubiquitin ligases, deubiquitinating enzymes, and kinases, which collectively decides, within a few minutes, whether the cell should activate NF-κB to survive, or die by apoptosis or necroptosis.

The activation of RIPK1 kinase is a key decisive event that can initiate cell death by either apoptosis or necroptosis. Activated RIPK1 can bind to FADD, which then serves as an adapter to recruit and activate caspase-8, an initiator caspase (Figure 22-44a). FADD can also be directly activated by binding to TRADD in RIPK1-independent apoptosis activation. Like the other initiator caspase, caspase-9, caspase-8 is activated by dimerization following binding of two molecules to the FADD proteins recruited to an active death

receptor trimer. Once activated, caspase-8 cleaves and thus activates several effector caspases which then cleave many cellular proteins to complete apoptotic cell death.

Caspase-8 also cleaves the BH3-only protein *B*H3-*i*nteracting-domain *d*eath agonist (Bid). The resulting tBid fragment (but not the full-length Bid) then binds to Bcl-2 on the outer mitochondrial membrane, leading to the dissociation of Bik and Bax and formation of Bak/Bax pores and holes, release of cytochrome *c* into the cytosol, and activation of the intrinsic apoptosis pathway (see Figure 22-42) as well.

Necroptosis, a genetically programmed type of necrosis, is a well-defined defense mechanism against certain viruses. As part of their replication cycle, many viruses produce an inhibitor of caspase-8, a key component of the TNFα cell murder pathway depicted in Figure 22-44a, thus preventing apoptosis and allowing continued replication of the virus and spread of the infection to neighboring cells. In cells lacking functional caspase-8, TNFα triggers an alternative pathway, necroptosis. In this pathway, activated RIPK1 binds to RIPK3, a family member of RIPK1, promoting activation of RIPK3 and triggering necroptosis. In this process, activated RIPK3 in turn phosphorylates another essential protein termed *MLKL*. Phosphorylation causes MLKL to alter its conformation and form oligomers that insert into the plasma membrane and form holes, allowing Ca^{2+} entry. The influx of Ca^{2+} causes the cell and its

organelles to swell and burst, releasing its contents into the extracellular space. Some of the released intracellular proteins trigger activation of immune-system cells and cause tissue inflammation and damage. Inflammation due to necroptosis has been implicated in mediating major human inflammatory diseases, including rheumatoid arthritis, inflammatory bowel disease, neurodegenerative diseases such as amyotrophic lateral sclerosis and Alzheimer's disease and progressive atherosclerotic lesions.

As we discussed in Chapter 16, protein inhibitors of TNFα are among the most widely used therapeutics for many inflammatory diseases; inhibiting RIPK1 kinase is another promising approach to treating human diseases characterized by necrosis and inflammation. Recall that TNFα activates multiple signal transduction pathways: one leads to activation of the transcription factor NF-κB (see Figure 16-30), a second to apoptosis, and the third to necroptosis (see Figure 22-44). Much work needs to be done to understand the regulation of each of these pathways and their interactions, as this cytokine is involved in many inflammatory diseases.

KEY CONCEPTS OF SECTION 22.4

Cell Death and Its Regulation

- All vertebrate cells require trophic factors to survive. In the absence of these factors, cells commit suicide.
- Genetic studies in *C. elegans* have defined an evolutionarily conserved apoptotic pathway with three major components: membrane-bound regulatory proteins,

cytosolic regulatory proteins, and apoptotic proteases (called *caspases* in vertebrates) (see Figure 22-36).

- Once activated, caspases cleave specific intracellular substrates, leading to the demise of a cell. Other proteins (e.g., CED-4, Apaf-1) that bind regulatory proteins and caspases are required for caspase activation (see Figures 22-36, 22-37, and 22-43).
- Survival of motor and sensory neurons during development is mediated by neurotrophins released from target tissues that bind to Trk receptor tyrosine kinases on the neuronal growth cones (see Figure 22-40), activating an anti-apoptotic response via the PI-3 kinase pathway (see Figure 22-42).
- The Bcl-2 family contains both pro-apoptotic and anti-apoptotic proteins; most are transmembrane proteins and engage in protein-protein interactions.
- In mammals, apoptosis can be triggered by oligomerization of Bax or Bak proteins in the outer mitochondrial membrane, leading to efflux of cytochrome *c* and SMAC/DIABLO proteins into the cytosol; these proteins then promote caspase activation and cell death.
- Bcl-2 proteins can restrain the oligomerization of Bax and Bak, inhibiting cell death.
- Pro-apoptotic BH3-only proteins (e.g., Puma, Bad) are activated by environmental stresses and stimulate the oligomerization of Bax and Bak, allowing cytochrome *c* to escape into the cytosol, bind to Apaf-1, and thus activate caspases.
- Direct interactions between pro-apoptotic and anti-apoptotic proteins lead to cell death in the absence of trophic factors. Binding of extracellular trophic factors can trigger changes in these interactions, resulting in cell survival (see Figure 22-42).
- Binding of extracellular death signals, such as tumor necrosis factor and Fas ligand, to their receptors oligomerizes an associated protein (FADD), which in turn triggers the caspase cascade, leading to cell murder by apoptosis (see Figure 22-44a).
- In the presence of inhibitors of caspase-8, tumor necrosis factor induces necroptosis (Figure 22-44b). Intracellular proteins released into the surroundings as a result can cause inflammation and tissue damage.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

apoptosis
asymmetric cell division
Bcl-2 family
BH3-only protein
blastocyst
caspases
cell lineage
cytokine
embryonic stem (ES) cell
germ line
induced pluripotent stem (iPS) cell
inner cell mass (ICM)
meristem

multipotent
necroptosis
necrosis
neoblast
neurotrophins
planar cell polarity (PCP)
pluripotent
polarity
progenitor (precursor) cell
programmed cell death
somatic cell
stem cell
stem-cell niche
symmetric cell division
totipotent
trophectoderm (TE)
trophic factor

Review the Concepts

1. What two properties define a stem cell? Distinguish between a totipotent stem cell, a pluripotent stem cell, and a precursor (progenitor) cell.
2. Where are stem cells located in plants? Where are stem cells located in adult animals? How does the concept of a stem cell differ between animal and plant systems?

3. In 1997, Dolly the sheep was cloned by a technique called *somatic-cell nuclear transfer* (or *nuclear-transfer cloning*). A nucleus from an adult mammary cell was transferred into an egg from which the nucleus had been removed. The egg was allowed to divide several times in culture, then the embryo was transferred to a surrogate mother who gave birth to Dolly. Dolly died in 2003 after mating and giving birth herself to viable offspring. What does the creation of Dolly tell us about the potential of nuclear material derived from a fully differentiated adult cell? Does the creation of Dolly tell us anything about the potential of an intact, fully differentiated adult cell?
4. Identify whether the following contain totipotent, pluripotent, or multipotent cells: (a) inner cell mass, (b) morula, (c) eight-cell embryo, (d) trophectoderm.
5. True or false: Differentiated somatic cells have the capacity to become reprogrammed to become other cell types. Provide one line of evidence discussed in the chapter that corroborates your response.
6. Explain how intestinal stem cells were first identified and then experimentally shown to be multipotent stem cells.
7. Explain how hematopoietic stem cells were experimentally shown to be both multipotent and capable of self-renewal.
8. The nematode *C. elegans* has proved to be a valuable model organism for studies of cell birth, cell asymmetry, and cell death. What properties of *C. elegans* render it so well suited for these studies? Why is so much information from *C.*

elegans experiments of use to investigators interested in mammalian development?

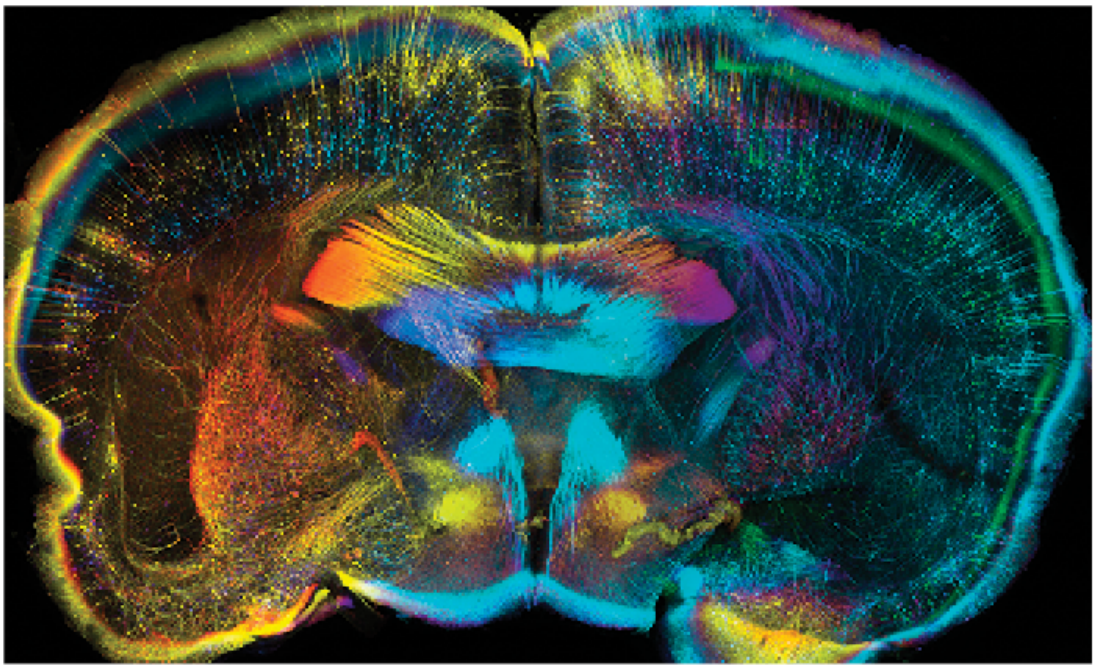
9. Asymmetric cell division often relies on cytoskeletal elements to generate or maintain the asymmetric distribution of cellular factors. In *Saccharomyces cerevisiae*, what factor is localized to the bud by myosin motors? In *Drosophila* neuroblasts, what factors are localized apically by microtubules?
10. Discuss the role of *par* genes in generating anterior/posterior polarity in the *C. elegans* embryo.
11. How do studies of nerve and muscle development in knockout mice support the statement that apoptosis is a default pathway in neuronal cells?
12. Compare and contrast cell death by apoptosis and by necrosis.
13. Identify and list the functions of the three general classes of proteins that control cell death.
14. Based on your understanding of the events surrounding cell death, predict the effect(s) of the following on the ability of a cell to undergo apoptosis:
    a. Functional CED-9; nonfunctional CED-3
    b. Active Bax and cytochrome *c*; nonfunctional caspase-9
    c. Inactive PI-3 kinase; active Bad
15. TNF and Fas ligand bind cell-surface receptors to trigger cell death. Although the death signal is generated external to the cell, why do we consider the death induced by these molecules to be apoptotic rather than necrotic?

16. Predict the effects of the following mutations on the ability of a cell to undergo apoptosis:
    a. Mutation in Bad such that it cannot be phosphorylated by protein kinase B (PKB)
    b. Overexpression of Bcl-2
    c. Mutation in Bax such that it cannot form homodimers

    One common characteristic of cancer cells is a loss of function in the apoptotic pathway. Which of the mutations listed above might you expect to find in some cancer cells?
17. How do IAPs (inhibitors of apoptosis proteins) interact with caspases to prevent apoptosis? How do mitochondrial proteins interact with IAPs to prevent inhibition of apoptosis?

CHAPTER 23

Cells of the Nervous System



Luis de la Torre-Ubieta, Geschwind Laboratory, UCLA.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021
W. H. Freeman and Company

Coronal slice of CLARITY-treated adult mouse brain expressing green fluorescent protein (GFP) in a subset of neurons (Thy1-GFP). CLARITY renders tissue optically transparent, permitting deep and complete imaging of tissues. Each section was stained with antibodies to GFP and color coded by depth to facilitate individual neuron visualization. This approach provides unprecedented opportunity to image intact brains at cellular resolution, paving the way to a comprehensive understanding of how the brain is wired.

OUTLINE

23.1 Neurons and Glia: Building Blocks of the Nervous System
23.2 Voltage-Gated Ion Channels and the Propagation of Action Potentials
23.3 Communication at Synapses
23.4 Sensing the Environment: Touch, Pain, Taste, and Smell

The nervous system, which consists of the central (brain and spinal cord) and peripheral (nerves outside of the brain and spinal cord) nervous systems, regulates all aspects of bodily function and is staggering in its complexity. The 1.3 kg adult human brain — the control center that stores, computes, integrates, and transmits information — contains nearly 100 billion nerve cells, called *neurons*. These neurons are interconnected by some 100 trillion synapses, the junction points where two or more neurons communicate. An individual neuron can form synapses with up to 10,000 other neurons.

Neurons are organized into interconnected units or circuits that have discrete functions. Some circuits sense features of both the external and internal environments of organisms and transmit this information to the brain for processing and storage. Others regulate the contraction of muscles and the secretion of hormones. Yet other circuits regulate cognition, emotion, and innate as well as learned behaviors. In addition to neurons, the nervous system contains glial cells. Historically considered to function simply as support cells for neurons, it is now recognized that glia play active roles in brain function.

The biology of the cells of the nervous system is remarkable on two levels. First, neurons are the most morphologically polarized and

compartmentalized cells in the body; their unique morphology is supported by remarkable cell biological processes, from cytoskeletal dynamics and membrane trafficking to signal transduction and activity-dependent gene regulation. Second, individual neurons and glia combine to form exquisitely complex and precise networks or circuits. Neural circuits are not completely hardwired; instead, the connectivity of neurons changes with experience through a process known as *synaptic plasticity*, in which experience modifies the strength and number of synaptic connections between neurons. A central focus of modern brain biology is to understand the logic underlying both the formation and the plasticity of neural circuits. While the structure and function of nerve cells is understood in great detail, the mechanisms by which neural circuits form, change with experience, and process and compute information remain a mystery. These issues represent some of the most exciting areas of twenty-first-century biology and inspired President Obama to launch the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative in 2013. One major aim of this ongoing, large-scale national effort is to catalogue and characterize the extraordinary diversity of cell types in the brain and to understand the logic by which these cells form circuits that underlie brain functions.

The vertebrate nervous system is anatomically divided into the central nervous system, which contains the nerves and glia located inside the brain and spinal cord, and the peripheral nervous system,

which contains the nerves and glia located outside the brain and spinal cord. Despite being anatomically separate, the central and peripheral nervous systems are functionally interconnected, with peripheral nerves serving as communication conduits between the brain and the body. The central nervous system itself can be divided into four primary components: the spinal cord, brain stem, cerebellum, and cerebrum (Figure 23-1a). Each region has discrete functions. The spinal cord conducts sensory and motor information from the body to the brain, the brain stem regulates basic functions such as breathing and blood pressure, the cerebellum controls motor function, and the cerebrum processes motor and sensory information, language, learning and memory, and other higher level functions. These primary components are further divided into distinct functional areas; for example, the cerebrum is organized into distinct frontal, temporal, parietal, and occipital lobes, each of which is involved in specific, though partially overlapping, functions (Figure 23-1b). Although there are neurons and glia that have distinct properties and characteristics specific to these various regions, the functional specialization of each brain region emerges primarily from differences in circuit connectivity rather than from differences in constituent cell types.

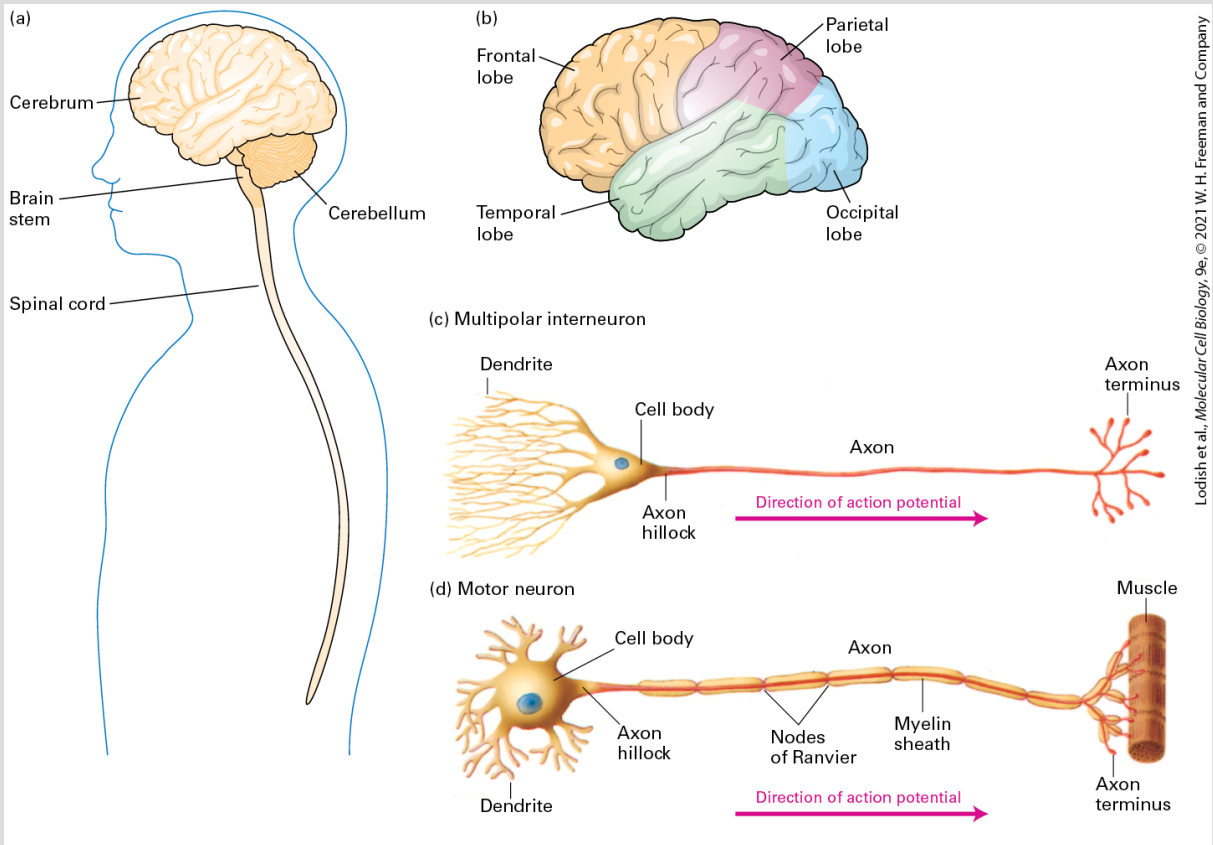


FIGURE 23-1 Major areas of the mammalian central nervous system and typical morphology of two types of mammalian neurons. (a) The mammalian central nervous system is composed of four major areas: the spinal cord, which relays sensory and motor information from the body to the brain; the brain stem, which controls many basic functions such as breathing and heart rate; the cerebellum, which coordinates movements; and the cerebrum, which controls higher functions such as language, learning, memory and emotion. (b) The mammalian cerebrum consists of four lobes, each of which mediates specific, partially overlapping functions. The lobes are: the frontal lobe, involved in executive functions such as problem solving, language, and impulse-control; the temporal lobe, involved in auditory processing, memory, and emotion; the parietal lobe, involved in touch, taste, temperature, and movement sensations; and the occipital lobe, involved primarily in vision. (c) A multipolar interneuron has profusely branched dendrites, which receive signals at synapses from several hundred other neurons. Small voltage changes imparted by inputs in the dendrites can sum to give rise to the more massive action potential, which starts in the hillock. A single, long axon that branches laterally at its terminus transmits signals to other neurons. (d) A motor neuron innervating a muscle cell typically has a single, long axon extending from the cell body to the effector cell. In

mammalian motor neurons, an insulating sheath of myelin usually covers all parts of the axon except at the nodes of Ranvier and the axon termini. The myelin sheath is composed of cells called *oligodendrocytes* in the central nervous system and Schwann cells in the peripheral nervous systems, both of which are glial cells (see Figure 23-18).

Indeed, despite the multiple types and shapes of neurons that are found in metazoan organisms, all nerve cells share common properties that make them specialized for communicating information using a combination of electrical and chemical signaling, as we will see in more detail in the next section. *Electrical signals* conduct information within neurons, which are usually highly polarized cells with extensions whose lengths are orders of magnitude greater than the cell soma (Figure 23-1c and d). The electrical pulses that travel along neurons are called *action potentials,* and information is encoded as the frequency at which action potentials are fired. Owing to the speed of electrical transmission, neurons are champion signal transducers, much faster than cells that secrete hormones. In contrast to the electrical signals that conduct information within a neuron, *chemical signals* transmit information between cells, utilizing processes similar to those employed by other types of signaling cells (see Chapters 15 and 16).

Taken together, the electrical and chemical signaling of the nervous system allows it to detect external stimuli, integrate and process the information received, relay it to higher brain centers, and generate an appropriate response to the stimulus. For example, **sensory**

neurons have specialized receptors that convert diverse types of stimuli from the environment (e.g., light, touch, sound, odorants) into electrical signals. These electrical signals are then converted into chemical signals that are passed on to other cells called **interneurons,** which convert the information back into electrical signals. In some circuits, this information is then transmitted to muscle-stimulating **motor neurons** or to other neurons that stimulate other types of cells, such as glands (an example of this type of simple neural circuit is shown in Figure 23-4).

In this chapter, we will focus on neurobiology at the cellular and molecular level. We will start by looking at the general architecture of neurons, at how they carry signals, and at how neurons and glia arise from stem cells. Next we will focus on ion flow, channel proteins, and membrane properties: how electrical pulses move rapidly along neurons. Third, we will examine communication between neurons where electrical signals traveling along a cell are translated into a chemical pulse between cells and then back into an electrical signal in the receiving cells. We will then examine neurons in several sensory tissues, including those that mediate our senses of touch, taste, and olfaction. The speed, precision, and integrative power of neural signaling enable the accurate and timely sensory perception of a swiftly changing environment. In the last section, we will turn to the circuits, neurons, and cell biological mechanisms underlying the storage of memories.

A great deal of information about nerve cells has been gleaned from analyses of humans, mice, nematodes, and flies bearing mutations that affect specific functions of the nervous system. In addition, molecular cloning and structural analysis of key neuronal proteins, such as voltage-gated ion channels and receptors, have helped elucidate the cellular machinery underlying complex brain functions such as instinct, learning, memory, and emotion. We will be referring to these model organisms and methodologies throughout this chapter.

23.1 Neurons and Glia: Building Blocks of the Nervous System

In this section, we examine the structure of neurons and connect this structure to the function of these cells. **Neurons** are distinguished by their elongated, asymmetric shape, by their highly localized proteins and organelles, and most of all by channel proteins that controls the flow of ions across the plasma membrane. Specific properties of neurons also give the nervous system considerable powers of signal analysis. One neuron can respond to the inputs from multiple neurons, generate electrical signals, and transmit the signals to multiple neurons. For example, a neuron might transmit a signal only if it receives five simultaneous activating signals from input neurons. The receiving neuron measures both the amount of incoming signal and whether the five signals are roughly *synchronous*. Fast synaptic input from one neuron to another can be either *excitatory* — combining with other signals to trigger electrical transduction in the receiving cell — or *inhibitory*, discouraging such transmission. In addition to excitatory and inhibitory synapses, neurons receive slower *neuromodulatory* inputs such as norepinephrine, dopamine, serotonin, and acetylcholine, which activate G protein–coupled receptors (see Chapter 15) to change the threshold for excitation or inhibition. Thus the properties and connections of individual neurons set the stage for integration and refinement of information. The output of a

nervous system is the result of its circuit properties, that is, the wiring or interconnections between neurons and the strength of these interconnections. We will begin by looking at how signals are received and sent, and in subsequent parts of the chapter we will look at the molecular details of the machinery involved.

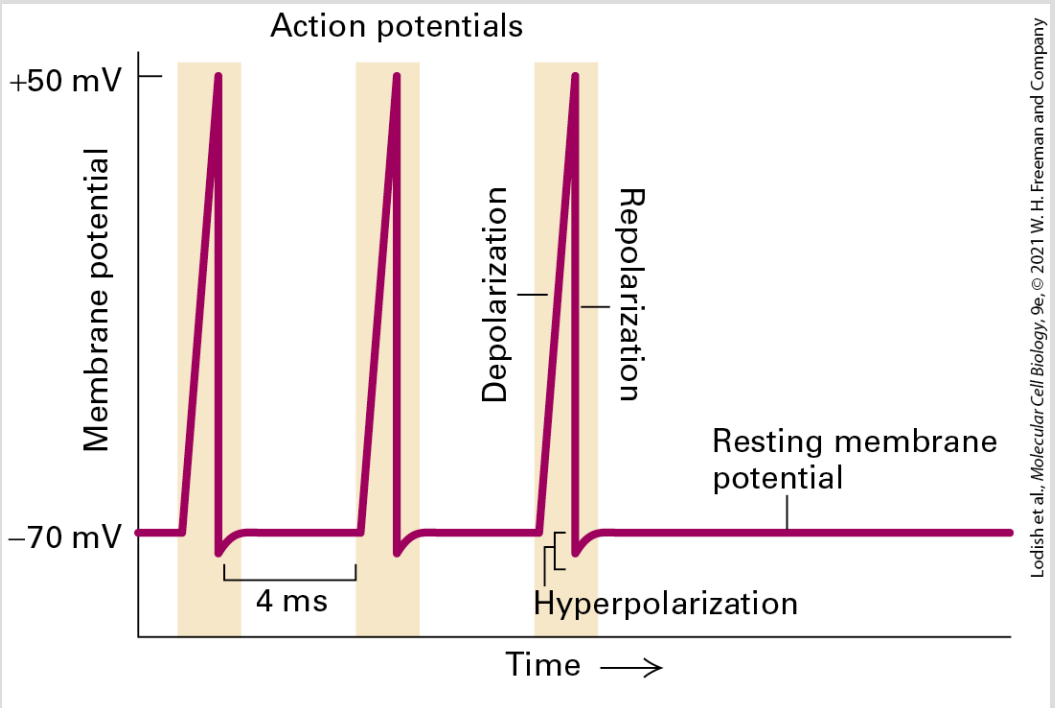
Information Flows Through Neurons from Dendrites to Axons

Neurons arise from roughly spherical *neuroblast* precursors. Newly born neurons can migrate long distances before growing into dramatically elongated cells. Fully differentiated neurons take many forms, but generally share certain key features (see Figure 23-1c and d). The nucleus is found in a rounded part of the cell called the *cell body*. Branching cell processes called **dendrites** (from the Greek for "treelike") are found at one end and are the main structures where signals are received from other neurons via synapses. Incoming signals are also received at synapses that form on neuronal cell bodies. Neurons often have extremely long dendrites with complex branches, particularly in the central nervous system (i.e., the brain and spinal cord). This allows them to form synapses with, and receive signals from, a large number of other neurons — up to tens of thousands. Thus the converging dendritic branches allow signals from many cells to be received and integrated by a single neuron.

When a neuron is first differentiating, the end of the cell opposite the dendrites undergoes dramatic outgrowth to form a long, extended arm called the **axon**, a structure that essentially acts as a transmission wire. The growth of axons must be controlled so that proper connections are formed; this is accomplished through a complex process called *axon guidance* that involves dynamic changes to the cytoskeleton (see Section 18.8). The diameters of axons vary from just a micrometer in certain neurons of the human brain to a millimeter in the giant axon of the squid. Axons can be meters in length (e.g., in giraffe necks) and are often partly covered with electrical insulation called the **myelin sheath** (see Figure 23-1d), which is made by specific classes of glial cells: oligodendrocytes (in the central nervous system) and Schwann cells (in the peripheral nervous system). The insulation speeds electrical transmission. The short, branched ends of the axon at the opposite end of the neuron from the dendrites are called the *axon termini.* This is where signals are passed along to the next neuron or to another type of cell, such as a muscle or hormone-secreting cell. The asymmetry of the neuron, with dendrites at one end and axon termini at the other, is indicative of the primarily unidirectional flow of information from dendrites to axons.

Information Moves Along Axons as Pulses of Ion Flow Called *Action Potentials*

Nerve cells are members of a class of *excitable cells,* which also includes muscle cells, cells in the pancreas, and some others. Like all metazoan cells, excitable cells have an inside-negative voltage or electric potential gradient across their plasma membranes, the **membrane potential** (see Chapter 11). Unlike other cells, neurons can use their unique electrical properties to send signals. The signals take the form of brief local voltage changes, from inside-negative to inside-positive, an event designated **depolarization**. A powerful surge of depolarizing voltage change, moving from one end of the neuron to the other, is called an **action potential**. A recording of several depolarization events, with the neuron going from inside-negative to neutral to inside-positive, is shown in Figure 23-2. At the peak of an action potential, the membrane potential can be as much as +50 mV (inside-positive), a net change of ~120 mV. As we will see in greater detail in Section 23.2, an action potential moves along the axon to the axon terminus at speeds of up to 100 meters per second. In humans, for instance, axons may be more than a meter long, yet it takes only a few milliseconds for an action potential to move along their length. Neurons can fire repeatedly after a brief recovery period, for example, every 4 milliseconds (ms), as in Figure 23-2. After the action potential passes through a section of a neuron, channel proteins and pumps restore the inside-negative resting potential (**repolarization**). The restoration process chases the action potential down the axon to the terminus, leaving the neuron ready to signal again.



EXPERIMENTAL FIGURE 23-2 Recording of an axonal membrane potential over time reveals the amplitude and frequency of action potentials. An action potential is a sudden, transient depolarization of the membrane, followed by repolarization to the resting potential of about −70 mV. The axonal membrane potential can be measured with a small electrode placed into it (see Figure 11-19). This recording shows the neuron generating one action potential about every 4 milliseconds.

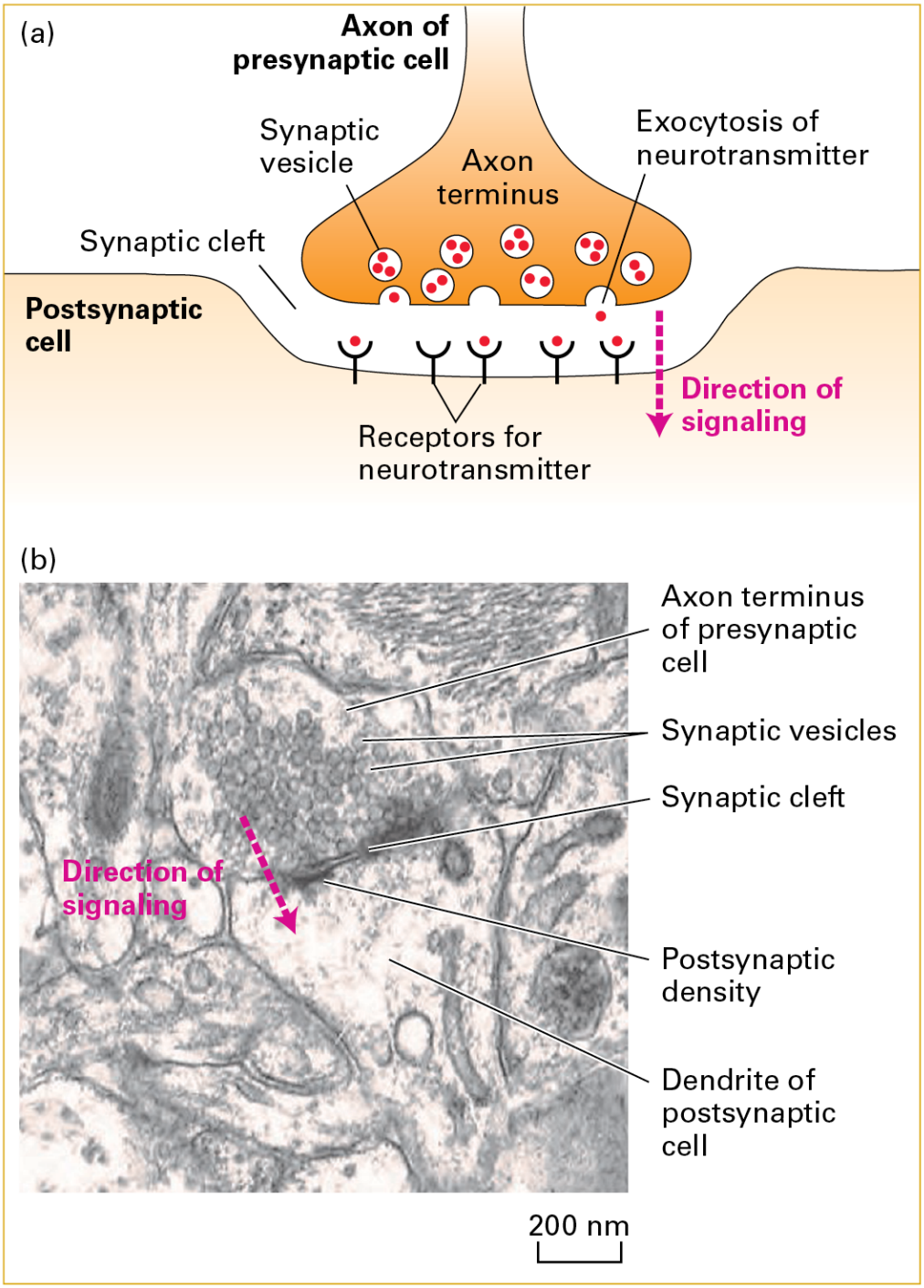
Action potentials are all or none. Once the threshold to start one is reached, a full firing occurs. The signal information is therefore carried primarily not by the intensity of the action potentials but by their timing and frequency. The stronger the stimulus, the higher the frequency of action potential firing.

As previously mentioned, neurons are not the only excitable cells in the body. Muscle contraction is triggered by motor neurons that

synapse directly on excitable muscle cells (see Figure 23-1d). Insulin secretion from the β islet cells of the pancreas is triggered by neurons. In both cases, the activating event involves an opening of plasma membrane channels that causes changes in the transmembrane flow of ions and in the electrical properties of the regulated, non-neuronal cells.

Information Flows Between Neurons via Synapses

How is information communicated between neurons? Axon termini from one neuron are closely apposed to dendrites of another, at junctions called *chemical synapses* or simply **synapses** (Figure 23-3). The axon terminus of the *presynaptic cell* contains many small vesicles, termed **synaptic vesicles**, each of which is filled with a single kind of small molecule known as a **neurotransmitter**. Arrival of an action potential at a presynaptic terminus causes an influx of calcium that triggers exocytosis of a small number of synaptic vesicles, releasing their content of neurotransmitter molecules.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 23-3 A chemical synapse. (a) A narrow region, the synaptic cleft, separates the plasma membranes of the presynaptic and postsynaptic cells. Arrival of action potentials in a presynaptic cell causes exocytosis at a synapse of a small number of synaptic vesicles, releasing their content of neurotransmitters (red circles). Following their diffusion across the synaptic

cleft, the neurotransmitters bind to specific receptors on the plasma membrane of the postsynaptic cell. These signals either depolarize the postsynaptic membrane (making the potential inside less negative), tending to induce an action potential in the cell, or hyperpolarize the postsynaptic membrane (making the potential inside more negative), inhibiting action potential induction. (b) Electron micrograph showing an axon terminus filled with synaptic vesicles forming a synapse with a postsynaptic dendrite, with its characteristic postsynaptic density. In the synaptic region, the plasma membrane of the presynaptic cell is specialized for vesicle exocytosis; synaptic vesicles containing a neurotransmitter are clustered in these regions. The opposing membrane of the postsynaptic cell (in this case, a neuron) contains receptors for the neurotransmitter. The postsynaptic density is a protein-dense specialization that contains receptors, adhesion molecules, scaffolding molecules, signaling molecules, and cytoskeletal elements.

[Part (b) republished with permission from Oxford University Press, from S. Okabe, 2013, "Fluorescence Imaging of Synapse Formation and Remodeling," *Microscopy* **62**(1):51–62; permission conveyed through Copyright Clearance Center, Inc.]

Neurotransmitters diffuse across the synapse in about 0.5 ms and bind to receptors on the dendrite of the adjacent neuron. Binding of neurotransmitter triggers the opening or closing of specific ion channels in the plasma membrane of *postsynaptic cell* dendrites, leading to changes in the membrane potential in this localized area of the postsynaptic cell. Generally these changes depolarize the postsynaptic membrane. The local depolarization, if large enough, triggers an action potential in the axon. This type of chemical neurotransmission is unidirectional, from the axon termini of the presynaptic cell to dendrites of the postsynaptic cell.

In some synapses, the effect of the neurotransmitters is to hyperpolarize and therefore lower the likelihood of an action potential in the postsynaptic cell. A single axon in the central nervous system can synapse with many neurons and induce responses in all of them simultaneously. Conversely, sometimes multiple neurons must act on the postsynaptic cell roughly synchronously to have a strong enough impact to trigger an action potential. Neuronal integration of depolarizing and hyperpolarizing signals determines the likelihood of an action potential. Thus neurons employ a combination of extremely fast electrical transmission along the axon with rapid chemical communication between cells. This is known as *electrochemical signaling*.

The Nervous System Uses Signaling Circuits Composed of Multiple Types of Neurons

In complex multicellular animals, neurons form signaling circuits composed of multiple, distinct types of neurons. In the spinal cord, signaling circuits that relay information between peripheral and central nervous systems consist of three basic types of nerve cells: afferent neurons, interneurons, and efferent neurons. **Afferent neurons**, also known as *sensory neurons*, carry nerve impulses from receptors or sense organs toward the central nervous system (i.e., the brain and spinal cord). These neurons report an event that has

happened, like the arrival of a flash of light or the movement of a muscle. A touch or a painful stimulus creates a sensation in the brain only after information about the stimulus travels there via afferent nerve pathways. **Efferent neurons**, also known as *effector neurons*, carry nerve impulses away from the central nervous system to generate a response. A motor neuron, for example, carries a signal to a muscle to stimulate its contraction (see Figure 23-1d); other effector neurons stimulate hormone secretion by endocrine cells. Interneurons, the largest group, relay signals from afferent to efferent neurons and to other interneurons as part of a neural pathway. An interneuron can bridge multiple neurons, allowing integration or divergence of signals and sometimes extending the reach of a signal. In a simple type of circuit called a *reflex arc*, interneurons connect multiple sensory and motor neurons, allowing one sensory neuron to affect multiple motor neurons and one motor neuron to be affected by multiple sensory neurons; in this way, interneurons integrate and enhance reflexes. For example, the knee-jerk reflex in humans, illustrated in Figure 23-4, involves a complex reflex arc in which one muscle is stimulated to contract while another is inhibited from contracting. The reflex also sends information to the brain to announce what happened. Such circuits allow an organism to respond to a sensory input by the coordinated action of sets of muscles that together achieve a single purpose.

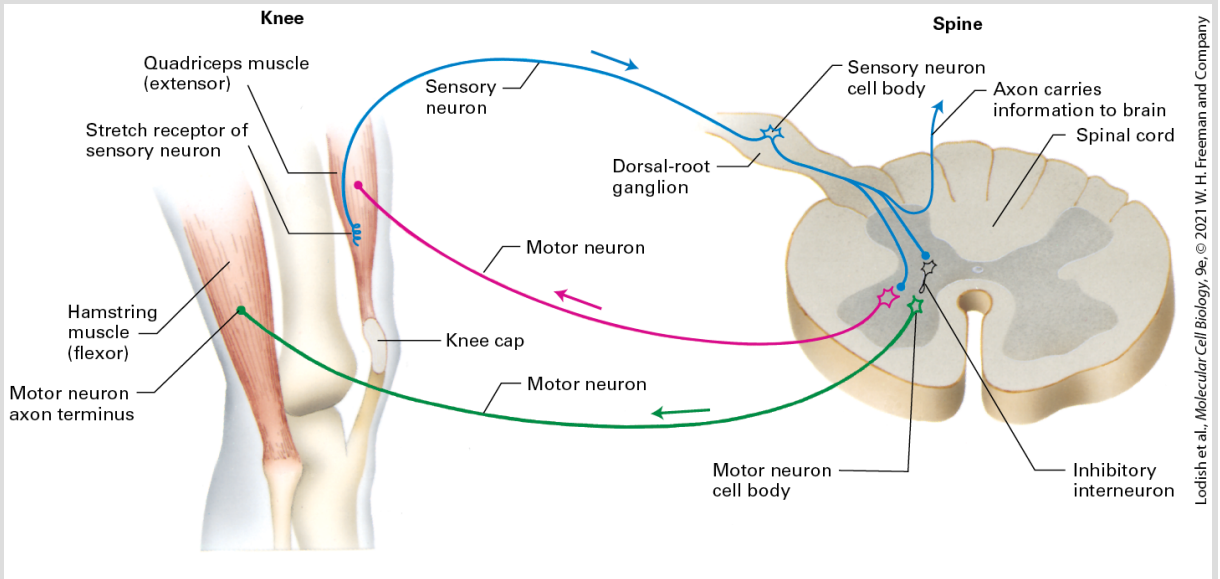


FIGURE 23-4 The knee-jerk reflex. A tap of the hammer stretches the quadriceps muscle, thus triggering electrical activity in the stretch receptor sensory neuron. The action potential, traveling in the direction of the top blue arrow, sends signals to the brain, so we are aware of what is happening, and also to two kinds of cells in the dorsal-root ganglion that is located in the spinal cord. One cell, a motor neuron that connects back to the quadriceps (red), stimulates muscle contraction so that you kick the person who hammered your knee. The second connection activates (or "excites") an inhibitory interneuron (black). The interneuron has a damping effect, blocking activity by a flexor motor neuron (green) that would in other circumstances activate the hamstring muscle that opposes the quadriceps. In this way, relaxation of the hamstring is coupled to contraction of the quadriceps. This is a reflex because movement requires no conscious decision.

The simple signaling circuits described above, however, do not directly explain higher order brain functions such as reasoning, computation, and memory development. Neurons in these circuits include an array of excitatory neurons, which promote action potential firing in a target neuron; inhibitory neurons, which inhibit action potential firing in a target neuron; and neuromodulatory neurons, which modulate the threshold for action potential firing in

a target neuron. Typical neurons in the brain receive signals from up to a thousand other neurons and, in turn, can direct chemical signals to many other neurons. The output of the nervous system depends on its circuit properties — the amount of wiring, or interconnections, between neurons and the strength of these interconnections. As complex and diverse as neural circuits are, they are comprised of a few basic patterns. These include divergence, in which one presynaptic neuron makes connections with many postsynaptic neurons; convergence, in which one postsynaptic neuron receives inputs from many presynaptic neurons; and feedback, in which the output of a postsynaptic neuron feeds back onto a presynaptic neuron or even onto itself (Figure 23-5). Feedback circuits form what are known as *closed loops*, in which the output of a system is used as the input. In a positive feedback circuit, the output sustains or increases the activity of the initial input. In a negative feedback circuit, the output inhibits the activity of the initial input.

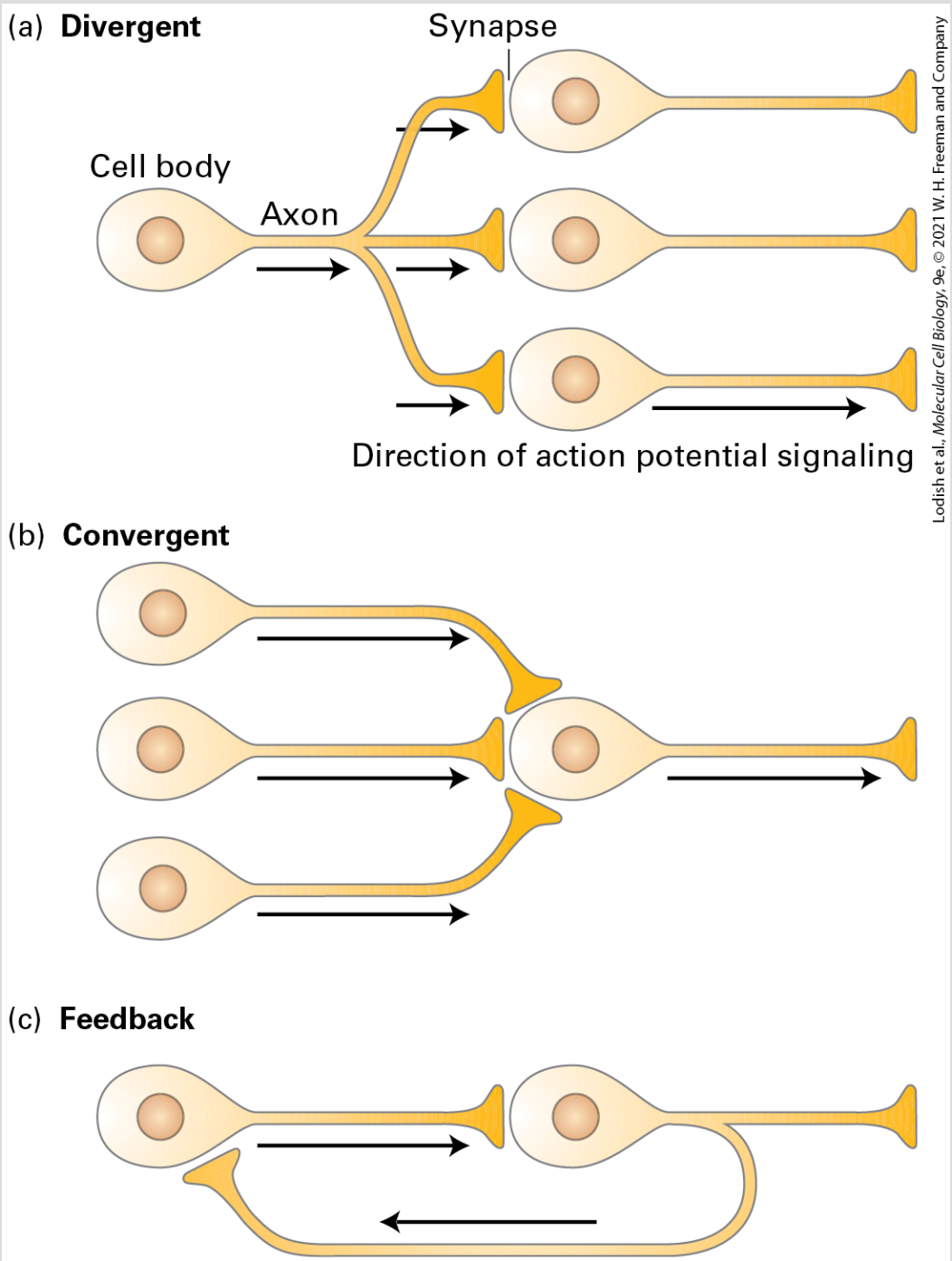


FIGURE 23-5 Common patterns in neural circuits. Neurons connect to one another to form functional circuits. Shown are three common patterns of connectivity that are found in many neural circuits. (a) In divergent neural circuits, a single neuron sends axonal branches to contact many different target neurons. (b) In convergent neural circuits, many different neurons send axonal branches that converge to contact a single target neuron. (c) In feedback

circuits, a neuron sends an axon to communicate with a neuron that is presynaptic to it. Combinations of these and other patterns of interconnectivity function to communicate information within neural circuits.

Glial Cells Form Myelin Sheaths and Support Neurons

Neurons are not the only essential cells in the human brain. **Glial cells** (also known as *neuroglia* or simply *glia*), which play many roles in the brain but do not themselves conduct electrical impulses, are present in the brain in large numbers. While many textbooks claim that glia outnumber neurons by 10 to 1, recent experiments have suggested that the ratio of glia to neurons in the human brain is closer to 1:1, although there is significant variability between species and brain areas. As described above, of the four principal types of glia, two produce myelin sheaths — the insulation that surrounds neuronal axons (see Figure 23-1d): **oligodendrocytes** make sheaths for the central nervous system (CNS), and **Schwann cells** make them for the peripheral nervous system (PNS). (Both types of glia are discussed in more detail in Section 23.2.) Astrocytes, a third type of glia, provide growth factors and other signals to neurons and also receive signals from neurons. A fourth type of glia, *microglia*, constitutes a part of the CNS immune system. While microglia are not related by lineage to neurons or to other glia, they do play important roles in brain development and health. We describe the function of astrocytes below; oligodendrocytes and

Schwann cells will be discussed in Section 23.2, and microglia in Section 23.3.

Astrocytes, named for their starlike shape (Figure 23-6), constitute about a third of the brain's mass and ~20 percent to 40 percent of the brain's cells. Recent studies have revealed a remarkable molecular, physiological, and functional diversity of astrocytes. Understanding the function of astrocytes in brain function is an area of active investigation. Astrocytes surround many synapses and dendrites; the Ca^{2+}, K^{+}, Na^{+}, and Cl^{-} channels found in astrocyte plasma membranes influence the concentration of free ions in the extracellular space, thus affecting the membrane potentials of neurons and of the astrocytes themselves. Astrocytes produce abundant extracellular matrix proteins, some of which are used as guidance cues by migrating neurons. They also produce a host of growth factors that carry a variety of types of information to neurons. In addition, astrocytes release several factors that are necessary for proper synapse formation between neurons, as discussed in Section 23.3. Astrocytes are joined to each other by gap junctions (see Figure 20-21 for structure of gap junctions), so changes in ionic composition in a given astrocyte are communicated to adjacent astrocytes. In this way, like neurons, astrocytes can also form networks, called *astrocyte syncytia*, within the nervous system.

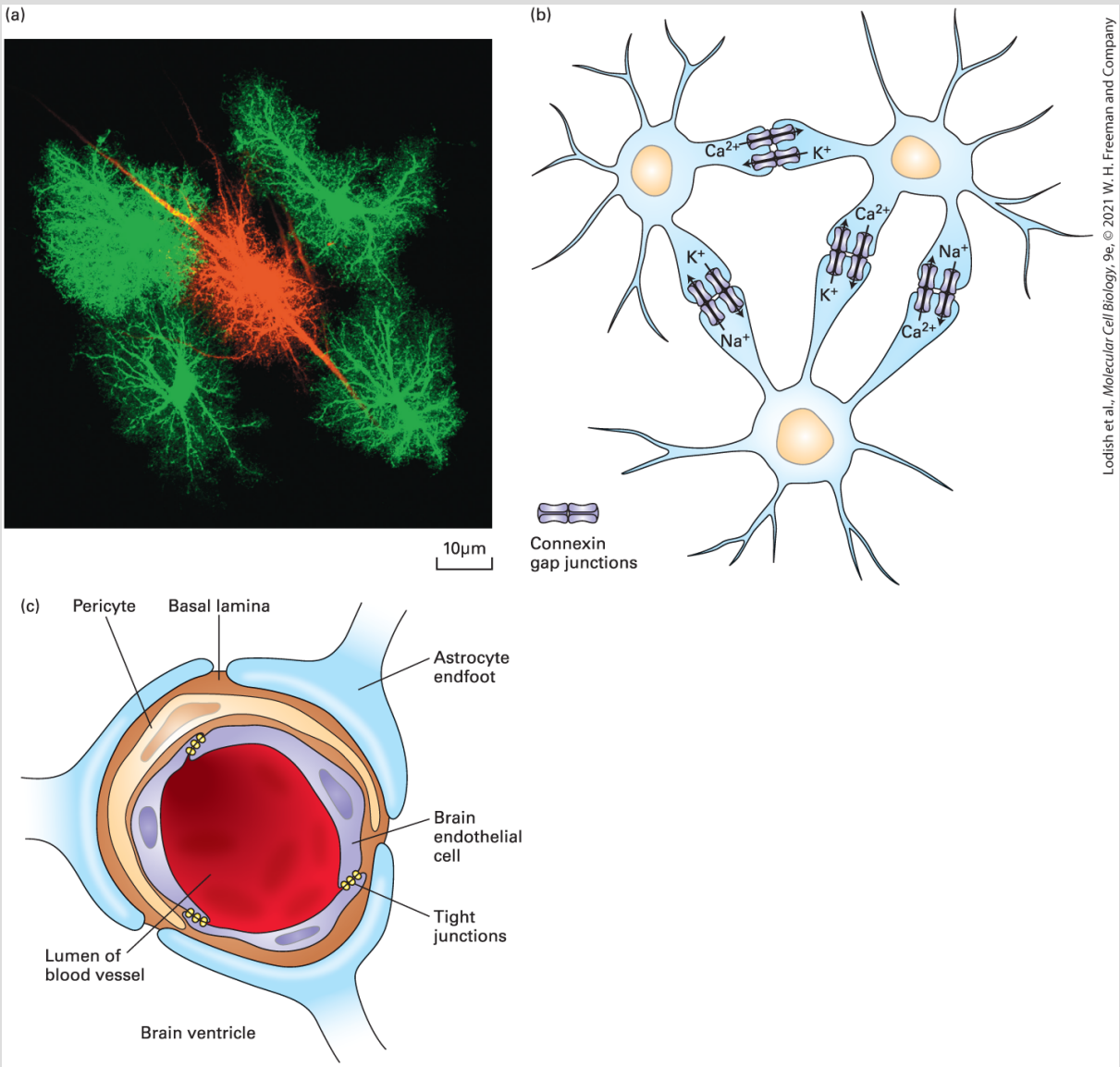


FIGURE 23-6 Astrocytes: morphology, gap-junction-mediated networks of astrocytes, and interaction with endothelial cells at the blood-brain barrier. (a) Astrocytes are so named for their starlike shape. In this photomicrograph, astrocytes are labeled in magenta with an antibody that recognizes glial fibrillary acidic protein (GFAP), an intermediate filament protein that reveals the major branches of the astrocytes. One astrocyte in the middle of the field was also microinjected with a soluble green fluorescent dye, which shows that, in addition to the major branches (seen as white from the merged magenta and green labels), astrocytes contain many fine branches called *branchlets*. (b) Astrocytes are connected to one another through gap junctions (see Figure 20-21), through which ions such as Ca^{2+}, K^+, and Na^+ can rapidly flow from one astrocyte to another to form networks or syncytia of astrocytes. Not shown here, astrocytes also form gap junctions with neurons and

oligodendrocytes. (c) Astrocytes contribute to the blood-brain barrier. Capillaries in the brain are formed by endothelial cells that are interconnected by tight junctions that are impermeable to most molecules. Transport between cells is blocked, so only small molecules that can diffuse across plasma membranes or substances specifically transported through cells can cross the barrier. Certain astrocytes surround the blood vessels, in contact with the endothelial cells, and send secreted protein signals to induce the endothelial cells to produce a selective barrier. The endothelial cells (purple) are ensheathed by a layer of basal lamina (brown), which contains contractile cells called *pericytes* (yellow) and are contacted on the outside by astrocyte processes (blue). See N. J. Abbott, L. Rönnbäck, and E. Hansson, 2006, *Nat. Rev. Neurosci.* **7**:41–53.

[Part (a) UC San Diego, Eric Bushong, Maryann Martone, Mark Ellisman (2002) CCDB:1053, rattus norvegicus, protoplasmic astrocyte. CIL. Dataset. https://doi:org/10.7295/W9CCDB1053.]

Some astrocytes are also critical regulators of the formation of the blood-brain barrier, which controls what types of molecules can travel out of the bloodstream into the brain and vice versa (see Figure 23-6c). Blood vessels in the brain supply oxygen and remove CO_2 and deliver glucose and amino acids, with capillaries found within a few micrometers of every cell. The blood-brain barrier is formed by tight junctions (Chapter 20) between endothelial cells in brain blood vessels. Endfeet projections from astrocytes interact closely with the brain endothelial cells and with the basal lamina (see Figures 20-23 and 20-24), which contains contractile cells called *pericytes*, to regulate blood-brain barrier integrity and function. Many scientists are interested in understanding the blood-brain barrier both because it appears to break down in certain neurological diseases and because it represents a major barrier to the development of effective neurotherapeutics.

Neural Stem Cells Form Nerve and Glial Cells in the Central Nervous System

The great interest in the formation of the nervous system and in finding better ways to prevent or treat neurodegenerative diseases through cell replacement therapies has made the characterization of neural stem cells and understanding what triggers their differentiation into mature neurons and glia an important goal. Much of what we understand about neural and glial stem cells comes from studies of embryonic brain development. The earliest stages of vertebrate neural development involve the rolling up of a tube of ectoderm, the cell layer that lines the outside of the embryo and extends the length of the embryo from head to tail (Figure 23-7a). This *neural tube* will form the brain and spinal cord. Initially, the thickness of the tube is a single layer of cells, and these cells, referred to as *neuroepithelial cells*, serve as the embryonic neural stem cells (NSCs) that will give rise to the entire central nervous system. The inside of the neural tube will expand in the forebrain to form the fluid-filled compartments called *ventricles*, and the cellular layer lining the neural tube, called the *ventricular zone (VZ)*, is where most cell division will take place.

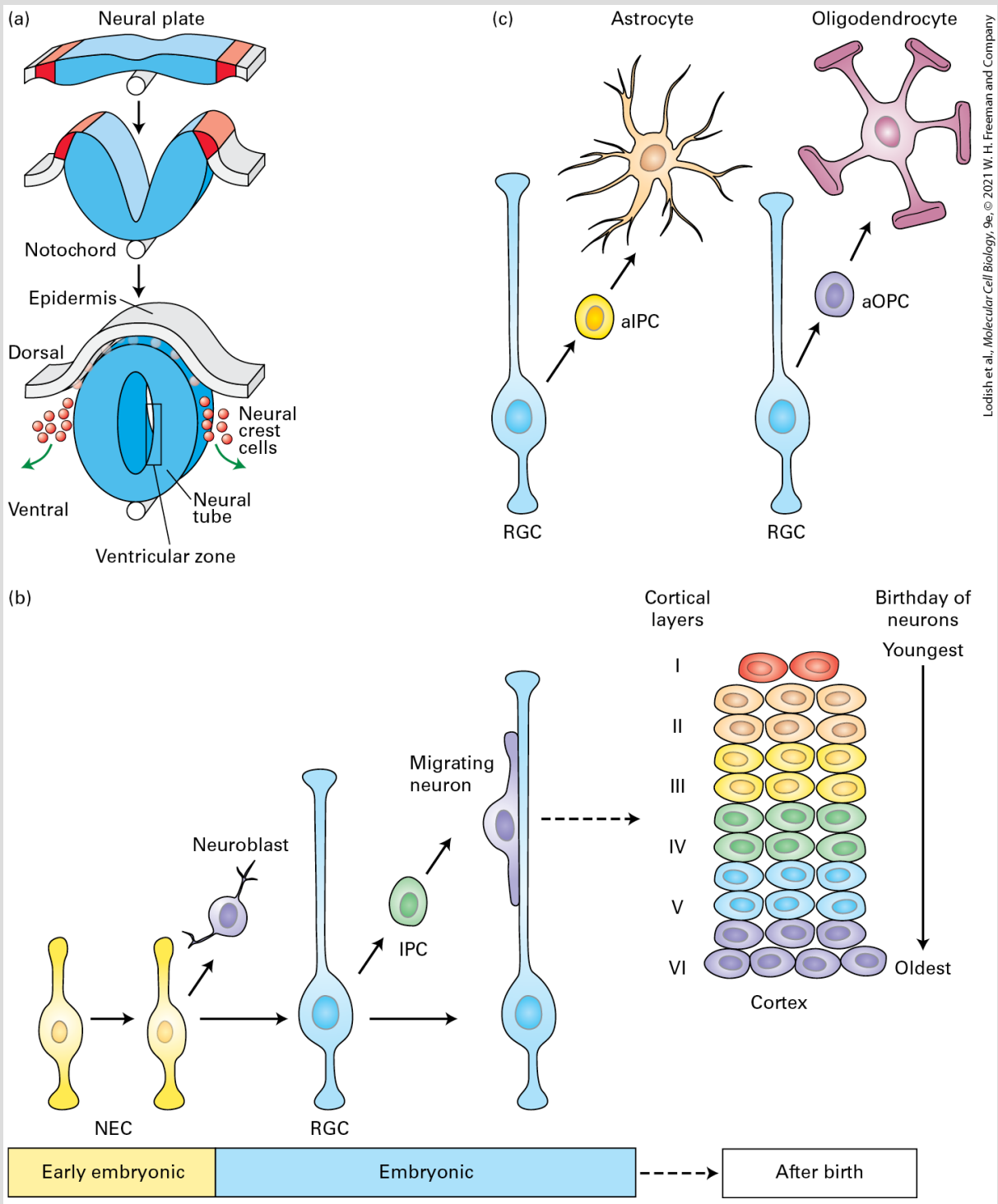


FIGURE 23-7 Formation of the neural tube and division of neural stem cells. (a) Early in vertebrate development a part of the ectoderm rolls up and separates from the rest of the cells. This forms the epidermis (gray) and the neural tube (blue). Near the interface between the two, neural crest cells form and then migrate to contribute to skin pigmentation, nerve

formation, craniofacial skeleton, heart valves, peripheral neurons, and other structures. The notochord, a rod of mesoderm for which chordates are named, provides signals that affect cell fates in the neural tube. The interior of the neural tube will become a fluid-filled series of chambers called *ventricles*. Neural stem cells located adjacent to the ventricles, described as being in the ventricular zone (VZ), will divide to form neurons that migrate radially outward to form the layers of the nervous system. (b) Early in development, in the neuroepithelium, neuroepithelial cells (NECs, yellow) divide to generate more NECs. Some are also thought to generate early neurons also called *neuroblasts* (purple). As development progresses, NECs convert into radial glial cells (RGCs). RGCs divide symmetrically or asymmetrically to generate neurons (purple) or intermediate progenitor cells (IPCs, green), which in turn generate neurons. The RGCs elongate and send a process down to the VZ and a process up to the meninges. Newly born neurons migrate along the elongated RGC to reach their final position, such that the oldest neurons are closest to the VZ and the youngest neurons are closest to the meninges. In the cortex, this means that the youngest neurons are in the most superficial cortical layers (I and II) and the oldest born neurons are in the deeper cortical layers (V and VI). (c) Near the end of embryonic development, RGCs convert into astrocyte intermediate precursor cells (aIPCs) and oligodendrocyte intermediate precursor cells (oIPCs), which in turn generate astrocytes and oligodendrocytes, respectively. See A. Kriegstein and A. Alvarez-Buylla, 2009, *Annu. Rev. Neurosci.* **32**:149–184.

Labeling and tracing experiments have been done in mouse models to determine how cells are born and where they go after birth. The embryonic neuroepithelial cells (NECs), the neural stem and progenitor cells that line the ventricle, can divide symmetrically, producing two daughter stem or progenitor cells side by side (Figure 23-7b), thereby expanding the progenitor population. Around the same time that neuron production begins, the NECs transform into radial glial cells (RGCs), which are the primary precursor cells during embryonic neurogenesis. Through a series of symmetric and asymmetric cell divisions, radial glial cells give rise to neurons and intermediate precursor cells (see Figure 23-7b). During later stages

of brain development, RGCs also give rise to glia, including both astrocytes and oligodendrocytes (Figure 23-7c).

For many years it was believed that no new nerve cells are formed in the adult brain. Most mammalian brain cells indeed stop dividing by adulthood, but some cells in the lateral ventricle, in a region called the *adult subventricular zone (SVZ)*, and in a region of the **hippocampus** (a region of the temporal lobe involved in memory formation) called the *subgranular zone (SGZ)* continue to act as stem cells to generate new neurons (Figure 23-8a). Similar to other types of stem cells, these neural stem cells are functionally defined by their ability to self-renew and differentiate into neural lineages, including neurons, astrocytes, and oligodendrocytes (Figure 23-8b). To identify and characterize neural stem cells, cells isolated from the SVZ were cultured with growth factors such as FGF2 or EGF. Some of the cells survived and proliferated in an undifferentiated state; that is, they could self-renew. In the presence of other growth factors, these undifferentiated cells gave rise to neurons, astrocytes, or oligodendrocytes. The successful establishment of self-renewing and multipotent cells from the adult brain provides strong evidence for the presence of nerve stem-cell populations. While the function of these new neurons in the adult brain is not yet understood, studies in rodents have shown that their survival is increased by enriched environments and by exercise (Figure 23-8b, bottom panel). Such studies may eventually provide the basis for therapeutic applications in neural repair in humans.

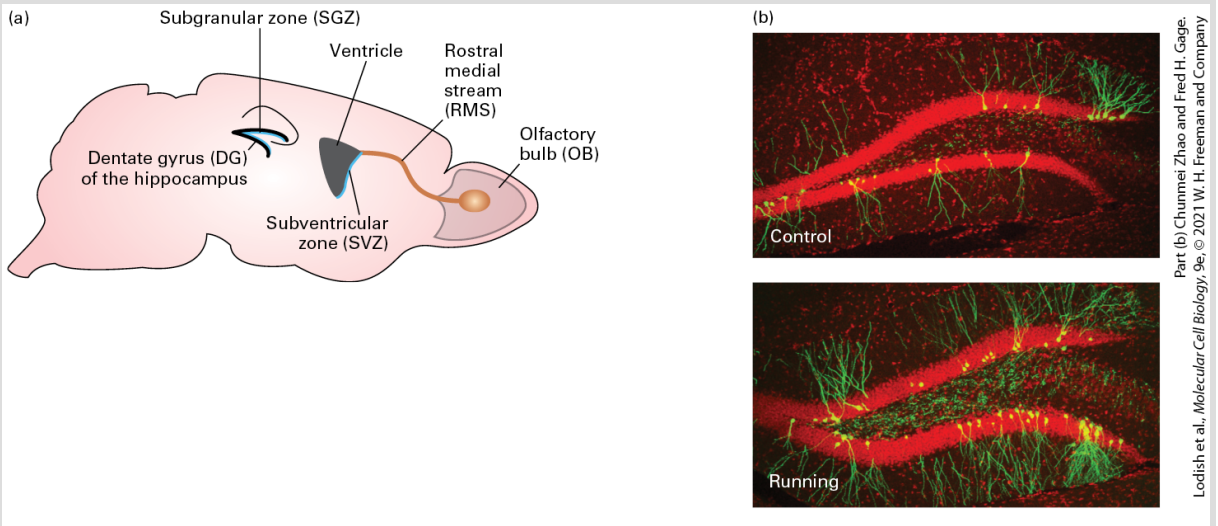


FIGURE 23-8 Neurogenesis in the adult brain. (a) New neurons are born in two regions of the adult brain, the subgranular zone (SGZ) within the dentate gyrus (DG) in the hippocampus and the subventricular zone (SVZ). Neurons derived from neural stem cells in the SVZ migrate to the olfactory bulb (OB) via the rostral medial stream (RMS) in mice. (b) Newly born neurons in the SGZ were labeled with a retrovirus that expresses GFP. Shown are sections of the DG from control mice and from mice that were allowed to exercise on a running wheel in their cages for a week. The newly born neurons are green, and the extensive dendritic branches they have developed indicate that they have survived and have become incorporated into the hippocampus. All the other cells in the hippocampus are labeled with a red nuclear marker. The dense red labeling in the DG (the sideways V-shaped structure) are the cell bodies of the granule cells. The other red cell bodies represent glial cells and inhibitory neurons. As this image illustrates, the percentage of granule cells in the DG that are newly born is very small and is significantly increased by running.

KEY CONCEPTS OF SECTION 23.1

Neurons and Glia: Building Blocks of the Nervous System

- Neurons are highly asymmetric cells composed of multiple dendrites at one end, a cell body containing the nucleus, a long axon, and axon termini.
- Neurons carry information from one end to the other using pulses of ion flow across the plasma membrane; signals are received by the dendrites and transmitted by the axons (see Figure 23-1).

- The resting potential in unstimulated neurons is −70 mV; a stimulus causes opening of channels that leads to depolarization to +50 mV (see Figure 23-2).
- The strong pulse of voltage change, called an *action potential*, travels down the axon by the opening of channels at speeds of up to 100 meters per second.
- Neurons connect to other cells via synapses. At the axon termini of the presynaptic cell the signal is converted from electrical to chemical to stimulate the postsynaptic cell.
- Upon stimulation by an action potential, axon termini release, by exocytosis, small packets of chemicals called *neurotransmitters*. Neurotransmitters diffuse across the synapse and bind to receptors in the postsynaptic cells (see Figure 23-3).
- Neurons form circuits that usually consist of afferent (sensory) neurons, interneurons, and efferent (motor) neurons.
- Neurons connect with one another to form circuits. Three fundamental patterns of neuronal connectivity include divergent, convergent, and feedback circuits (see Figure 23-5).
- Glial cells are abundant in the nervous system and serve many purposes. Two types of glial cells, oligodendrocytes and Schwann cells, build the myelin insulation that coats many neurons. Astrocytes, another type of glial cell, wrap their processes around synapses and blood vessels and promote formation of the blood-brain barrier (see Figure 23-6). Astrocytes also secrete proteins that stimulate synapse formation and participate in the formation and function of neural circuits.
- Embryonic neural stem cells in the ventricular zone give rise to all cells in the central nervous system. These stem and progenitor cells undergo a series of symmetric and asymmetric cells to produce more progenitor cells, glia, and neurons (see Figure 23-7).
- In the adult brain, new neurons are born in the subventricular zone (SVZ) and in the subgranular zone (SGZ) in the dentate gyrus region of the hippocampus (see Figure 23-8).

23.2 Voltage-Gated Ion Channels and the Propagation of Action Potentials

Information processing in the nervous system depends on signaling in neurons, which in turn depends on the function of a set of membrane channels. In Chapter 11, we learned that an electric potential of ~70 mV (cytosolic face negative) exists across the plasma membrane of all cells, including resting nerve cells. This resting membrane potential is generated by outward movement of K^+ ions through open nongated K^+ channels in the plasma membrane and is driven by the K^+ concentration gradient (greater in the cytosol than in the extracellular medium). The high cytosolic K^+ and low cytosolic Na^+ concentrations, relative to their concentrations in the extracellular medium, are generated by the plasma membrane Na^+/K^+ pump (see Figure 11-12), which uses the energy released by hydrolysis of phosphoanhydride bonds in ATP to pump Na^+ outward and K^+ inward. The entry of Na^+ ions into the cytosol from the medium is thermodynamically favored, driven both by the Na^+ concentration gradient (greater in the extracellular medium than in the cytosol) and the inside-negative membrane potential (see Figure 11-24). However, most Na^+ channels in the plasma membrane are closed in resting cells (including neurons), so little inward movement of Na^+ ions can occur (Figure 23-9a).

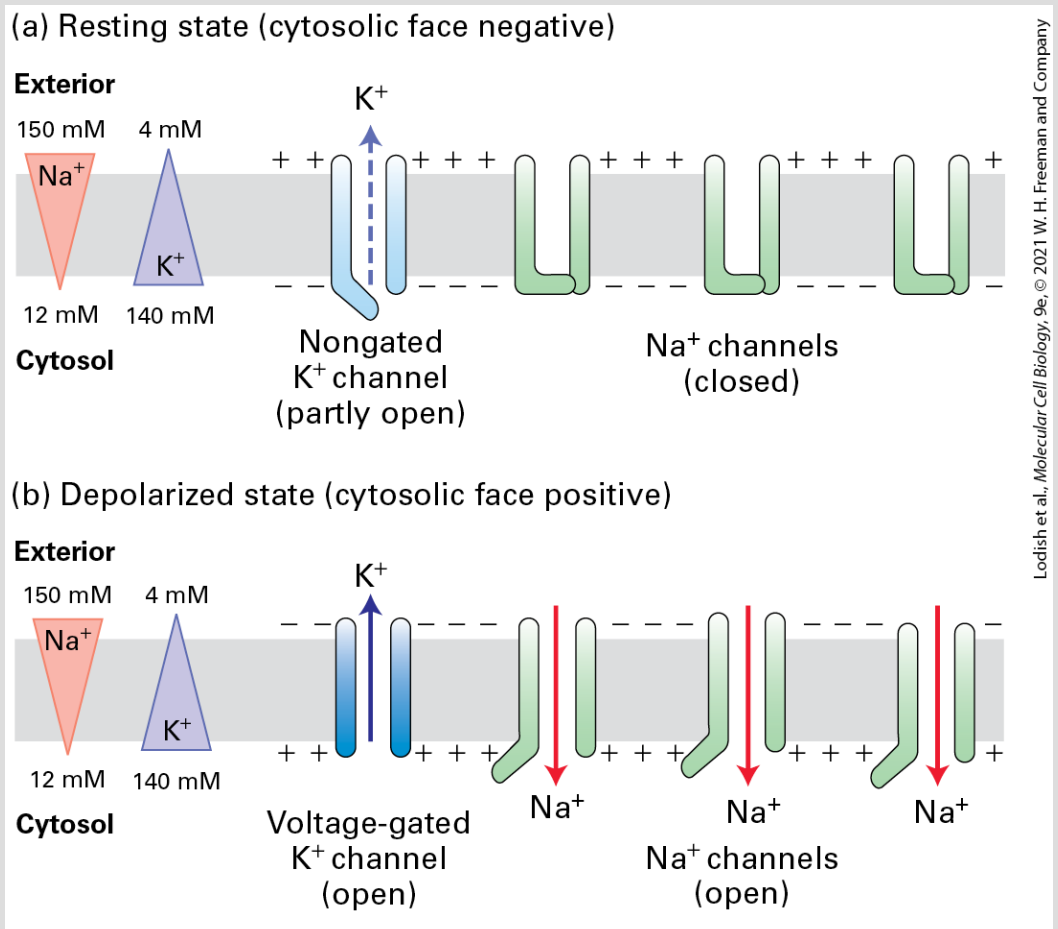


FIGURE 23-9 Depolarization of the plasma membrane is due to opening of gated Na^+ channels. (a) In resting neurons, a type of nongated K^+ channel (blue) is open part of the time, but the more numerous gated Na^+ channels (green) are closed. The movement of K^+ ions outward establishes the inside-negative membrane potential characteristic of most cells. (b) Opening of gated Na^+ channels permits an influx of sufficient Na^+ ions to cause a reversal of the membrane potential. In the depolarized state, voltage-gated K^+ channels open and subsequently repolarize the membrane. Note that the flows of ions are too small to have much effect on the overall concentration of either Na^+ or K^+ in the cytosol or exterior fluid.

During an action potential, some of these Na^+ channels open, allowing inward movement of Na^+ ions, which depolarizes the

membrane. Action potentials are propagated down the axon because a change in voltage in one part of the axon triggers the opening of channels in the next section of the axon. Such **voltage-gated channels** therefore lie at the heart of neural transmission. In this section, we first introduce some of the key properties of action potentials, which move rapidly along the axon from the cell body to the axon termini. We then describe how the voltage-gated channels responsible for propagating action potentials in neurons operate. In the last part of the section, we will see how the myelin sheath, produced by glial cells, increases the speed and efficiency of electrical transmission in nerve cells.

The Magnitude of the Action Potential Is Close to E_{Na} and Is Caused by Na^+ Influx Through Open Na^+ Channels

The depolarization of the plasma membrane is driven by the opening of Na^+ channels. Figure 23-9b illustrates how the membrane potential will change if enough Na^+ channels in the plasma membrane open. The resulting influx of positively charged Na^+ ions into the cytosol will more than compensate for the efflux of K^+ ions through open resting K^+ channels. The result will be a net inward movement of cations, generating an excess of positive charges on the cytosolic face of the plasma membrane and a corresponding excess of negative charges on the extracellular face

(owing to the Cl^- ions "left behind" in the extracellular medium after influx of Na^+ ions). In other words, the plasma membrane becomes depolarized to such an extent that the inside face becomes positive with respect to the external face.

Recall from Chapter 11 that the equilibrium potential of an ion is the membrane potential at which there is no net flow of that ion from one side of the membrane to the other due to the balancing of two opposing forces, the ion concentration gradient and the membrane potential. At the peak of depolarization in an action potential, the magnitude of the membrane potential is very close to the Na^+ equilibrium potential E_{Na} given by the Nernst equation (Equation 11-2), as would be expected if opening of voltage-gated Na^+ channels is responsible for generating action potentials.

Sequential Opening and Closing of Voltage-Gated Na^+ and K^+ Channels Generate Action Potentials

The cycle of changes in membrane potential and return to the resting value that constitutes an action potential lasts 1–2 milliseconds and can occur hundreds of times a second in a typical neuron (see Figure 23-2). The rapidity of these changes in membrane potential is critical to the ability of neurons to encode the intensity of stimuli by modulating the frequency of action

potential firing, with low-intensity stimuli driving lower frequency of action potential firing, and stronger stimuli driving higher frequency of action potential firing. The cyclical changes in the membrane potential result first from the opening and closing of a number of *voltage-gated Na^+ channels* (i.e., channels opened by a change in membrane potential) in a segment of the axonal plasma membrane and then from the opening and closing of *voltage-gated K^+ channels.* The role of these channels in the generation of action potentials was elucidated in classic studies done on the giant axon of the squid, in which multiple microelectrodes can be inserted without causing damage to the integrity of the plasma membrane. However, the same basic mechanism, involving voltage-gated Na^+ and K^+ channels, is used by all neurons.

Voltage-Gated Na^+ Channels

As just discussed, voltage-gated Na^+ channels are closed in resting neurons. Depolarization of the membrane (as occurs when neurotransmitter stimulates a postsynaptic cell) causes a conformational change in these channel proteins that opens a gate on the cytosolic surface of the pore, permitting Na^+ ions to pass through the pore into the cell. The greater the initial membrane depolarization, the more voltage-gated Na^+ channels that open and the more Na^+ ions that enter.

As Na^+ ions flow inward through opened channels, the excess positive charges on the cytosolic face and negative charges on the exoplasmic face diffuse a short distance away from the initial site of depolarization. This *passive spread* of positive charges on the cytosolic face and negative charges on the external face depolarizes adjacent segments of the plasma membrane, causing opening of additional voltage-gated Na^+ channels in these segments and an increase in Na^+ influx. As more Na^+ ions enter the cell, the inside of the cell membrane becomes more depolarized, causing the opening of yet more voltage-gated Na^+ channels and even more membrane depolarization, setting into motion an explosive entry of Na^+ ions. For a fraction of a millisecond, the permeability of this small segment of the membrane to Na^+ becomes vastly greater than that for K^+, and the membrane potential approaches E_{Na}, the equilibrium potential for a membrane permeable only to Na^+ ions. At this point, however, further net inward movement of Na^+ ions ceases, since the concentration gradient of Na^+ ions (greater outside than inside) is now offset by the inside-positive membrane potential. The action potential is, at its peak, close to the value of E_{Na}.

Figure 23-10 schematically depicts the critical structural features of voltage-gated Na^+ channels and the conformational changes that cause their opening and closing. In the resting state, a segment of the protein on the cytosolic face — the *gate* — obstructs the central pore, preventing passage of ions. The channel contains four

positively charged *voltage-sensing α helices*; in the resting state these helices are attracted to the inside-negative surface of the plasma membrane. A small depolarization of the membrane triggers movement of these voltage-sensing helices toward the negative charges that are building up on the exoplasmic surface, causing a conformational change in the gate that opens the channel and allows Na^+ ion flow. After about 1 ms, further Na^+ influx is prevented by movement of the cytosol-facing *channel-inactivating segment* into the open channel, blocking any further movement of Na^+ ions. As long as the membrane remains depolarized, the channel-inactivating segment remains in the channel opening; during this **refractory period**, the channel is inactivated and cannot be reopened. A few milliseconds after the inside-negative resting potential is reestablished, the channel-inactivating segment swings away from the pore and the voltage-sensing α helices return to their resting position near the cytosolic surface of the membrane. Thus the channel returns to the closed resting state, once again able to be opened by depolarization. Note the important distinction between "closed" channels and those that are "inactive" as depicted in Figure 23-10.

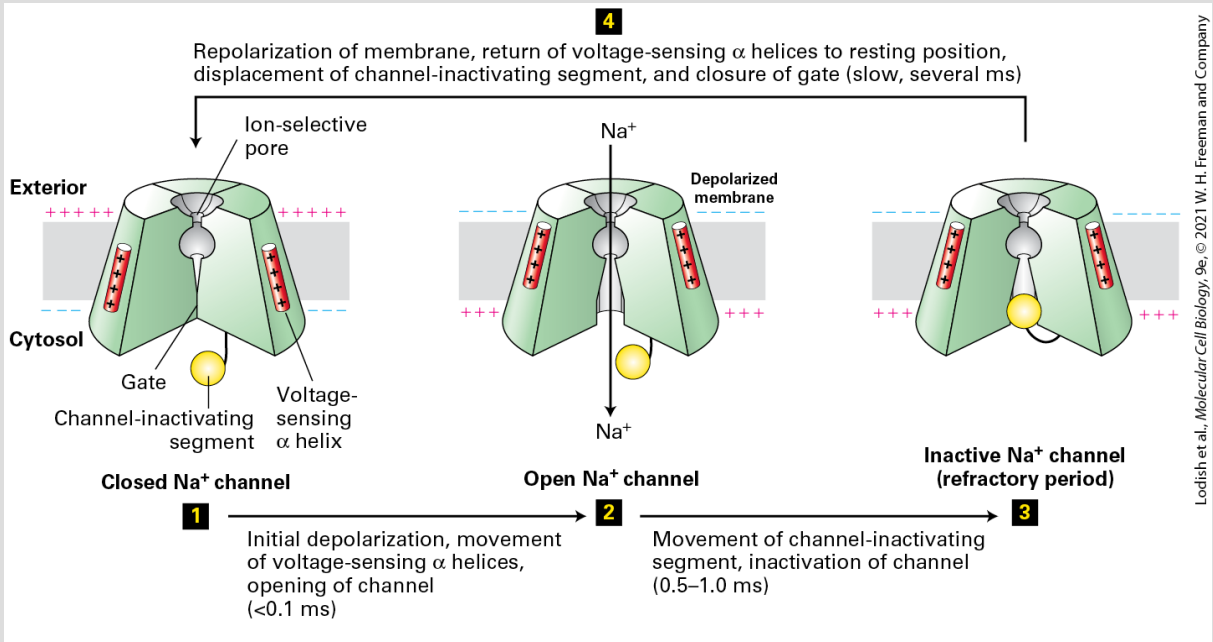


FIGURE 23-10 Operational model of the voltage-gated Na^+ channel. As in the K^+ channel depicted in Figure 11-20, four transmembrane domains in the protein contribute to the central pore through which ions move. The critical components that control movement of Na^+ ions are shown here in the cutaway views depicting three of the four transmembrane domains. In the closed, resting state, the voltage-sensing α helices, which have positively charged side chains every third residue, are attracted to the negative charges on the cytosolic side of the resting membrane. This keeps the gate segment near the cytosolic face in a "closed" position that blocks the channel, preventing entry of Na^+ ions (step 1). In response to a small depolarization, the voltage-sensing helices move through the phospholipid bilayer toward the outer membrane surface, causing an immediate conformational change in the gate at the cytosolic face of the protein that opens the channel (step 2). Within a fraction of a millisecond the channel-inactivating segment moves into the open channel, preventing passage of further ions (step 3). Once the membrane is repolarized, the voltage-sensing helices return to the resting position, the channel-inactivating segment is displaced from the channel opening, and the gate closes; the protein reverts to the closed, resting state and can be opened again by depolarization (step 4). See W. A. Catterall, 2001, *Nature* **409**:988; and S. B. Long et al., 2007, *Nature* **450**:376.

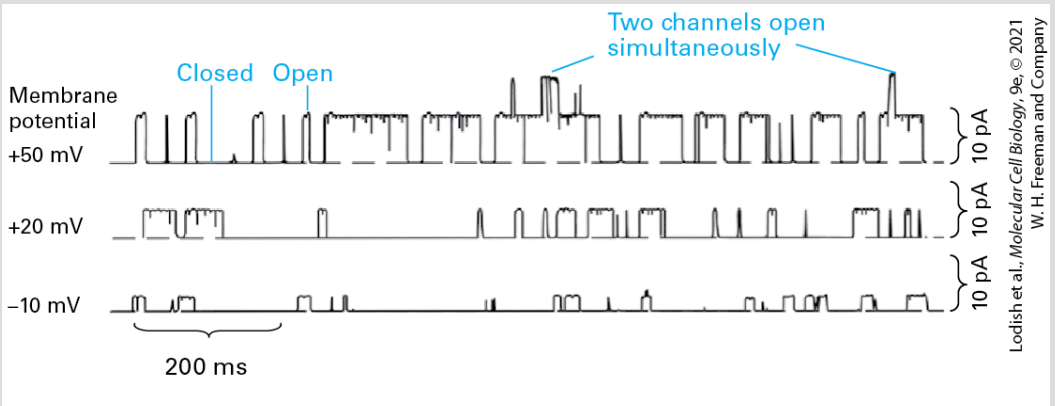
Voltage-Gated K^+ Channels

The repolarization of the membrane that occurs during the refractory period is due largely to opening of voltage-gated K^+ channels. The subsequent increased efflux of K^+ from the cytosol removes the excess positive charges from the cytosolic face of the plasma membrane, thereby restoring the inside-negative resting potential. For a brief instant, the membrane actually becomes hyperpolarized; at the peak of this **hyperpolarization**, the potential approaches E_K, which is more negative than the resting potential (see Figure 23-2).

Opening of the voltage-gated K^+ channels is induced by the large depolarization resulting from the action potential. Unlike voltage-gated Na^+ channels, most types of voltage-gated K^+ channels remain open as long as the membrane is depolarized and close only when the membrane potential has returned to an inside-negative value. Because the voltage-gated K^+ channels open slightly after the initial depolarization, at the height of the action potential, they sometimes are called *delayed rectifier K^+ channels*. Eventually all the voltage-gated K^+ and Na^+ channels return to their closed resting states. The only open channels in this baseline condition are the nongated or background K^+ channels, which lack voltage-sensitive transmembrane segments, and which contribute to the resting membrane potential of −60 to −70 mV (see Figure 23-9a).

The patch-clamp tracings in Figure 23-11 reveal the essential properties of voltage-gated K^+ channels (see Figure 11-22 for a

description of patch clamping). In this experiment, small segments of a neuronal plasma membrane were held (clamped) at different voltages and the flux of electric charges through the patch due to flow of K^+ ions through open K^+ channels was measured. At the modest depolarizing voltage of −10 mV, the channels in the membrane patch open infrequently and remain open for only a few milliseconds, as shown, respectively, by the number and width of the upward steps on the tracings. Further, the ion flux through them is rather small, as measured by the electric current passing through each open channel (the height of the steps). Depolarizing the membrane further to +20 mV causes these channels to open about twice as frequently; also, more K^+ ions move through each open channel (the height of the steps is greater) because the force driving cytosolic K^+ ions outward is greater at a membrane potential of +20 mV than at −10 mV. Depolarizing the membrane further to +50 mV, the value at the peak of an action potential, causes opening of more K^+ channels and also increases the flux of K^+ ions through them. Thus by opening during the peak of the action potential, these K^+ channels permit the outward movement of K^+ ions and repolarization of the membrane potential while the voltage-gated Na^+ channels are being closed and inactivated.



EXPERIMENTAL FIGURE 23-11 Probability of channel opening and current flux through individual voltage-gated K^+ channels increases with the extent of membrane depolarization. These patch-clamp tracings were obtained from patches of neuronal plasma membrane clamped at three different potentials, +50, +20, and −10 mV. The upward deviations in the current indicate the opening of K^+ channels and movement of K^+ ions outward (cytosolic to exoplasmic face) across the membrane. Increasing the membrane depolarization (i.e., the clamping voltage) from −10 mV to +50 mV increases the probability a channel will open, the time it stays open, and the amount of electric current (numbers of ions) that passes through it. pA = picoamperes.

[Data from B. S. Pallotta, K. L. Magleby, and J. N. Barrett, 1981, *Nature* **293**:471; as modified by B. Hille, 1992, *Ion Channels of Excitable Membranes,* 2d ed., Sinauer, p. 122.]

More than 100 voltage-gated K^+ channel proteins have been identified in humans and other vertebrates. As we discuss later, all these channel proteins have a similar overall structure, but they exhibit different voltage dependencies, conductivities, channel kinetics, and other functional properties.

Given the fundamental role that voltage-gated Na^+ and K^+ channels play in determining action potential firing, it is not

surprising that mutations in these channels have been identified in inherited, monogenic human epilepsies. Epilepsies are seizure disorders that affect about 1 percent of the population and that result from excessive synchronized neuronal activity in the brain. While epilepsy can arise from a variety of causes, including abnormal brain development, brain injury, and drug and alcohol abuse, some forms of epilepsy are caused by mutations in genes that encode ion channels. These diseases are called *channelopathies*. Human genetic studies have identified specific missense mutations in the Nav1.1 voltage-gated Na^+ channel that cause generalized epilepsy with febrile seizures, while loss-of-function mutations in the same channel cause Dravet's syndrome, also known as *severe myoclonic epilepsy of infancy*. A variety of mutations in the Kv7.2 and Kv7.3 voltage-gated K^+ channels cause another form of epilepsy called *benign familial neonatal convulsions*. Mutations in other voltage-gated Na^+ and K^+ channels have been identified in other forms of epilepsy and cause neuronal hyperexcitability in a variety of ways. For example, altering Na^+ channel inactivation or blocking K^+ channel–dependent repolarization of neurons both prolong the duration of action potentials, whereas decreasing the ratio of inhibitory to excitatory inputs onto neurons can lower the threshold for triggering action potentials. ■

Action Potentials Are Propagated Unidirectionally Without Diminution

An action potential begins with changes that occur in a small patch of the axonal plasma membrane near the cell body. At the peak of the action potential, passive spread of the membrane depolarization is sufficient to depolarize a neighboring segment of membrane. This causes a few voltage-gated Na^+ channels in this region to open, thereby increasing the extent of depolarization in the region and causing an explosive opening of more Na^+ channels and generation of an action potential. This depolarization soon triggers opening of voltage-gated K^+ channels and restoration of the resting potential. The action potential thus spreads as a traveling wave away from its initial site without diminution.

As noted earlier, during the refractory period voltage-gated Na^+ channels are inactivated for several milliseconds. Such refractory channels cannot conduct ion movements and cannot open during this period even if the membrane is depolarized owing to passive spread. As illustrated in Figure 23-12, the inability of Na^+ channels to reopen during the refractory period ensures that action potentials are propagated only in one direction, from the initial axon segment where they originate to the axon termini. Because the Na^+ channels upstream of the location of the action potential are still inactivated, they cannot be reopened by the small depolarization caused by passive spread. In contrast, the Na^+ channels downstream of the action potential begin to open. The refractory period of the Na^+ channels also limits the number of action potentials that a neuron

can conduct per second. This is important, since it is the frequency of action potentials that carries the information.

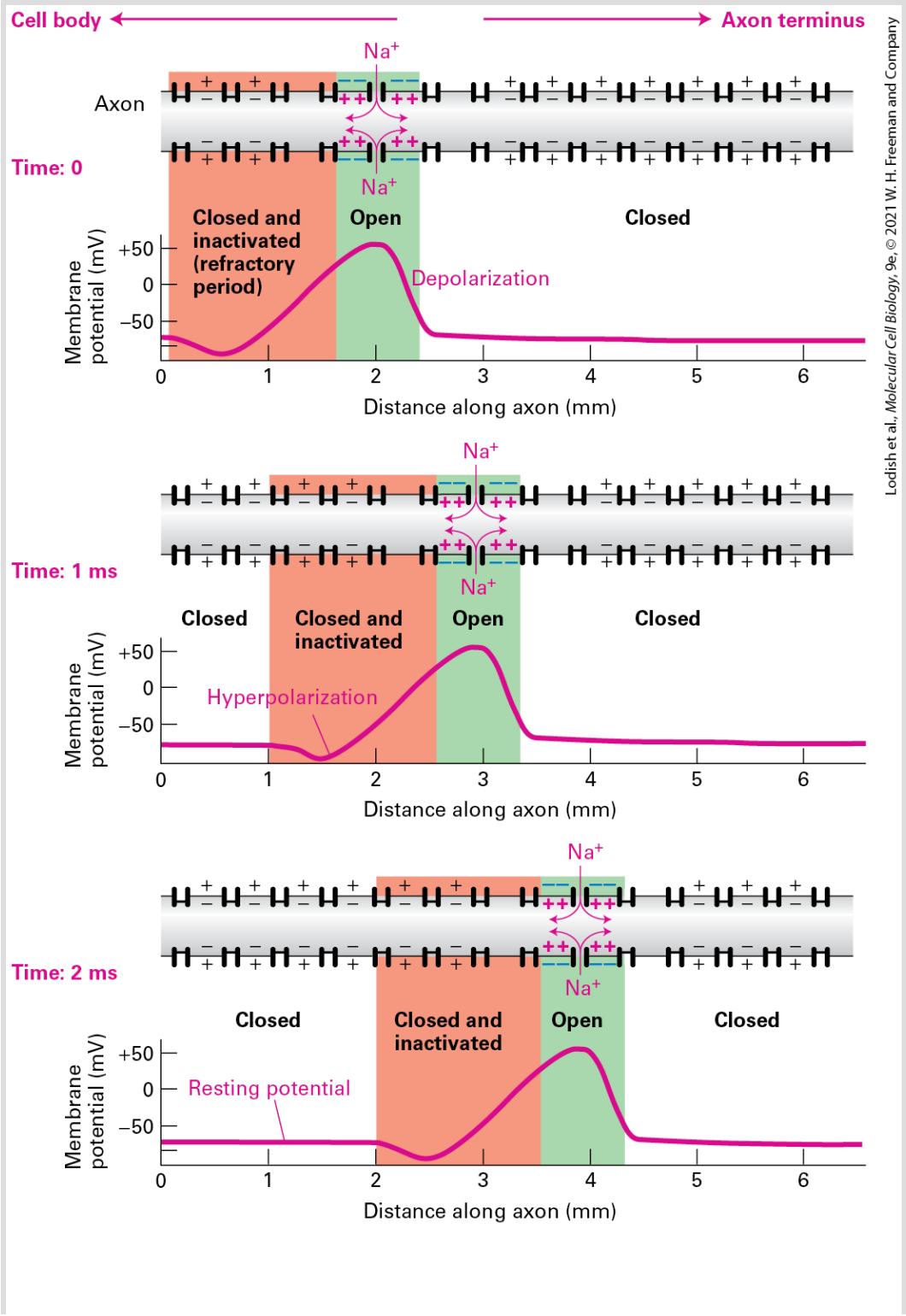


FIGURE 23-12 Unidirectional conduction of an action potential is due to the transient inactivation of voltage-gated Na^+ channels. At time 0, an action potential (pink line) is at the 2 mm position on the axon; the Na^+ channels at this position are open (green shading), and Na^+ ions are flowing inward. The excess Na^+ ions diffuse in both directions along the inside of the membrane, passively spreading the depolarization in both directions (curved pink arrows). Because the Na^+ channels at the 1 mm position are still inactivated (red shading), they cannot yet be reopened by the small depolarization caused by passive spread; the Na^+ channels at the downstream 3 mm position, in contrast, begin to open. Each region of the membrane is refractory (inactive) for a few milliseconds after an action potential has passed. Thus the depolarization at the 2 mm site at time 0 triggers action potentials only downstream; at 1 ms an action potential is passing the 3 mm position, and at 2 ms an action potential is passing the 4 mm position.

The depolarization of the membrane during an action potential results from movement of just a small number of Na^+ ions into a neuron and does not significantly affect the intracellular Na^+ concentration. While the flow of Na^+ and K^+ ions alters membrane potential dramatically as it is depolarized, hyperpolarized, and repolarized during an action potential cycle, it is important to note that the exchange of these ions across the membrane is small compared to the overall numbers of Na^+ and K^+ ions in the cytosol and extracellular space.

All Voltage-Gated Ion Channels Have Similar Structures

Having explained how the action potential is dependent on regulated opening and closing of voltage-gated channels, we turn to a molecular dissection of these remarkable proteins. After describing the basic structure of these channels, we will focus on three questions:

- How do these proteins sense changes in membrane potential?
- How is this change transduced into opening of the channel?
- What causes these channels to become inactivated shortly after opening?

The initial breakthrough in understanding voltage-gated ion channels came from the analysis of fruit flies (*Drosophila melanogaster*) carrying the *shaker* mutation. These flies shake vigorously under ether anesthesia, reflecting a loss of motor control and a defect in certain motor neurons that have an abnormally prolonged action potential. Researchers suspected that the *shaker* mutation caused a defect in channel function. The cloning of the gene involved confirmed that the defective protein was a voltage-gated K^+ channel. The *shaker* mutation prevents the mutant channel from opening normally immediately upon depolarization. To establish that the wild-type *shaker* gene encoded a K^+ channel, cloned wild-type *shaker* cDNA was used as a template to produce *shaker* mRNA in a cell-free system. Expression of this mRNA in frog oocytes and patch-clamp measurements on the newly synthesized channel protein showed that its functional properties were identical with those of the voltage-gated K^+ channel in the neuronal

membrane, demonstrating conclusively that the *shaker* gene encodes this K^+-channel protein.

The Shaker K^+ channel and most other voltage-gated K^+ channels that have been identified are tetrameric proteins composed of four identical subunits arranged in the membrane around a central pore. Each subunit is constructed of six membrane-spanning α helices, designated *S1–S6*, and a P segment (Figure 23-13a). The S5 and S6 helices and the P segment are structurally and functionally homologous to those in the voltage-independent background K^+ channel discussed earlier (see Figure 11-20); the S5 and S6 helices form the lining of the K^+ selectivity filter through which the ion travels. The S1–S4 helices form a rigid complex that functions as a voltage sensor (with positively charged side chains in S4 acting as the primary sensor). The N-terminal ball extending into the cytosol from S1 is the channel-inactivating segment.

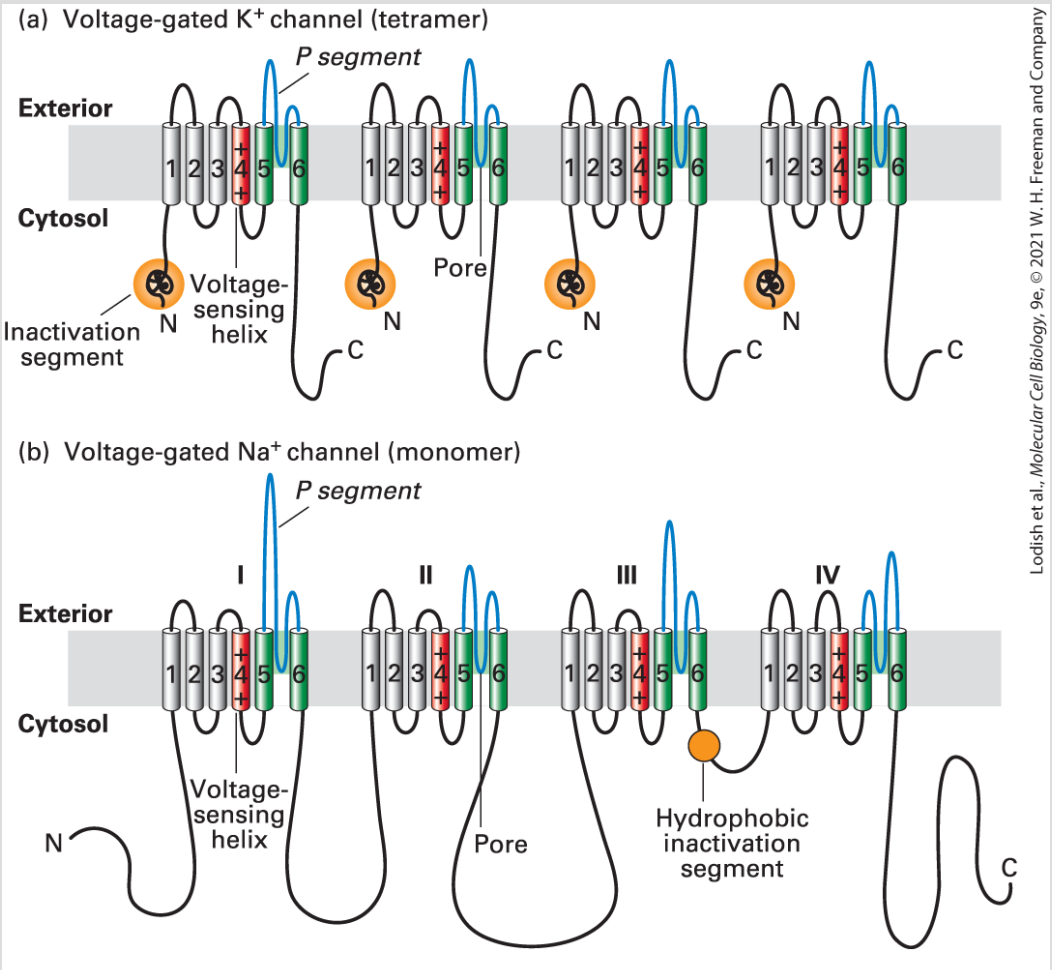


FIGURE 23-13 Schematic depictions of the secondary structures of voltage-gated K^+ and Na^+ channels. (a) Voltage-gated K^+ channels are tetramers composed of four identical subunits, each containing 600–700 amino acids, and six membrane-spanning α helices, S1–S6. The N-terminus of each subunit, located in the cytosol and labeled *N*, forms a globular domain (orange ball) essential for inactivation of the open channel. The S5 and S6 helices (green) and the P segment (blue) are homologous to those in voltage-independent background K^+ channels, but each subunit contains four additional transmembrane α helices. One of these, S4 (red), is the primary voltage-sensing α helix and is assisted in this role by forming a stable complex with helices S1–S3. See C. Miller and H. Hughes, 1992, *Curr. Biol.* **2**:573; and H. P. Larsson et al., 1996, *Neuron* **16**:387. (b) Voltage-gated Na^+ channels are monomers containing 1800–2000 amino acids organized into four transmembrane domains (I–IV) that are similar to the subunits in voltage-gated K^+ channels. The single hydrophobic channel-inactivating segment (orange ball) is located in

the cytosol between domains III and IV. Voltage-gated Ca^{2+} channels have a similar overall structure. Most voltage-gated ion channels also contain regulatory (β) subunits, which are not depicted here. See W. A. Catterall, 2001, *Nature* **409**:988.

Voltage-gated Na^+ channels are monomeric proteins organized into four homologous domains, I–IV (Figure 23-13b). Each of these domains is similar to a subunit of a voltage-gated K^+ channel. However, in contrast to voltage-gated K^+ channels, which have four channel-inactivating segments, the monomeric voltage-gated channels have a single channel-inactivating segment. Except for this minor structural difference and their varying ion permeabilities, all voltage-gated ion channels are thought to function in a similar manner and to have evolved from a monomeric ancestral channel protein that contained six transmembrane α helices. In the next section, we will focus on the structures of the voltage-gated K^+ channels and briefly discuss the more recently solved Na^+ channel structure.

Voltage-Sensing S4 α Helices Move in Response to Membrane Depolarization

Over the past 20 years, determination of the crystal structures for bacterial and Shaker potassium channels and other channels has greatly advanced our understanding of channel-protein biochemistry and function. More recently, powerful developments

in cryoelectron microscopy (see Chapter 11) have enabled the determination of the structures of many additional integral membrane proteins, including voltage-gated Na^+ and K^+ channels, in their native lipid environment.

The structures of the channels reveal remarkable arrangements of the voltage-sensing domains and suggest how parts of the protein move in order to open the channel. As already noted, the K^+-channel tetramer, like the Na^+-channel monomer, has a pore whose walls are formed by helices S5 and S6 (Figure 23-14). Outside that core structure, four arms (or "paddles"), each containing helices S1–S4, protrude into the surrounding membrane and also interact with the outer sides of the S5 and S6 helices; these are the voltage sensors, and they are in minimal contact with the pore. Sensitive electrical measurements suggested that the opening of a voltage-gated Na^+ or K^+ channel is accompanied by the movement of positive charges from the cytosolic to the exoplasmic surface of the membrane. The moving parts of the protein are the rigid complexes composed of helices S1–S4; S4 accounts for much of the positive charge and is the primary voltage sensor, with a positively charged lysine or arginine every third or fourth residue (Figure 23-14d). Arginines in S4 have been measured moving as much as 1.5 nm as the channel opens, which can be compared with the ~5 nm thickness of the membrane or the 1.2 nm diameter of the α helix itself; these significant conformational changes underlie the opening of the channel pore.

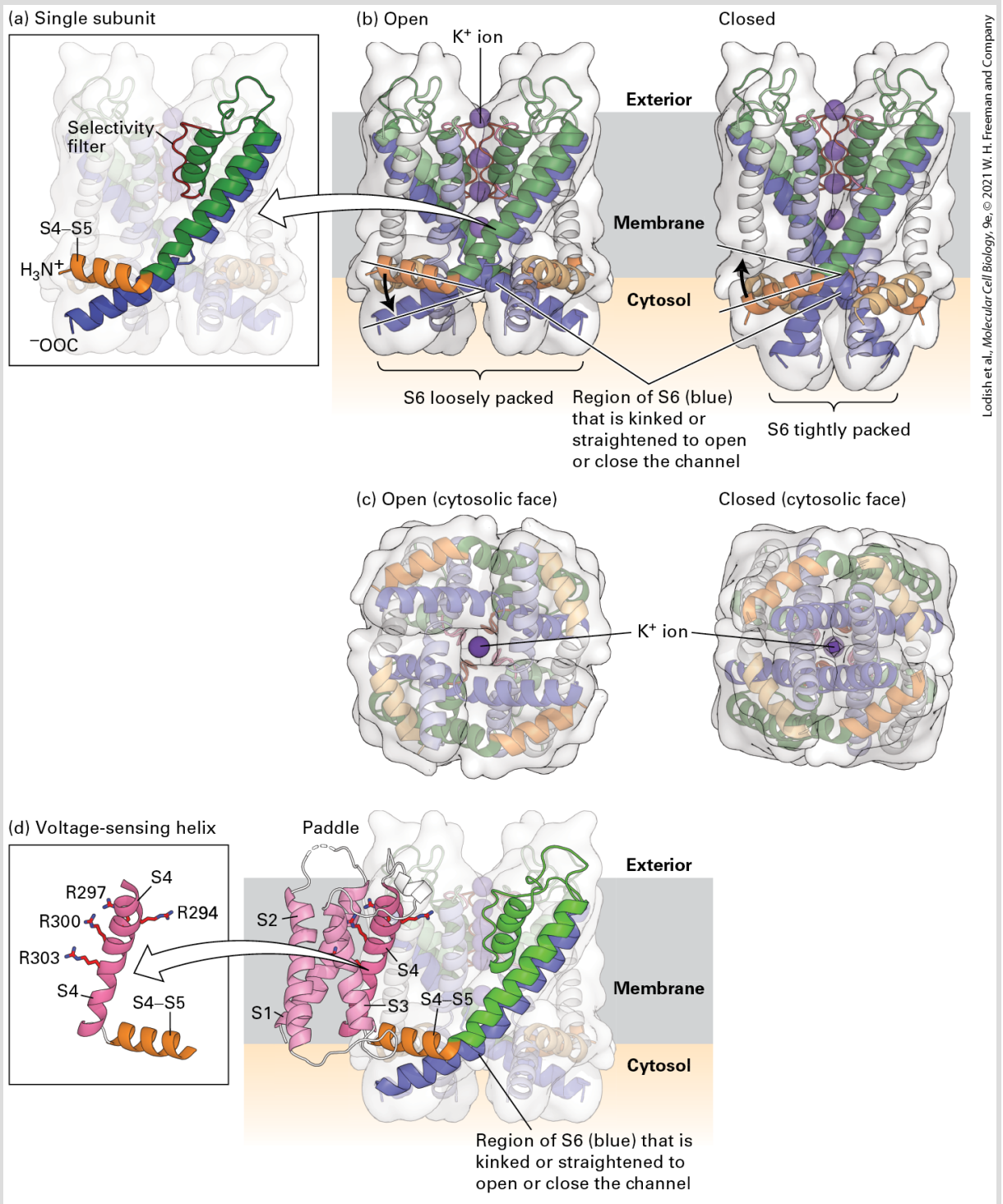


FIGURE 23-14 Molecular structure of a voltage-sensitive K^+ channel. Models of the potassium channel single subunit (a) and tetramer (b) as viewed from the side, in open and closed states. The four green (S5) and four blue (S6) α helices span the membrane, with the interior of the cell at the bottom and exterior at the top. Note how the helices are tightly

packed at the bottom in the closed conformation, so that the K^+ ion cannot pass through. [Compare the distances between S5 helices as shown by the curly brackets below (a) and (b).] The S4–S5 linker (orange), located in the cytoplasm, connects the S4 helix (not shown) to the S5 helix. For clarity, helices S1 through S4 have been omitted from the model; they would normally be attached to the end of the S4–S5 linker and protrude from the molecule. (c) Ribbon diagrams of the open and closed states of the channel as viewed from the cytoplasmic face of the membrane. In the open state, but not in the closed, potassium ions (dark purple) can pass through the pore. (d) Three-dimensional structure of the voltage-sensing "paddles" comprising helixes S1–S4, with the four voltage-sensing arginine (R) residues in S4. These paddles move from near the interior to the exterior of the membrane in response to depolarization. Since each one is attached to an S4–S5 linker, each linker and its attached S5 helix is moved, in turn moving S6 helices, which opens the pore. Note that as shown in (b), the linker between S4 and S5 is pointed upward toward the exoplasmic (exterior) surface in the open channel, pulled upward by the outward movement of the S1–S4 paddles; in contrast the S4–S5 linker is pointing downward in the closed channel when the S1–S4 paddles are nearer the cytosolic surface.

[Data from X. Chen et al., 2010, *Proc. Nat'l Acad. Sci. USA* **107**:11352, PDB ID 3lut; and Y. Zhou et al., 2001, *Nature* **414**:43–48, PDB ID based on 1k4c.]

In the resting state, the positive charges on the S1–S4 complexes (the paddles) are attracted to the negative charges on the cytosolic face of the membrane. In the depolarized membrane, these same positive charges become attracted to the negative charges on the exoplasmic (outer) surface of the membrane, causing the S1–S4 paddles to move partly across the membrane from the cytosolic to the exoplasmic surface. This movement is depicted schematically for the voltage-gated Na^+ channel in Figure 23-10 and triggers a conformational change in the protein that opens the channel.

The most unusual aspect of the voltage-sensitive channel structures is the presence of charged groups, such as arginines, in contact with lipid. The location of the voltage sensor helps to explain earlier experiments in which a non-voltage-sensitive channel was converted into a voltage-sensing channel by adding voltage-sensing domains. Such a result would seem unlikely if the voltage sensors had to be deeply embedded in the core structure.

Studies with mutant Shaker K^+ channels support the importance of the S4 helix in voltage sensing. When one or more arginine or lysine residues in the S4 helix of the Shaker K^+ channel were replaced with neutral or acidic residues, fewer positive charges than normal moved across the membrane in response to a membrane depolarization, indicating that arginine and lysine residues in the S4 helix do indeed move across the membrane and are necessary for channel function. The structure of the open form of a mammalian voltage-gated K^+ channel has been contrasted with the closed structure of another K^+ channel family member. The results suggest a model for the opening and closing of the channel in response to movements of the voltage sensors across the membrane (see Figure 23-14a, b). In the model, the voltage sensors, composed of helices S1–S4, move in response to voltage and exert a torque on a linker helix that connects S4 to S5:

- In the open-channel conformation, the position of the S4–S5 linker forces the S6 helix to form a kink near the cytosolic surface (blue in Figure 23-14a), and the pore inside, near the

cytosolic surface, is open. The pore's 1.2 nm diameter is sufficient to accommodate hydrated K^+ ions (see Figure 23-14c).

- When the cell membrane is repolarized and the voltage sensor moves toward the cytosolic membrane surface, the S4–S5 linkers (orange in Figure 23-14b) are twisted down, toward the inside of the cell. The S6 helices are consequently straightened, squeezing the bottom of the channel closed. Thus the gate is composed of the cytosol-facing ends of the S5 and S6 helices, where the pore is narrowest.

Although voltage-gated K^+ and voltage-gated Na^+ channels share similar voltage sensor and pore structures, the structure of their ion selectivity filters and the way that they conduct ions differ significantly (as also discussed in Chapter 11). The selectivity filter of the Na^+ channel is much larger than that of the K^+ channel, even though the diameter of a Na^+ ion (0.102 nm) is smaller than that of a K^+ ion (0.138 nm). The pore of K^+ channels contains conserved amino acids that form a lining of carbonyl oxygen atoms (see Figure 11-21). As K^+ ions enter the pore, these oxygen atoms replace its waters of hydration; the smaller Na^+ ions would be too small to interact with the backbone carbonyls of the K^+ channel pore. In contrast, Na^+ ions pass through the Na^+ channel pore as water-hydrated ions. The pore of the Na^+ channel is lined by conserved, negatively charged amino acids and is large enough to fit a single, water-hydrated Na^+ ion, with the positively charged Na^+ ion interacting with the negatively charged pore residues through its

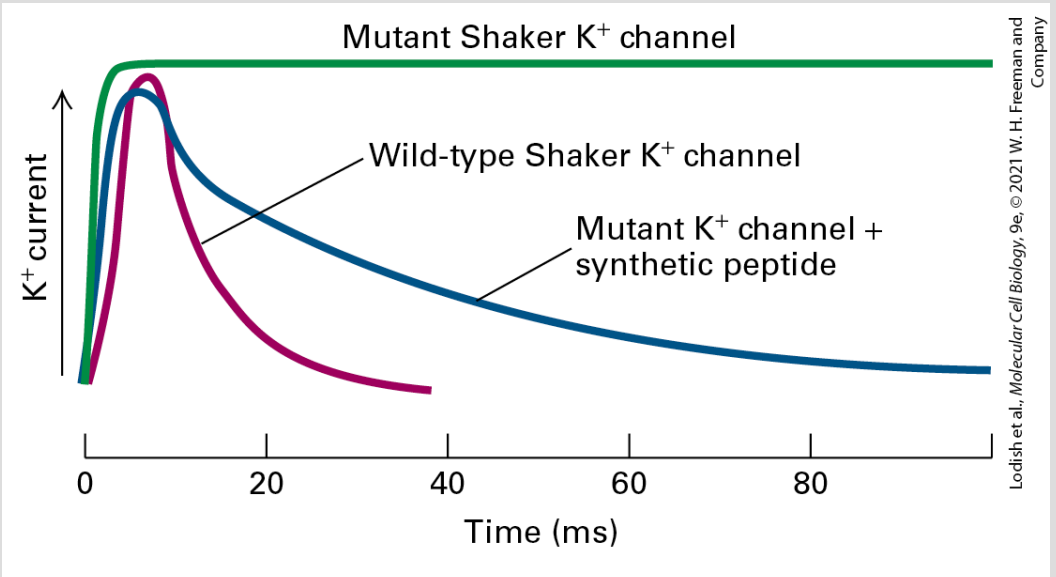
inner shell of bound water molecules. The hydrated K^+ ion is too large to fit through this pore.

The topical anesthetic lidocaine, which is commonly used to reduce pain during dental procedures or during minor surgeries (e.g., to suture cuts), works by blocking the flow of Na^+ ions through the voltage-gated Na^+ channel. Lidocaine binds to amino acid residues that line the channel pore and prevents the influx of Na^+ and thus the generation of an action potential. The binding sites for lidocaine are only available in the open state of the channel, and binding of lidocaine appears to lock the channel in the open but occluded state. ■

Movement of the Channel-Inactivating Segment into the Open Pore Blocks Ion Flow

An important characteristic of most voltage-gated channels is inactivation; that is, soon after opening they close spontaneously, forming an inactive channel that will not reopen until the membrane is repolarized. In the resting state, the globular balls at the N-termini of the four subunits in a voltage-gated K^+ channel are free in the cytosol (see Figure 23-13). Several milliseconds after the channel is opened by depolarization, one ball moves through an opening between two of the subunits and binds in a hydrophobic

pocket in the pore's central cavity, blocking the flow of K^+ ions (see Figure 23-10). After a few milliseconds, the ball is displaced from the pore, and the protein reverts to the closed, resting state. The ball-and-chain domains in K^+ channels are functionally equivalent to the channel-inactivating segment in Na^+ channels. The experimental results shown in Figure 23-15 demonstrate that inactivation of K^+ channels depends on the ball domains, occurs after channel opening, and does not require the ball domains to be covalently linked to the channel protein.



EXPERIMENTAL FIGURE 23-15 Experiments with a mutant K^+ channel lacking the N-terminal globular domains support the ball-and-chain inactivation model. The wild-type Shaker K^+ channel and a mutant form lacking the amino acids composing the N-terminal ball were expressed in *Xenopus* oocytes. The activity of the channels was monitored by the patch-clamp technique. When patches were depolarized from 0 to +30 mV, the wild-type channel opened for ~5 ms and then closed (red curve). The mutant channel opened normally but could not close (green curve). When a chemically synthesized ball peptide was added to the cytosolic face of the patch, the

mutant channel opened normally and then closed (blue curve). This demonstrated that the added peptide inactivated the channel after it opened and that the ball does not have to be tethered to the protein in order to function.

[Data from W. N. Zagotta, T. Hoshi, and R. W. Aldrich, 1990, *Science* **250**:568.]

The single channel-inactivating segment in voltage-gated Na^+ channels contains a conserved hydrophobic motif composed of isoleucine, phenylalanine, methionine, and threonine (see Figure 23-13b). Like the longer ball-and-chain domain in K^+ channels, this segment folds into and blocks the Na^+-conducting pore until the membrane is repolarized.

Myelination Increases the Velocity of Impulse Conduction

As we have seen, action potentials are triggered by the passage of ions through Na^+ and K^+ channels in the axon membrane. Action potentials can move down an unmyelinated axon without diminution at speeds up to 1 meter per second. Even such fast speeds are insufficient to permit the complex movements typical of animals. In adult humans, for instance, the cell bodies of motor neurons innervating leg muscles are located in the spinal cord, and the axons are about a meter in length. The coordinated muscle contractions required for walking, running, and similar movements would be impossible if it took 1 second for an action potential to

move from the spinal cord down the axon of a motor neuron to a leg muscle. The solution is to wrap cells in insulation that increases the rate of movement of an action potential. The insulation is called a *myelin sheath* (see Figure 23-1d). The presence of a myelin sheath around an axon increases the velocity of impulse conduction to 10–100 meters per second. As a result, in a typical human motor neuron, an action potential can travel the length of a 1-meter-long axon and stimulate a muscle to contract within 0.01 seconds.

In nonmyelinated neurons, the conduction velocity of an action potential is roughly proportional to the diameter of the axon, because a thicker axon will have a greater number of ions that can diffuse. The human brain is packed with relatively small, myelinated neurons. If the neurons in the human brain were not myelinated, their axonal diameters would have to increase about 10,000-fold to achieve the same conduction velocities as myelinated neurons. Thus vertebrate brains, with their densely packed neurons, never could have evolved without myelin.

The vertebrate brain and spinal cord can be divided into gray and white matter (Figure 23-16), with the gray matter consisting of neuronal and glial cell bodies and unmyelinated processes, and the white matter consisting of myelin-ensheathed axon tracts, with the lipid-rich myelin imparting the white color. The difference in tissue density can also be detected using brain imaging techniques such as computerized tomography (CT) scanning and magnetic resonance

imaging (MRI; Figure 23-16b). These methods can be used to differentiate disease processes that affect myelin from those that affect neurons.

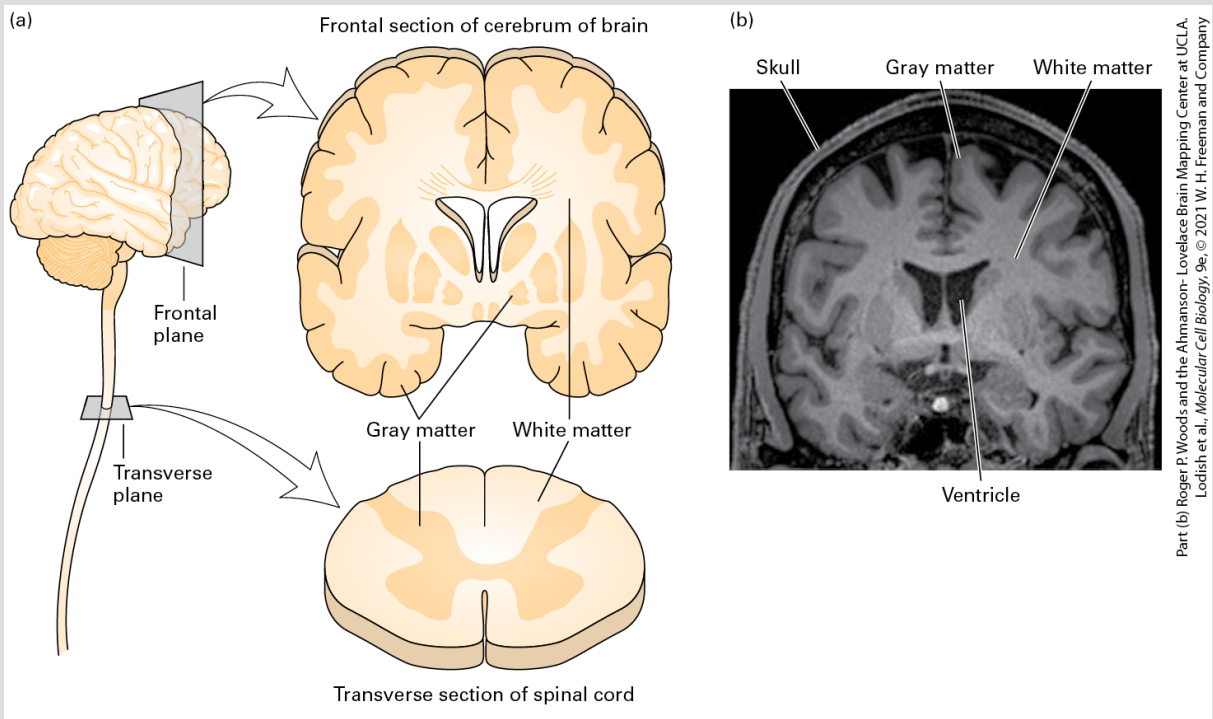


FIGURE 23-16 Gray and white matter in the central nervous system. The vertebrate brain and spinal cord can be divided into "gray matter," consisting of neuronal and glial cell bodies and unmyelinated processes, and "white matter" consisting of myelin-ensheathed axon tracts. The differences in density and lipid composition results in visible differences at the level of tissue histology, shown for the spinal cord and the cerebrum in (a). The differences in tissue density are also detected using imaging techniques such as magnetic resonance imaging (MRI) of the cerebrum (b).

Action Potentials "Jump" from Node to Node in Myelinated Axons

The myelin sheath surrounding an axon is formed from many glial cells. Each region of myelin formed by an individual glial cell is separated from the next region by an unmyelinated area of axonal membrane about 1 μm in length called the **node of Ranvier** (or simply, "node"; see Figure 23-1d). The axonal membrane is in direct contact with the extracellular fluid only at the nodes, and the myelin covering prevents any ion movement into or out of the axon except at the nodes. Moreover, voltage-gated Na^+ and Na^+/K^+ pumps, which maintain the ionic gradients in the axon, are present in high concentrations in the nodes, while K^+ channels are in high concentrations adjacent to the nodes.

As a consequence of this localization, the inward movement of Na^+ ions that generates the action potential can occur only at the myelin-free nodes (Figure 23-17). The excess cytosolic positive ions generated at a node during the membrane depolarization associated with Na^+ movement into the cytosol as part of an action potential spread passively through the axonal cytosol to the next node with very little loss or attenuation, since they cannot cross the myelinated axonal membrane. In this way, a depolarization at one node spreads rapidly to the next node and induces an action potential there, effectively permitting the action potential to jump from node to node, in a kind of transmission called **saltatory conduction**. This mode of transmission allows for both faster signaling and narrower axons seen in vertebrate animals.

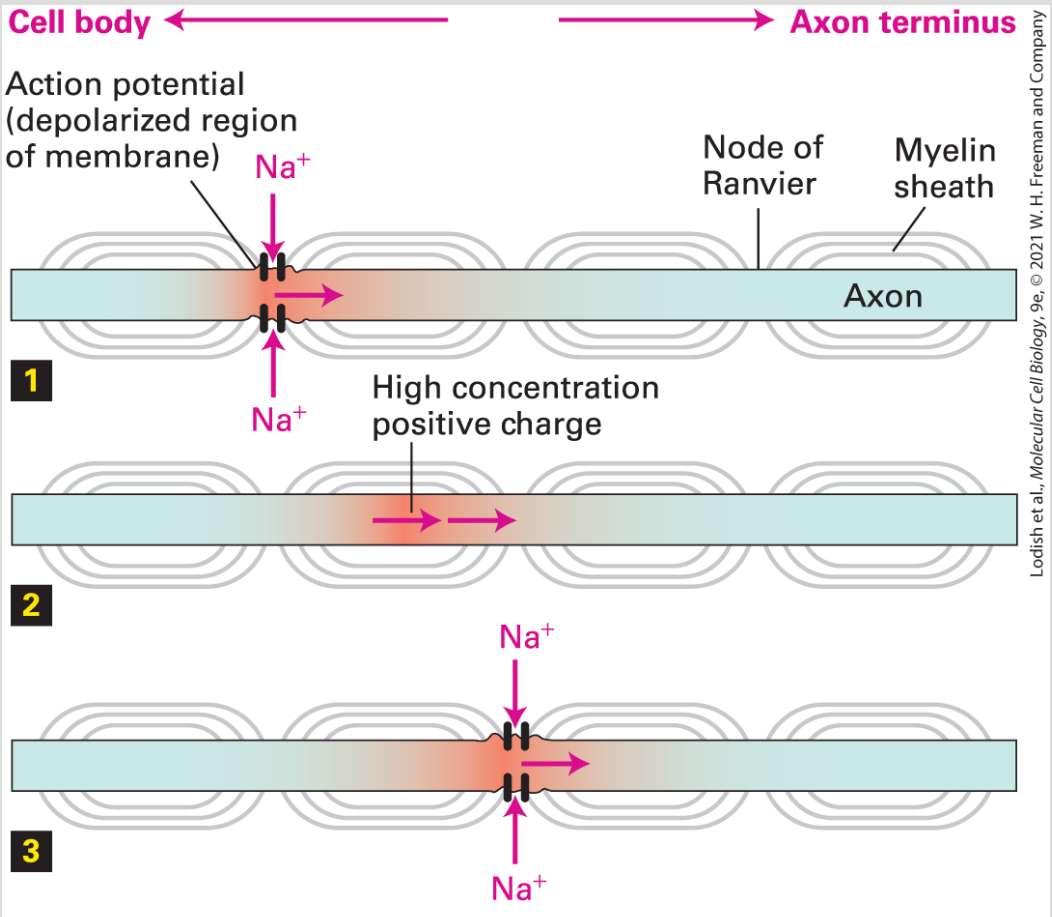


FIGURE 23-17 Conduction of action potentials in myelinated axons. Because the myelin layer renders the axon impermeable to ion movement across its membrane and because voltage-gated Na^+ channels are found only on axonal membrane at the nodes of Ranvier, the influx of Na^+ ions associated with an action potential can occur only at nodes. When an action potential is generated at one node (step 1), the excess positive ions in the cytosol, which cannot move outward across the sheath, diffuse rapidly down the axon, causing sufficient depolarization at the next node (step 2) to induce an action potential at that node (step 3). By this mechanism the action potential jumps from node to node along the axon.

Two Types of Glia Produce Myelin Sheaths

Myelin sheaths are produced by two types of glial cells present in the nervous system: oligodendrocytes make sheaths for the central nervous system (CNS), and Schwann cells make them for the peripheral nervous system (PNS) (Figure 23-18). For reasons that are not understood, not all axons are myelinated.

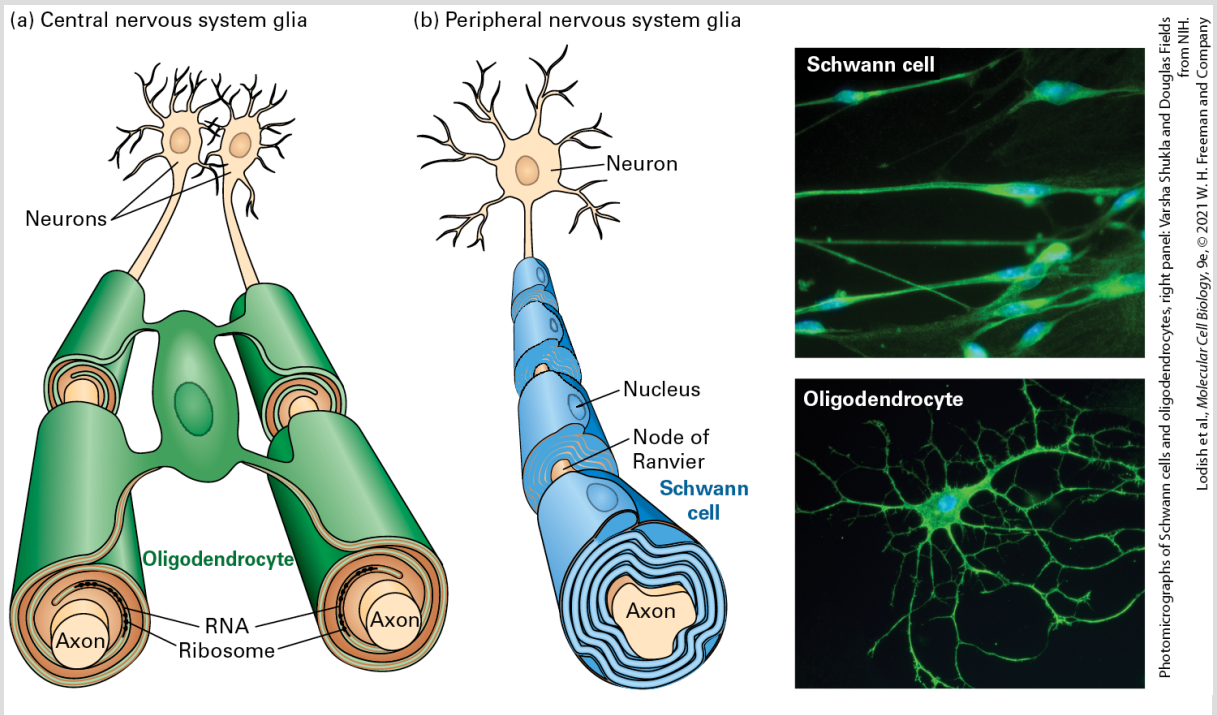


FIGURE 23-18 Two types of glial cells make myelin sheaths. (a) A single oligodendrocyte in the central nervous system can myelinate segments of multiple axons. (b) Each Schwann cell insulates a section of a single peripheral nervous system axon. See B. Stevens, 2003, *Curr. Biol.* **13**:R469, and D. L. Sherman and P. J. Brophy, 2005, *Nat. Rev. Neurosci*. **6**:683–690.

Oligodendrocytes

Oligodendrocytes form the spiral myelin sheath around axons of the central nervous system (Figure 23-18a). Each oligodendrocyte provides myelin sheaths to segments of multiple neurons. The major protein constituents are myelin basic protein (MBP) and proteolipid protein (PLP). MBP, a peripheral membrane protein found in both the central and peripheral nervous systems, has seven RNA-splicing variants that encode different forms of the protein. It is synthesized by ribosomes located in the growing myelin sheath, an example of specific transport of mRNAs to a distal cell region (Chapter 10). MBP mRNA undergoes microtubule-dependent transport to distal oligodendrocyte processes, where its local translation contributes to the formation of the myelin sheath.

Damage to proteins produced by oligodendrocytes underlies a prevalent human neurological disease, multiple sclerosis (MS). MS is usually characterized by spasms and weakness in one or more limbs, bladder dysfunction, local sensory losses, and visual disturbances. This disorder, the prototype *demyelinating disease*, is caused by patchy loss of myelin in areas of the brain and spinal cord. In MS patients, conduction of action potentials by the demyelinated neurons is slowed, and the Na^+ channels spread outward from the nodes, lowering their nodal concentration. The cause of the disease is not known but is believed to involve the body's production of auto-antibodies (antibodies that bind to normal body proteins) that react with and attack myelin. A mouse mutant, *shiverer,* has a deletion of much of the *MBP* gene, leading to tremors,

convulsions, and early death. Similarly, a human mutation (Pelizaeus–Merzbacher disease) and a mouse mutation (*jimpy*) in the gene coding for the other major protein of CNS myelin, PLP, cause loss of oligodendrocytes and inadequate myelination. A number of new drugs to treat MS, many targeting the immune system, have recently been approved by the FDA, and many more MS drugs are currently in clinical trials. ■

Schwann Cells

Schwann cells form myelin sheaths around the axons of neurons in the PNS. A Schwann cell myelin sheath is a remarkable spiral wrap (Figure 23-18b). A long axon can have as many as several hundred Schwann cells along its length, each contributing myelin insulation to an *internode* stretch of about 1–1.5 μm of axon. Mutations in mice that eliminate Schwann cells cause the death of most neurons.

In contrast to oligodendrocytes, each Schwann cell myelinates only one axon. The sheaths are composed of about 70 percent lipid (rich in cholesterol) and 30 percent protein. In the PNS, the principal protein constituent (~80 percent) of myelin is called *protein 0* (P_0), an integral membrane protein that has immunoglobulin (Ig) domains. MBP is also an abundant component. The extracellular Ig domains of P_0 bind together the surfaces of sequential wraps around the axon to compact the spiral of myelin sheath (Figure 23-19). Other proteins play this kind of role in the CNS.

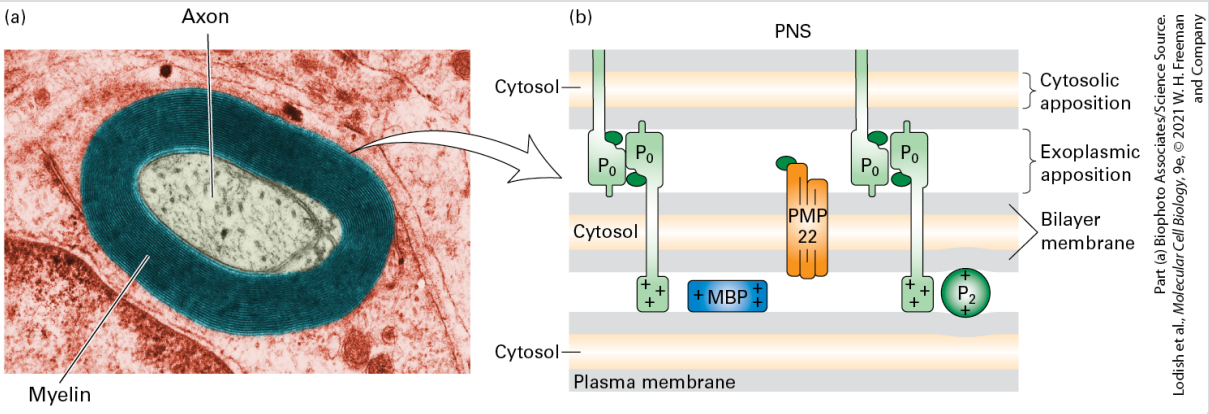


FIGURE 23-19 Formation and structure of a myelin sheath in the peripheral nervous system (PNS). (a) At high magnification the specialized spiral myelin membrane appears as a series of layers, or lamellae, of phospholipid bilayers wrapped around the axon. (b) Close-up cartoon of three layers of the myelin membrane spiral. The two most abundant integral myelin membrane proteins, P_0 and PMP22, are produced only by Schwann cells. The exoplasmic domains of P_0 proteins, on opposite membrane surfaces associate with one another, thereby zipping together the exoplasmic membrane surfaces in close apposition. Close apposition of the cytosolic faces of the membrane may result from binding of the cytosolic tail of each P_0 protein to phospholipids in the opposite membrane. PMP22 may also contribute to membrane compaction. Myelin basic protein (MBP), a cytosolic protein, remains between the closely apposed membranes as the cytosol is squeezed out. See L. Shapiro et al., 1996, *Neuron* **17**:435; and E. J. Arroyo and S. S. Scherer, 2000, *Histochem. Cell Biol.* **113**:1.

In humans, peripheral myelin, like CNS myelin, is a target of autoimmune disease, mainly involving the formation of antibodies against P_0. The Guillain-Barre syndrome (GBS), also known as *acute inflammatory demyelinating polyneuropathy,* is one such disease. GBS is the most common cause of rapid-onset paralysis, occurring at a frequency of one person out of 100,000. The cause is unknown, although it usually follows an acute infectious illness and is thought to involve an immune attack on the PNS. The common inherited

neurological disorder called *Charcot-Marie-Tooth disease type 1A,* which damages peripheral motor and sensory nerve function, is due to overexpression of the gene that encodes PMP22 protein, another constituent of peripheral nerve myelin. ■

Interactions between glia and neurons control the placement and spacing of myelin sheaths and the assembly of nerve-transmission machinery at the nodes of Ranvier. For example, a high concentration of voltage-gated Na^+ channels and Na^+/K^+ pumps, which is critical to the rise in membrane voltage during action potential generation, congregates at the nodes of Ranvier through interactions with cytoskeletal proteins. The activity of these cytoskeletal proteins is modulated by interactions between cell-adhesion molecules on glia and neurons. The regions adjacent to the myelin-free nodes are called *paranodes* and consist of specialized junctions between the myelin and the axonal membrane formed by interactions between cell-adhesion molecules described in the next paragraph. These junctions serve to restrict the movement of ion channels within the axonal membrane; they separate the nodes from a region called the *juxtaparanode,* which is enriched in voltage-gated K^+ channels. These K^+ channels mediate the return of the membrane voltage to baseline after an action potential, and the physical separation of the Na^+ and K^+ channels contributes to the efficient propagation of action potentials.

Light-Activated Ion Channels and Optogenetics

Unicellular flagellates such as *Chlamydomonas* (Figure 1-23b) respond rapidly to light through *phototaxis*, which guides them toward the light, or *photophobia*, which prevents them from moving toward the light. These rapid *photomotility* responses are mediated by light-activated cation channels that reside in a specialized photoreceptive organelle called the *eyespot*. The first light-activated channel to be cloned was from the green algae *Chlamydomonas reinhardtii*. This channel, called *channelrhodopsin*, consists of a seven-transmembrane protein covalently linked to a photo-isomerizable chromophore, all-*trans*-retinal (Figure 23-20a, b). This is the same chromophore that is used to detect photons in the eye, as discussed in Chapter 15 (see Figure 15-34). When all-*trans*-retinal absorbs a photon, it converts to 13-*cis*-retinal, which induces a conformational change in the channel protein, leading to the opening of a pore with a diameter of approximately 0.6 nm and the rapid influx of cations. Channelrhodopsins are nonselective cation channels that conduct H^+, Na^+, K^+, and Ca^{2+} ions. Many other light-activated cation channels have now been identified, and genetic engineering approaches have generated a series of designer channelrhodopsins that are activated or inactivated by specific wavelengths and intensities of light and selective for distinct cations. These advances have given rise to the exciting new field of **optogenetics**, in which channelrhodopsins are expressed in

electrically excitable cells, permitting the use of light to rapidly and selectively manipulate the membrane potential of the cell (Figure 23-20c).

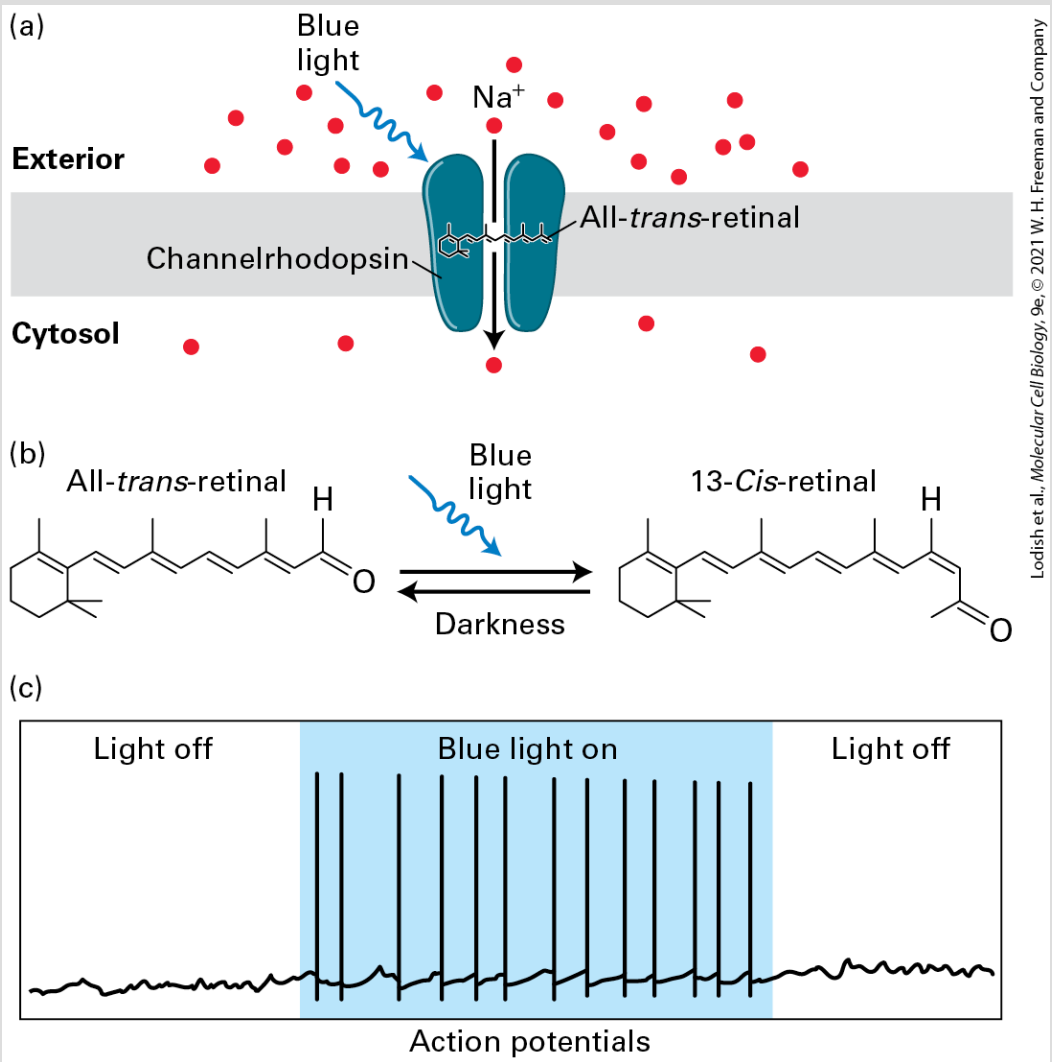


FIGURE 23-20 Channelrhodopsins and optogenetics: activating neurons with light. (a) Light-activated ion channels. Channelrhodopsins are light-activated cation channels. The channel has seven transmembrane domains and is covalently linked to the photo-isomerizable chromophore all-*trans*-retinal. (b) All-*trans*-retinal absorbs blue light (~470 nm) and changes to the 13-*cis*-retinal conformation. This leads to a conformational change in the channel, opening a pore through which cations can flow. Opening of channelrhodopsin, shown in

(a), leads to influx of Na^+ from the extracellular solution into the cell, resulting in rapid depolarization of excitable cells. When the light is removed, retinal returns to the all-*trans* conformation and the channel closes. This is shown in (c) where illumination of a neuron expressing channelrhodopsin with blue light triggers action potentials as long as the light is on. See J. Wong, O. J. Abilez, and E. Kuhl, 2012, *J. Mech. Phys. Solids* **60**:1158–1178.

Optogenetic approaches have revolutionized the study of neuronal circuits in the brain because they allow neuroscientists to directly test the effect of neural circuit activity on behavior, rather than to simply record neural activity and correlate it with behavior. A common experimental approach is to make transgenic mice in which channelrhodopsin is expressed under the control of a cell-type-specific promoter (described in Chapter 6) so that only a subset of neurons in the brain express the protein. A window is then made in the skull, and lasers are used to illuminate channelrhodopsin-expressing cells near the surface of the brain, or alternatively a fiber-optic cable is used to deliver light to channelrhodopsin-expressing cells in deeper brain regions. The light triggers action potentials in the neurons, activating their circuits. By examining the behavior of the mouse, one can link the circuit with a specific behavior.

As one example of this type of experiment, the identity of the neurons in the brain that mediate thirst was discovered using optogenetics. To do this, channelrhodopsin was expressed in a region of the hypothalamus called the *subfornical organ (SFO)* in

mice (Figure 23-21). In one mouse line, channelrhodopsin was expressed exclusively in excitatory neurons, and in the other mouse line, it was expressed in inhibitory neurons. Light was delivered to the SFO using a fiber-optic cable; when the excitatory neurons were depolarized with light, the mice exhibited thirst (seeking water and drinking intensely), but when the inhibitory neurons were activated, thirst was suppressed. Together, these experiments demonstrate that neurons in the SFO of the hypothalamus regulate thirst, with activation of excitatory neurons promoting thirst and activation of inhibitory neurons repressing thirst. Similar approaches have been used to uncover circuits underlying many other behaviors, including locomotion, feeding and overeating, anxiety, aggression, and other social behaviors. In addition, optogenetic approaches can be used with other methods, such as calcium imaging (see Chapter 4 and Section 23.3) to map out the anatomy of specific circuits in the brain. This is done by activating a specific neuron with light and then using imaging approaches to visualize activity (e.g., by monitoring calcium dynamics) in downstream neurons.

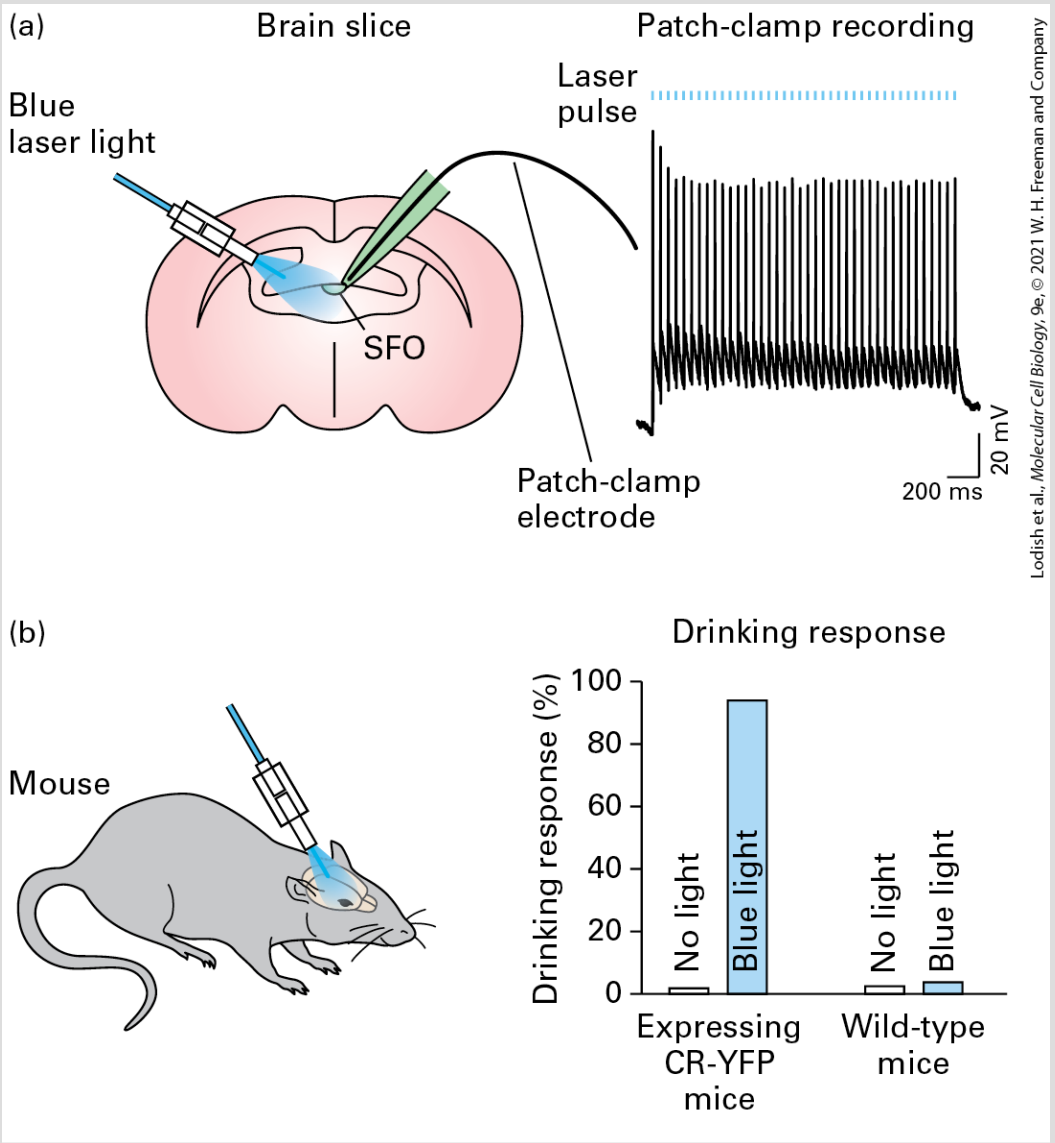


FIGURE 23-21 Using optogenetics to dissect neural circuits mediating thirst. (a) Channelrhodopsin tagged with a fluorescent protein (YFP) was expressed in excitatory neurons in the subfornical organ (SFO) of the hypothalamus using an excitatory neuron–specific promoter. Neurons expressing channelrhodopsin-YFP were confirmed to be excitable by illumination with blue light in acute hypothalamic slices. (b) When excitatory neurons in the SFO of living mice are activated by blue light, the mice seek water and drink large volumes even if they are well hydrated.

[Data from Y. Oka, M. Ye, and C. S. Zuker, 2015, *Nature* **520**:349–352.]

KEY CONCEPTS OF SECTION 23.2

Voltage-Gated Ion Channels and the Propagation of Action Potentials

- Action potentials are sudden membrane depolarizations followed by rapid repolarization.
- An action potential results from the sequential and rapid opening and closing of voltage-gated Na^+ and K^+ channels (see Figure 23-9) in the plasma membrane of neurons and muscle cells.
- The excess cytosolic cations associated with an action potential generated at one point on an axon spread passively to the adjacent segment, triggering opening of voltage-gated Na^+ channels in the vicinity and thus propagation of the action potential along the axon.
- As the action potential reaches its peak, opening of voltage-gated K^+ channels permits efflux of K^+ ions, which repolarizes and then hyperpolarizes the membrane. As these channels close, the membrane returns to its resting potential, followed by a refractory period of several milliseconds (see Figures 23-2, 23-9, and 23-10).
- The action potential is propagated in only one direction, toward the axon termini, because of the absolute refractory period of the voltage-gated Na^+ channels and the brief hyperpolarization resulting from K^+ efflux (see Figure 23-12).
- Voltage-gated Na^+ channels are monomeric proteins containing four domains that are structurally and functionally similar to each of the subunits in the tetrameric voltage-gated K^+ channels. Each domain or subunit in voltage-gated cation channels contains six transmembrane α helices and a nonhelical P segment that forms the ion-selectivity pore (see Figure 23-13).
- Opening of voltage-gated channels results from movement of the positively charged S1−S4 paddles toward the extracellular side of the membrane in response to a depolarization of sufficient magnitude (see Figure 23-14).
- Closing and inactivation of voltage-gated cation channels result from movement of a cytosolic ball segment into the open pore (see Figure 23-10).
- While the voltage sensor and inactivation gate of voltage-gated K^+ channels and voltage-gated Na^+ channels are similar, the structure of the selectivity filter is different and provides specificity for the type of ion that is conducted through the channel.
- Myelination increases the rate of impulse conduction up to a hundredfold and permits the close packing of neurons characteristic of vertebrate brains.

- The central nervous system (CNS) can be divided into gray matter, composed of neuronal and glial cell bodies, and white matter, enriched in myelin-ensheathed axons (see Figure 23-16).
- In myelinated neurons, voltage-gated Na^+ channels are concentrated at the nodes of Ranvier. Depolarization at one node spreads rapidly with little attenuation to the next node, so that the action potential jumps from node to node (see Figure 23-17).
- Myelin sheaths are produced by glial cells that wrap themselves in spirals around neurons. Oligodendrocytes produce myelin for the CNS; Schwann cells, for the peripheral nervous system (PNS; see Figure 23-18).
- The field of optogenetics is revolutionizing the study of neural circuits. It involves the genetic expression of light-activated cation channels, called *channelrhodopsins*, in neurons and the use of light to specifically activate or inhibit that population of neurons (see Figure 23-20). In this way, neuroscientists can directly link specific neural circuits with specific behaviors.

23.3 Communication at Synapses

As we have discussed, electrical pulses transmit signals along neurons, but signals are transmitted between neurons and other excitable cells mainly by chemical signals. Synapses are the junctions where presynaptic neurons release these chemical signals, or neurotransmitters, which then act on postsynaptic target cells (see Figure 23-3). A target cell may be another neuron, a muscle, or a gland cell. Communication at chemical synapses usually goes in only one direction: pre- to postsynaptic cell.

Arrival of an action potential at an axon terminus in a presynaptic cell leads to opening of voltage-sensitive Ca^{2+} plasma-membrane channels and an influx of Ca^{2+}, causing a localized rise in the cytosolic Ca^{2+} concentration in the axon terminus. In turn, the rise in Ca^{2+} triggers fusion of small (40–50 nm) neurotransmitter-containing synaptic vesicles with the plasma membrane, releasing neurotransmitters into the synaptic cleft, the narrow space separating the presynaptic from the postsynaptic cell. The membrane of the postsynaptic cell is located within approximately 20 nm of the presynaptic membrane, reducing the distance the neurotransmitter must diffuse.

Neurotransmitters, small, water-soluble molecules such as glutamate (excitatory) or gamma-amino butyric acid (GABA, inhibitory), bind to receptors on the postsynaptic neuron that, in

turn, induce localized changes in the potential across its plasma membrane. If the membrane potential becomes less negative (i.e., becomes depolarized) an action potential will tend to be induced in the postsynaptic cell. Such synapses are called **excitatory synapses**, and in general involve the opening of Na^+ channels in the postsynaptic plasma membrane. In contrast, at an **inhibitory synapse**, binding of the neurotransmitter to a receptor on the postsynaptic cell causes hyperpolarization of the plasma membrane, that is, generation of a more inside-negative potential. Typically, hyperpolarization is the result of opening of Cl^- or K^+ channels in the postsynaptic plasma membrane, which tends to hinder generation of an action potential.

Neurotransmitter receptors fall into two broad classes: ligand-gated ion channels, also known as *ionotropic receptors*, which open immediately upon neurotransmitter binding, and metabotropic G protein–coupled receptors (GPCRs). Neurotransmitter binding to a GPCR induces the opening or closing of a separate ion-channel protein over a period of seconds to minutes. These so-called slow neurotransmitter receptors were discussed in Chapter 15 along with GPCRs that bind different types of ligands and modulate the activity of cytosolic proteins other than ion channels. In the central nervous system, *glutamate* and GABA bind primarily to ionotropic receptors to mediate excitation and inhibition, respectively, while neuromodulators like serotonin and dopamine usually bind to metabotropic receptors. In the peripheral nervous system, the main

neurotransmitters are acetylcholine and norepinephrine (also called *noradrenaline*), both of which are also expressed in the central nervous system.

The duration of the neurotransmitter signal depends on the amount of transmitter released by the presynaptic cell, which in turn depends on the amount of transmitter that had been stored as well as the frequency of action potentials arriving at the synapse. The duration of the signal also depends on how rapidly any unbound neurotransmitter is degraded in the synaptic cleft or transported back into the presynaptic cell. Presynaptic cell plasma membranes, as well as glia, contain transporter proteins that pump neurotransmitters across the plasma membrane back into the cell, thus keeping the extracellular concentrations of transmitter low.

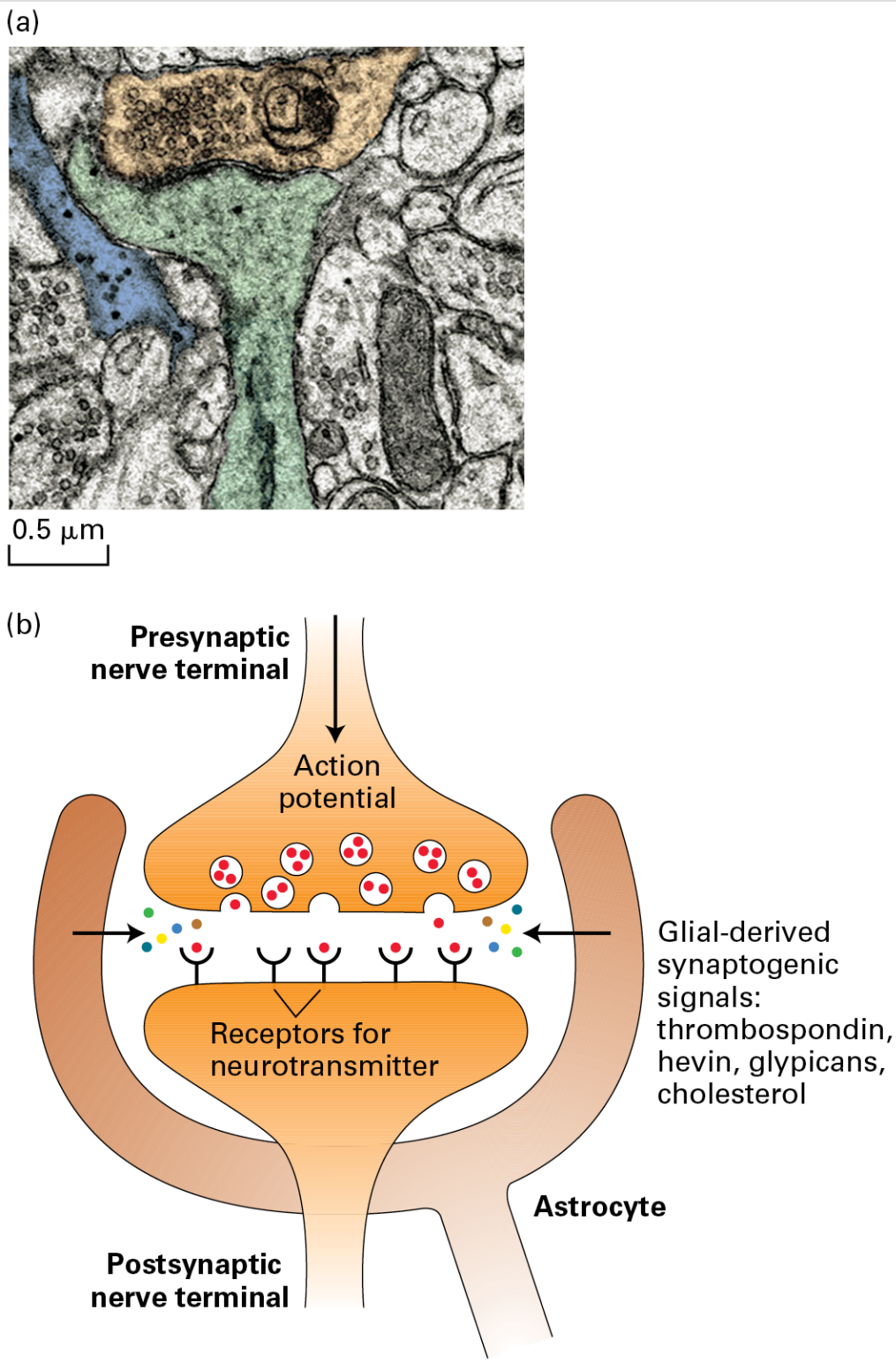
In this section, we focus first on how synapses form and how they control the regulated secretion of neurotransmitters in the context of the basic principles of vesicular trafficking outlined in Chapter 14. Next we look at the mechanisms that limit the duration of the synaptic signal, and how neurotransmitter signaling is received and interpreted by the postsynaptic cell.

Formation of Synapses Requires Assembly of Presynaptic and Postsynaptic Structures

Axons extend from the cell body during development, guided by signals from other cells along the way, so that the axon termini will reach the correct location (see Section 18.8). As axons grow, they come into contact with their potential target cells, such as dendrites of other neurons, where synapses then form. In the CNS, presynaptic release sites occur frequently all along an axon and are called *en passant* ("in passing") *synapses*; in contrast, motor neurons form synapses with muscle cells only at the axon termini.

Glial cells play an essential role in synapse formation. Neurons cultured in isolation will not form synapses very efficiently, but when glia are added, the rate of synapse formation increases substantially. Astrocytes and Schwann cells send protein signals to neurons to stimulate the formation of synapses and then help to preserve them. One such signal is thrombospondin (TSP), a component of the extracellular matrix; mice lacking two *thrombospondin* genes have only 70 percent of the normal number of synapses in their brains. Mutual communication between neurons and the glia that surround them is frequent and complex, making the signals and information they carry an area of active research. New imaging approaches have revealed that astrocytes form a multitude of small branches that intercalate with neurons and ensheath synapses (see Figure 23-6a). Due to the function of astrocytes in providing glial-derived factors to promote synapse formation, many neuroscientists now propose that the synapse should be considered a tripartite structure, composed not only of

presynaptic and postsynaptic partners but also of astrocytes (see Figure 23-22).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 23-22 Astrocytes and the tripartite synapse. (a) Many synapses are ensheathed by astroglial processes, as shown in this electron micrograph of a

synapse in the rodent hippocampus. The postsynaptic compartment (dendrite and dendritic spine) is highlighted in green, the presynaptic terminal in orange, and astroglial process in blue. (b) Astrocytes not only ensheath the synapse but also secrete a number of factors that promote correct synapse formation. These include thrombospondin, hevin, glypicans, and cholesterol. Indicative of the importance of astrocytes in neuronal synapse formation, when neurons are grown in cell culture, they require astrocytes for proper synapse formation and development.

[Part (a) republished with permission from Elsevier, from J. Bourne and K. M. Harris, 2007, "Do Thin Spines Learn to Be Mushroom Spines That Remember?" *Curr. Opin. Neurobiol*. **17**(3):381–386; permission conveyed through Copyright Clearance Center, Inc.]

At the site of a synapse, the presynaptic neuron has hundreds to thousands of synaptic vesicles, some docked at the membrane and others waiting in reserve. The release of neurotransmitter into the synaptic cleft occurs in the *active zone,* a specialized region of the plasma membrane containing a remarkable assemblage of proteins whose functions include modifying the properties of the synaptic vesicles and bringing them into position for docking and fusing with the plasma membrane. Viewed by electron microscopy, the active zone has electron-dense material and fine cytoskeletal filaments (Figure 23-23). A similarly dense region of specialized structures is seen across the synapse in the postsynaptic cell, the *postsynaptic density (PSD).* Cell-adhesion molecules that connect pre- and postsynaptic cells keep the active zone and PSD aligned. After release of synaptic vesicles in response to an action potential, the presynaptic neuron retrieves synaptic vesicle membrane proteins by **endocytosis** both within and outside the active zone.

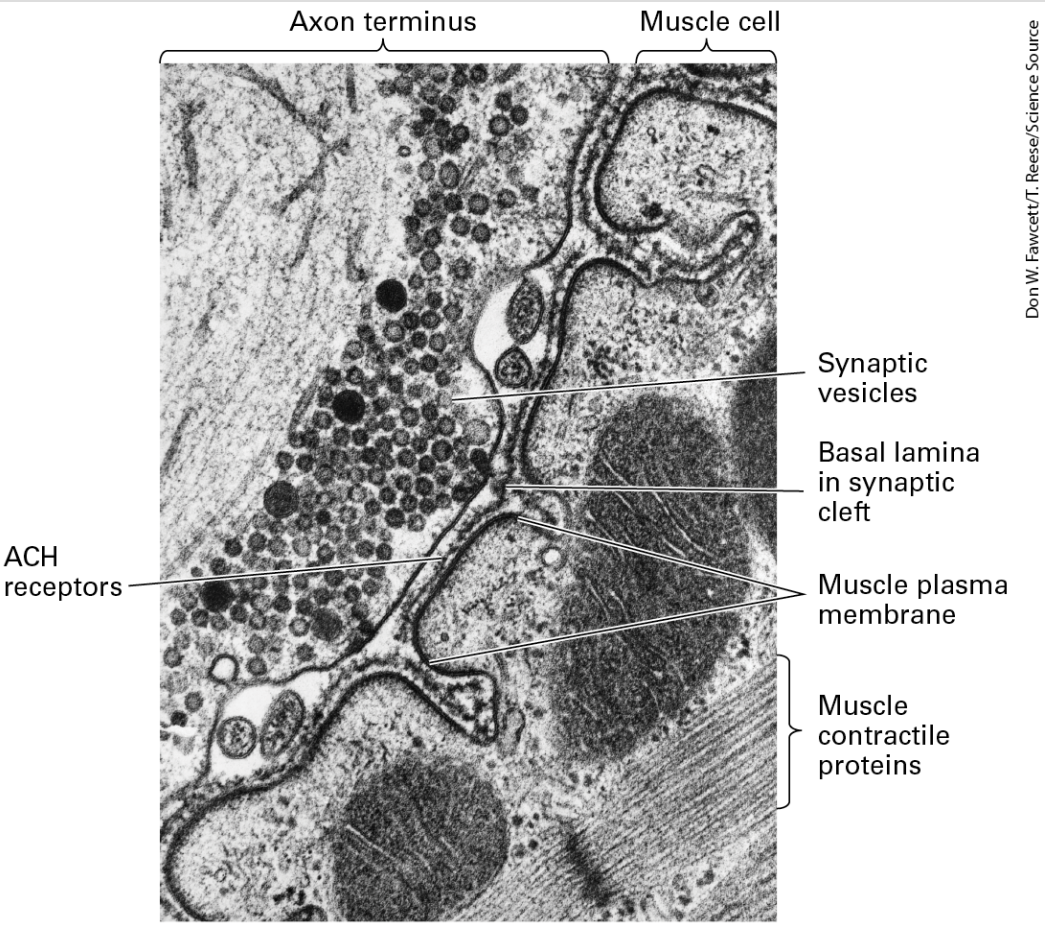
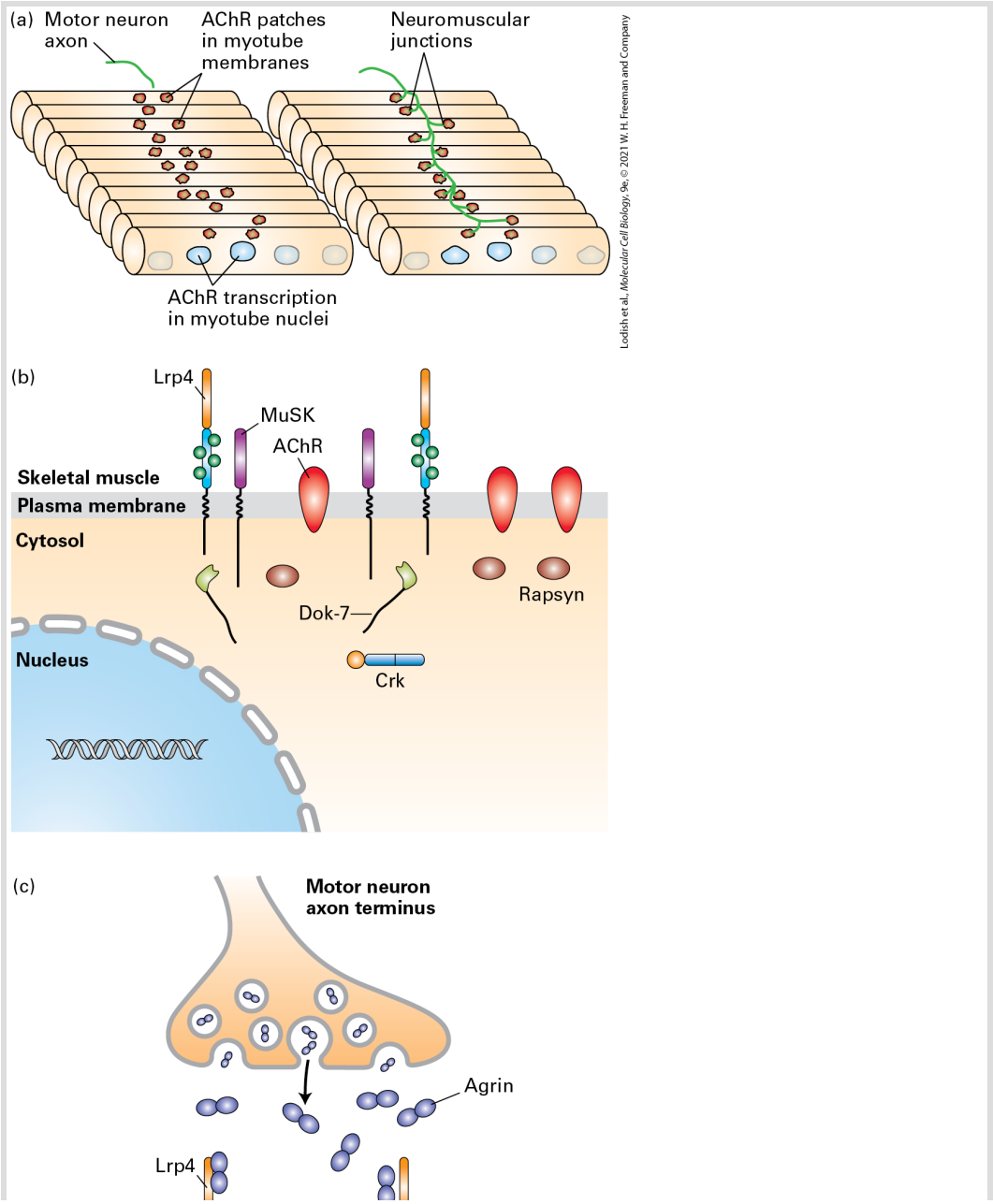


FIGURE 23-23 Synaptic vesicles are localized in the axon terminus near the region where neurotransmitter is released. In this longitudinal section through a neuromuscular junction, the basal lamina lies in the synaptic cleft separating the neuron from the muscle membrane, which is extensively folded. Acetylcholine receptors are concentrated in the postsynaptic muscle membrane at the top and partway down the sides of the folds in the membrane.

Synapse assembly has been extensively studied at the **neuromuscular junction (NMJ)** (Figure 23-23); we will use this example in the discussion below. At NMJ synapses, **acetylcholine** is the neurotransmitter produced by motor neurons, and the acetylcholine receptor, AChR, is produced by the postsynaptic

muscle cell. Muscle cell precursors, myoblasts, when put into culture, will spontaneously fuse into multinucleate myotubes that look similar to normal muscle cells. As myotubes form, AChR is produced near the center of the cell and inserted into the myotube plasma membrane, forming diffuse membrane patches (Figure 23-24a).



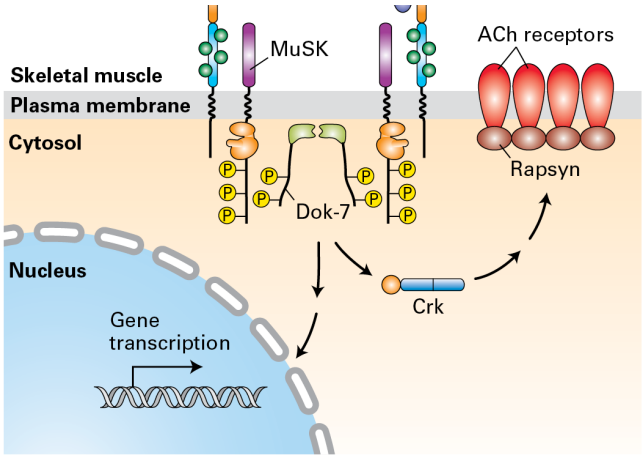


FIGURE 23-24 Formation of the neuromuscular junction. (a) Motor neuron-myotube interactions initiate formation of neuromuscular junctions (NMJs). Following fusion of myoblasts to form multinucleate myotubes, the nuclei synthesize acetylcholine receptor (AChR) mRNA. The nuclei near the center of each muscle fiber synthesize significantly more AChR mRNA than other nuclei. AChRs together with MuSK receptor kinases accumulate in membrane patches near the center of the cell, the prospective synaptic region of the muscle, prior to and independent of innervation; the cell is said to be "prepatterned." The motor neuron axon termini grow toward these AChR clusters and secrete the glycoprotein Agrin. Agrin, in turn, induces clustering of the AChRs (dark red) and MuSKs around the axon termini (green), forming the NMJ. (b and c) Signaling downstream of Agrin receptors is mediated by seven proteins: Agrin, LRP4, MuSK, Dok-7, Crk, rapsyn, and acetylcholine (ACh) receptors. (b) In the absence of motor neuron innervation, LRP4, MuSK, and ACh receptors diffuse within the plasma membrane and Dok-7, Crk, and rapsyn diffuse within the cytosol of the muscle cells. (c) Upon motor neuron innervation, motor axons secrete Agrin, which stabilizes postsynaptic differentiation by binding LRP4 and activating MuSK kinase activity. Phosphorylation of tyrosines in the MuSK juxtamembrane region stimulates recruitment and tyrosine phosphorylation of Dok-7, an adapter protein that is expressed selectively in muscle, which forms a dimer, stimulates MuSK kinase activity, and recruits the adapter protein Crk. Crk is essential to activate a rapsyn-dependent pathway for clustering AChRs opposite the presynaptic axon termini. Recent work has shown that rapsyn is an E3 ligase (see Figure 3-32) that adds a ubiquitin-like modification to the ACh receptors, which is necessary for receptor clustering. The pathway for synapse-

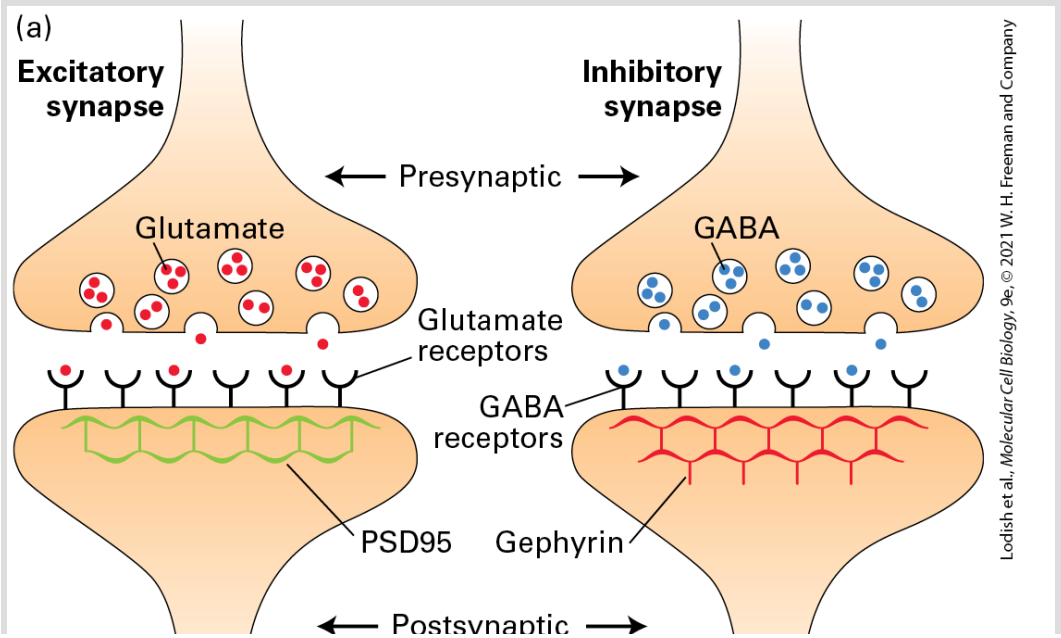
specific transcription is less well understood but likely involves kinase-dependent activation of transcription factors that stimulate expression of multiple genes encoding proteins required for the formation of cholinergic NMJ synapses.

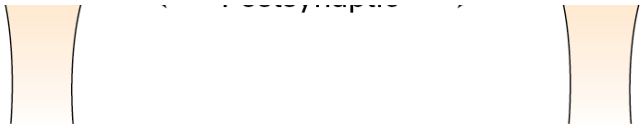
The formation of the neuromuscular synapse is a multistep process requiring signaling interactions between motor neurons and muscle fibers. A key player is **MuSK**, a receptor tyrosine kinase that is localized in the diffuse AChR-rich patches of the myotube plasma membrane. In ways that are not known, MuSK both induces clustering of AChRs and serves to attract the termini of growing motor neuron axons. For example, knockdown of MuSK inhibits both processes, while overexpression of MuSK in cultured muscle cells induces motor neuron growth throughout the muscle and formation of excess synapses.

Another key player is **Agrin**, a glycoprotein synthesized by developing motor neurons, transported in vesicles along axon microtubules, and secreted near the developing myotubes. Agrin binds to LRP4, a single-pass membrane-spanning protein; this stimulates an association between LRP4 and MuSK and increases MuSK kinase activity (Figure 23-24b). This leads to activation of several downstream signal transduction pathways, one of which leads to the formation of clusters of AChRs with the cytoskeletal protein rapsyn; this interaction, together with binding of other cytoskeletal proteins including actin, leads to localization of AChRs opposite the nerve termini at the neuromuscular junction. The

density of acetylcholine receptors in a mature synapse reaches is 1000-fold greater than in the plasma membrane.

While the molecular mechanisms underlying the formation of synapses in the central nervous system are less well understood, the process appears to follow a similar logic as that at NMJs, in which interactions between the pre- and postsynaptic compartments trigger a reorganization of already synthesized synaptic components. Analogous to the function of rapsyn in clustering AChRs at the neuromuscular synapse, distinct scaffolding proteins cluster neurotransmitter receptors at excitatory and inhibitory synapses in the central nervous system. A large PDZ-containing protein (see Chapter 20 for definition of PDZ domains) called *PSD95* clusters glutamate receptors at excitatory synapses, while another scaffolding protein called *gephryin* clusters GABA and glycine receptors at inhibitory synapses (Figure 23-25).





(b)

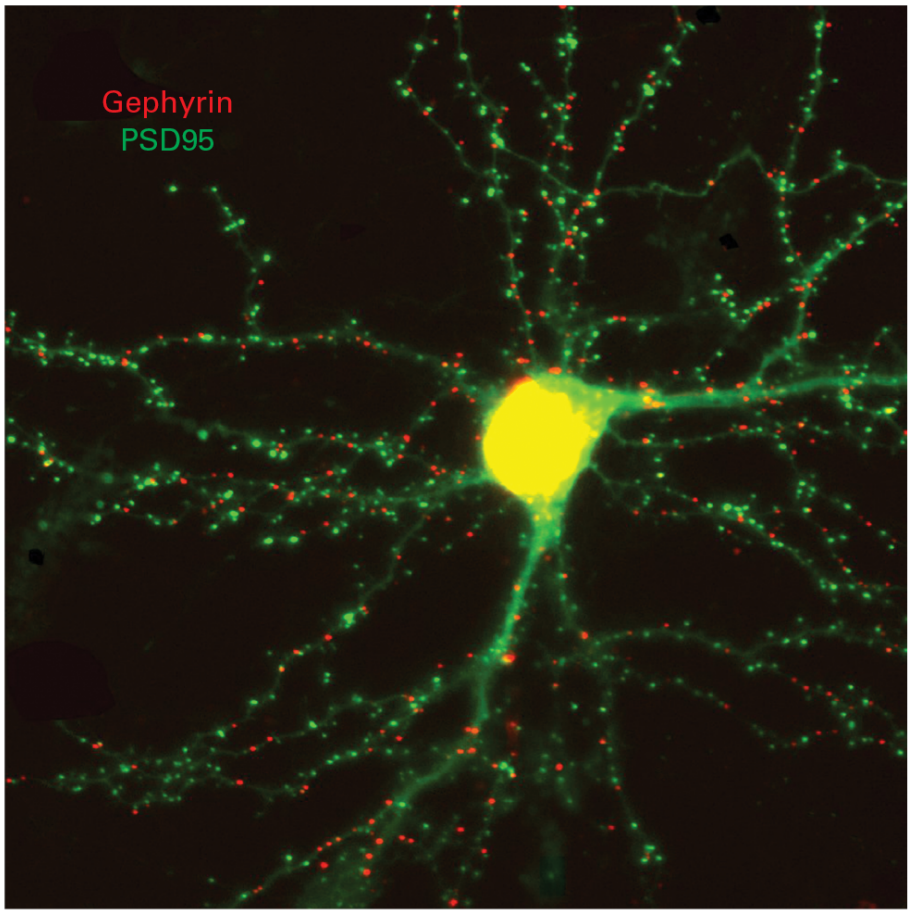


FIGURE 23-25 PSD95 and gephyrin are scaffolding proteins for excitatory and inhibitory postsynaptic compartments, respectively. (a) The PDZ-containing protein PSD95 is part of the postsynaptic density at excitatory synapses and serves as a scaffolding protein to cluster glutamate receptors at the postsynaptic membrane. The scaffolding protein gephyrin plays an analogous role at inhibitory synapses, where it clusters GABA receptors at the postsynaptic membrane. (b) PSD95 and gephyrin can be used to mark excitatory and inhibitory synapses, respectively. Shown in green are the PSD95-containing excitatory synapses and, in red, the gephyrin-containing inhibitory synapses that form on a single mouse cortical neuron in culture.

[Part (b) republished with permission from Elsevier, from G. G. Gross et al., 2013, "Recombinant Probes for Visualizing Endogenous Synaptic Proteins in Living Neurons," *Neuron* **78**(6):971–985; permission conveyed through Copyright Clearance Center, Inc.]

Pre- and postsynaptic compartments are linked by a network of trans-synaptic cell-adhesion molecules, whose adhesive interactions are so strong that it is not possible to biochemically separate the presynaptic compartment from the postsynaptic compartment. Synaptic adhesion molecules bind scaffolding proteins and cytoskeletal elements through their intracellular domains, promoting organization of protein complexes in both presynaptic and postsynaptic compartments. These adhesion molecules include cadherins, immunoglobulin-containing cell-adhesion molecules (described in Chapter 20), neurexins and neuroligins, ephrins, and Eph receptors.

Precise wiring in the nervous system involves not only synapse formation but also **synapse elimination**. At birth, most animals are born with an excess number of synapses, and neural circuit maturation involves a critical process of synaptic pruning that eliminates a large number of synapses. Synaptic pruning is dependent on neural activity, and studies indicate that the weaker synapses are eliminated while the stronger synapses are maintained. For example, at birth, each muscle cell is innervated by multiple motor neurons, but over time many of the motor neuron contacts are eliminated so that eventually each muscle cell is

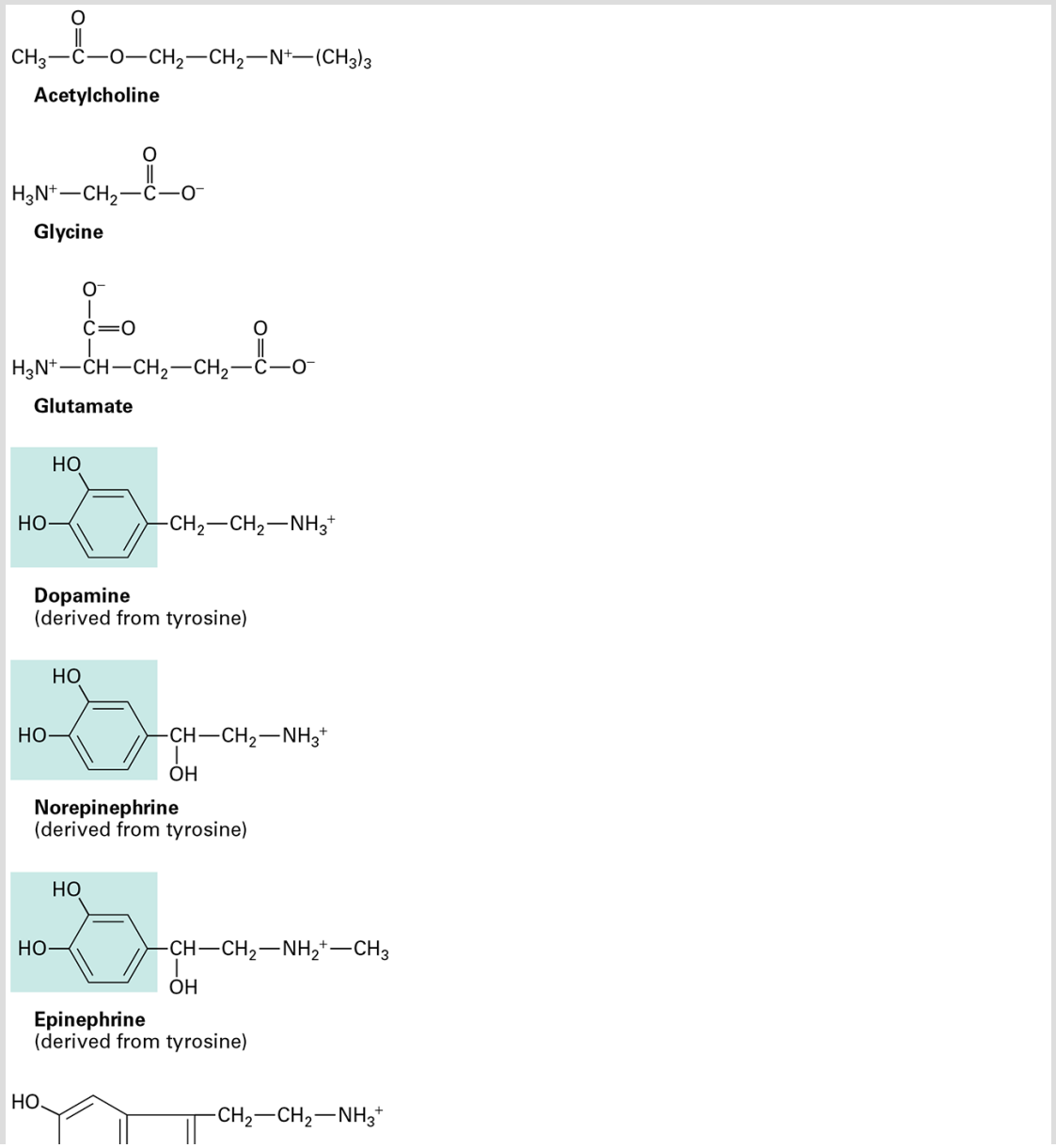
innervated by a single motor neuron. This process of synapse elimination is dependent on activity in the muscle cell: if the muscle cell is pharmacologically silenced, it never loses its polyneuronal innervation.

The cell biological mechanisms underlying synapse elimination, including the mechanisms that specify which synapses should be eliminated and the process of elimination itself, are areas of active research. Many recent studies have uncovered roles for astrocytes and microglia in the phagocytosis of eliminated synapses during synaptic pruning. In addition, human genetic studies have identified specific alleles in immune genes expressed by microglia, including in components of the complement system (see Chapter 24), as major risk alleles for schizophrenia, consistent with the hypothesis that aberrant synapse elimination contributes to this psychiatric disease.

Neurotransmitters Are Transported into Synaptic Vesicles by H^+-Linked Antiport Proteins

Numerous small molecules function as neurotransmitters at various synapses. With the exception of acetylcholine, the most common neurotransmitters, as shown in Figure 23-26, are amino acids or derivatives of amino acids. Nucleotides such as ATP and the corresponding nucleosides, which lack phosphate groups, also

function as neurotransmitters. While neurons were previously thought to make only one type of neurotransmitter, more recent studies indicate that some neurons can produce and release two or more. While the types of neurotransmitters are varied and while they operate in different parts of the nervous system, their release results in one of two outcomes: either the induction of an electrical signal or its inhibition.



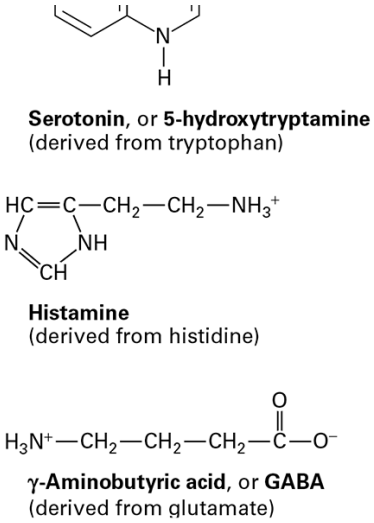
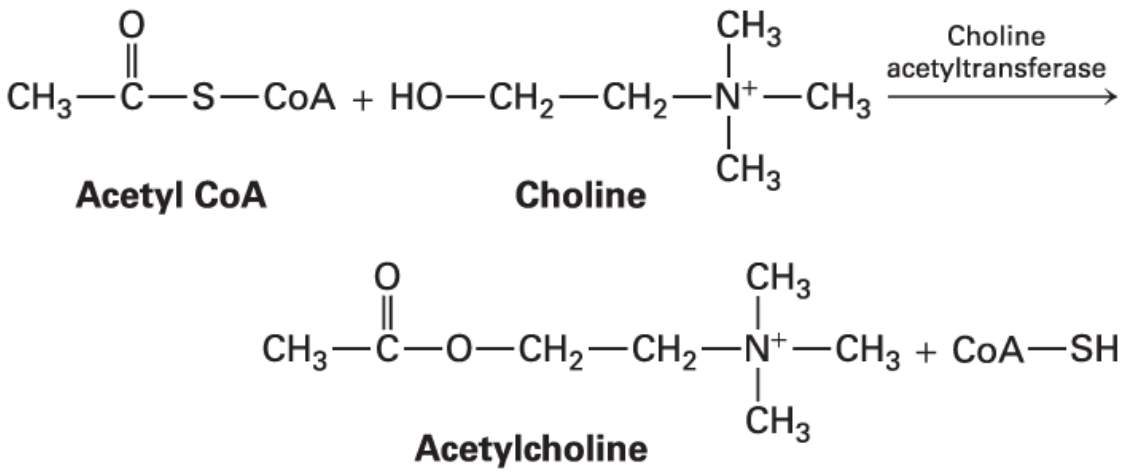


FIGURE 23-26 Structures of several small molecules that function as neurotransmitters. Except for acetylcholine, all of these molecules are amino acids (glycine and glutamate) or derived from the indicated amino acids. The three transmitters synthesized from tyrosine, which contain the catechol moiety (blue highlight), are referred to as *catecholamines*.

All of these neurotransmitters are synthesized in the cytosol and imported into membrane-bound synaptic vesicles within axon termini, where they are stored. These vesicles are 40–50 nm in diameter, and their lumen has a low pH, generated by operation of a V-class proton pump in the vesicle membrane. Similar to the accumulation of metabolites in plant vacuoles (see Figure 11-29), this proton concentration gradient (greater in vesicle lumen than in cytosol) drives neurotransmitter import by ligand-specific H^+-linked neurotransmitter antiporters in the vesicle membrane (see Figure 23-26).

For example, acetylcholine is synthesized in the cytosol from a precursor molecule, acetyl coenzyme A (acetyl CoA), and choline in a reaction catalyzed by choline acetyltransferase:



Synaptic vesicles take up and concentrate acetylcholine from the cytosol against a steep concentration gradient, using an H^+/acetylcholine antiporter in the vesicle membrane. As with other antiporters, the export of protons from the forming vesicle down its electrochemical gradient powers the uptake of the neurotransmitter. As an example of an evolutionary mechanism to ensure coordinate expression of these two proteins, the gene encoding this antiporter is contained entirely within the first intron of the gene encoding choline acetyltransferase.

Different H^+/neurotransmitter antiport proteins are used for import of other neurotransmitters into synaptic vesicles. For example, glutamate is imported into synaptic vesicles by a family of proteins called *vesicular glutamate transporters (VGLUTs)*. VGLUTs are highly

specific for glutamate but have rather low substrate affinity. Another family of transporters, the *vesicular GABA transporters (VGATs)* transport GABA and glycine into synaptic vesicles. Like the acetylcholine transporter, VGLUTs and VGATs are antiporters, moving glutamate and GABA or glycine into synaptic vesicles while protons move in the other direction. VGLUT and VGAT serve as useful markers for excitatory and inhibitory synaptic terminals, respectively.

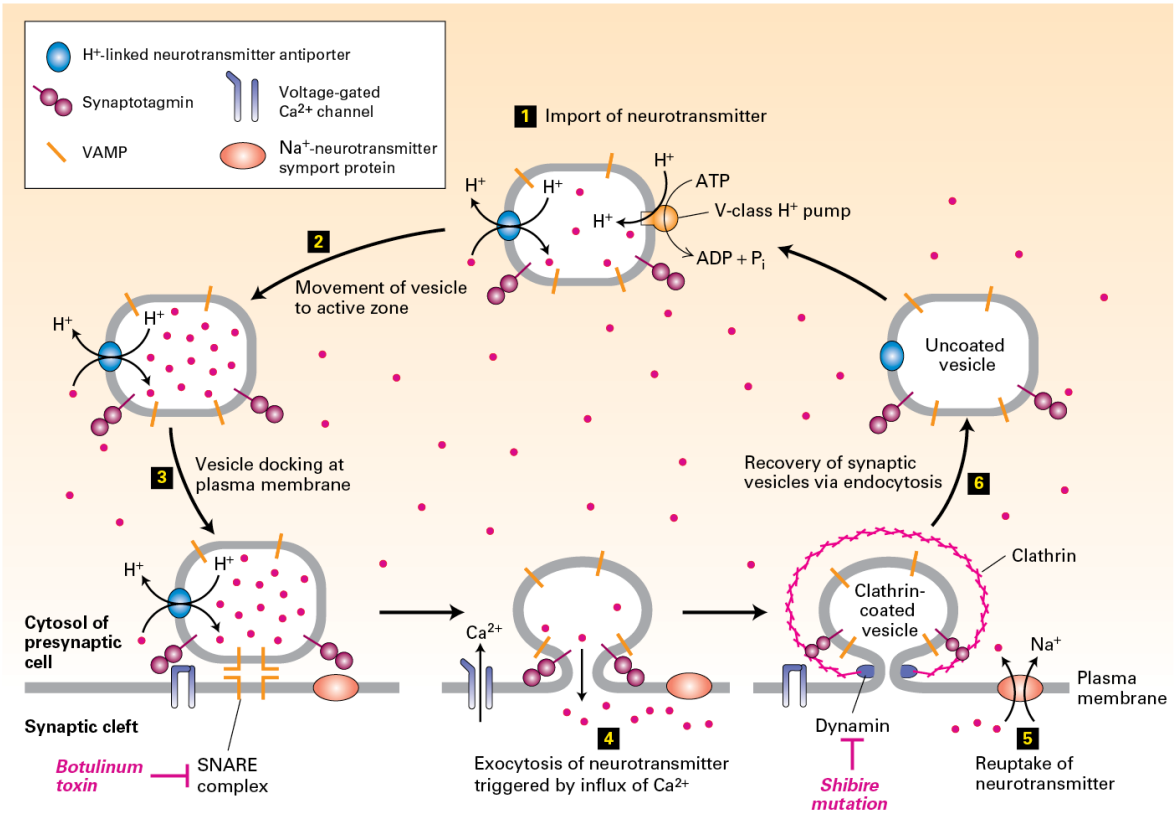
Three Pools of Synaptic Vesicles Loaded with Neurotransmitter Are Present in the Presynaptic Terminal

A highly organized arrangement of cytoskeletal fibers in the axon terminus helps localize synaptic vesicles within the presynaptic terminal. The population of synaptic vesicles has been proposed to exist in three states: a small readily releasable pool, which is docked at the active zone near the plasma membrane; a larger recycling pool, which is proximal but not docked at the plasma membrane and is released with moderate stimulation; and a reserve pool, which includes the majority of synaptic vesicles in the terminal. The reserve pool is the most distal from the active zone and is released only in response to particularly strong stimuli. A family of phosphoproteins called *synapsins* tether synaptic vesicles to the actin cytoskeleton and to one another. Neuronal stimulation activates

kinases that phosphorylate synapsins to modulate synaptic vesicle tethering and thereby alter the number of synaptic vesicles available for release. Indeed, synapsin knockout mice, although viable, are prone to seizures; during repetitive stimulation of many neurons in such mice, the number of synaptic vesicles that fuse with the plasma membrane is greatly reduced.

Influx of Ca^{2+} Triggers Release of Neurotransmitters

The exocytosis of neurotransmitters from synaptic vesicles involves vesicle-targeting and fusion events similar to those that occur during the intracellular transport of secreted and plasma-membrane proteins (see Chapter 14). However, two unique features critical to synapse function differ from secretory pathways in non-neuronal cells: (1) secretion is tightly coupled to arrival of an action potential at the axon terminus, and (2) synaptic vesicles are recycled locally within the axon terminus after fusion with the plasma membrane. Figure 23-27 shows the entire cycle whereby synaptic vesicles are filled with neurotransmitter, release their contents, and are recycled.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 23-27 Cycling of neurotransmitters and of synaptic vesicles in axon termini. Most synaptic vesicles are formed by endocytic recycling as depicted here (although note that some synaptic vesicles are transported from the cell body by axonal transport along microtubules). The entire cycle typically takes about 60 seconds. Step **1**: The uncoated vesicles express a V-type proton pump (orange) and a single type of H^+-neurotransmitter antiporter (blue) specific for the particular neurotransmitter, to import neurotransmitters (red dots) from the cytosol. Step **2**: Synaptic vesicles loaded with neurotransmitter move to the active zone. Step **3**: Vesicles dock at defined sites on the plasma membrane of the presynaptic cell, and the vesicle v-SNAREs called *VAMP* bind to the plasma membrane t-SNAREs, forming a SNARE complex. Synaptotagmin prevents membrane fusion and release of neurotransmitter. Botulinum toxin prevents exocytosis by proteolytically cleaving VAMP, the v-SNARE on vesicles. Step **4**: In response to a nerve impulse (action potential), voltage-gated Ca^{2+} channels in the plasma membrane open, allowing an influx of Ca^{2+} from the extracellular medium. The elevation in cytosolic Ca^{2+} leads to fusion of docked vesicles with the plasma membrane and release of neurotransmitters into the synaptic cleft (via mechanisms highlighted in Figure 23-29). Step **5**: Na^+ symporter proteins take up neurotransmitter from the synaptic cleft into the cytosol, which limits the duration of the

action potential and partially recharges the cell with transmitter. Step **6**: Vesicles are recovered by endocytosis, creating uncoated vesicles, ready to be refilled and begin the cycle anew. After clathrin/AP vesicles containing neurotransmitter transporter proteins bud inward and are pinched off in a dynamin-mediated process, they lose their coat proteins. Dynamin mutations such as *shibire* in *Drosophila* block the re-formation of synaptic vesicles, leading to paralysis. Unlike most neurotransmitters, acetylcholine is not recycled. See K. Takei et al., 1996, *J. Cell Biol.* **133**:1237; V. N. Murthy and C. F. Stevens, 1998, *Nature* **392**:497; and R. Jahn, T. Lang, and T. C. Südhof, 2003, *Cell* **112**:519.

Depolarization of the plasma membrane cannot, by itself, cause synaptic vesicles to fuse with the plasma membrane. In order to trigger vesicle fusion, an action potential must be converted into a chemical signal — namely, a localized rise in the cytosolic Ca^{2+} concentration. The transducers of the electrical signals are *voltage-gated* Ca^{2+} channels localized to the region of the plasma membrane adjacent to the synaptic vesicles. The membrane depolarization due to the arrival of an action potential opens these channels, permitting an influx of Ca^{2+} ions from the extracellular medium into the region of the axon terminus near the docked synaptic vesicles. Importantly, the rise in cytosolic Ca^{2+} is localized; it is also transient, as the excess Ca^{2+} is rapidly pumped out of the cell by plasma membrane Ca^{2+} pumps.

A single action potential leads to exocytosis of about 10 percent of synaptic vesicles in a presynaptic terminal. Membrane proteins unique to synaptic vesicles then are specifically internalized by endocytosis, usually via the same types of clathrin-coated vesicles used to recover other plasma membrane proteins by other types of

cells. After the endocytosed vesicles lose their clathrin coat, they are rapidly refilled with neurotransmitter. The ability of many neurons to fire 50 times a second is clear evidence that the recycling of vesicle membrane proteins occurs quite rapidly (see Figure 23-27). The machinery of endocytosis and exocytosis is highly conserved; it is described in more detail in Chapter 14.

The development of fluorescent Ca^{2+} indicators has provided a powerful means of visualizing synaptic activity in neurons in culture and in intact neural circuits. As discussed in Chapter 4, these indicators are fluorescent molecules that change their fluorescence emission upon Ca^{2+} binding and include both chemical indicators and genetically encoded indicators. Delivery or expression of Ca^{2+} indicators to neurons in a circuit allows experimenters to use time-lapse microscopy to monitor Ca^{2+} transients in hundreds of neurons and glia in real time. For example, expression of the genetically encoded Ca^{2+} indicator GCaMP6 in the visual cortex of mice, combined with presentation of visual stimuli and in vivo two-photon microscopy, has been used to identify the population of neurons that respond to specific orientations of visual information (Figure 23-28).

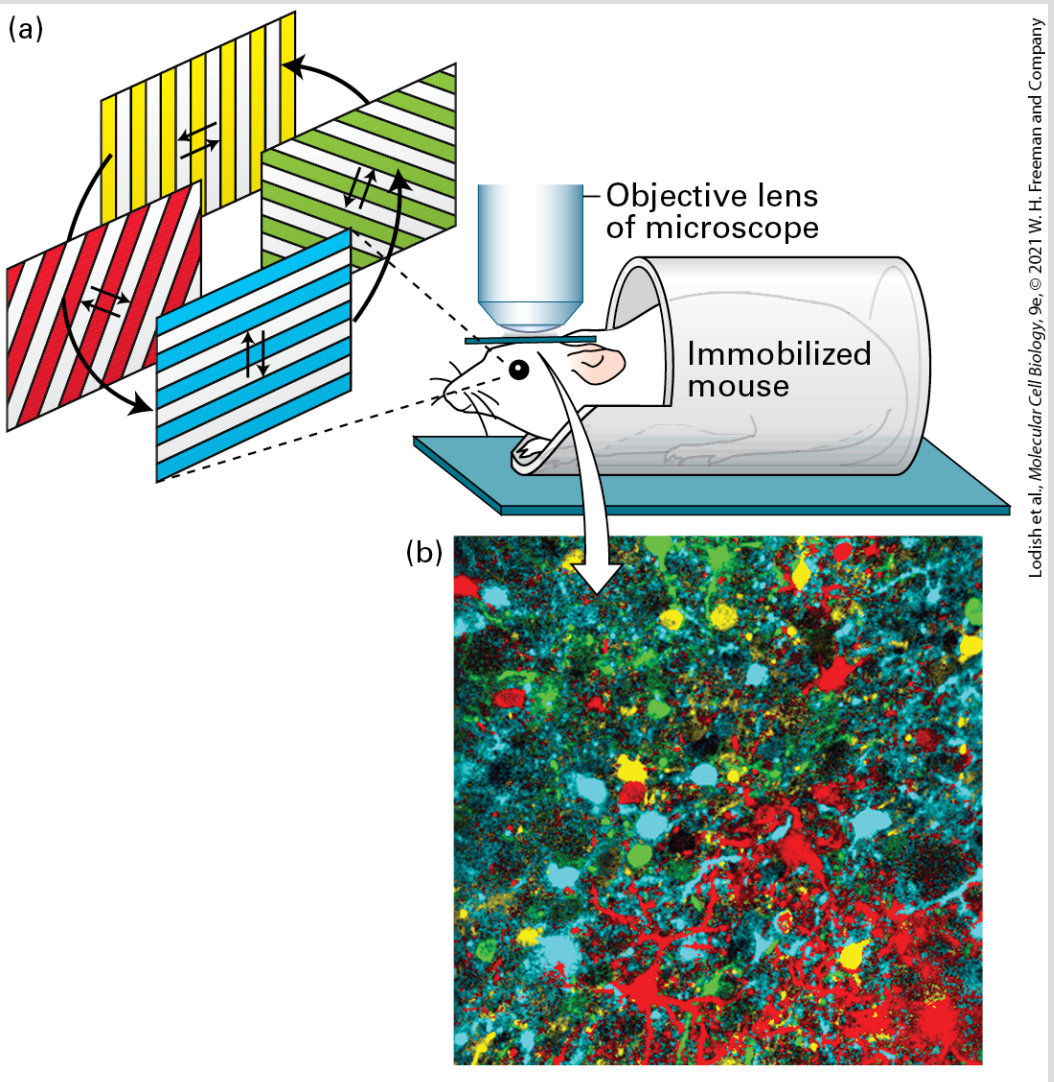


FIGURE 23-28 Calcium indicators allow visualization of activity in neural circuits. A genetically encoded calcium indicator was expressed in neurons in the mouse visual cortex. (a) A window was made in the skull of the mouse, and a microscope (indicated by the objective lens) was used to visualize calcium transients in populations of neurons in the visual cortex while the mouse was looking at gratings that moved in different directions. Individual neurons within the visual cortex respond to specific orientations of the gratings, as detected by elevations in calcium that are visualized as increases in the fluorescence of the calcium indicator. (b) Neurons were color coded according to the orientation that elicited increases in calcium (as shown below the photo). The neurons shown in yellow respond to horizontally moving gratings, and the neurons shown in cyan respond to vertically moving gratings, while the neurons shown

in green and red respond to diagonally oriented gratings. This type of experiment reveals that individual neurons are tuned to specific orientations of visual stimuli.

[Part (b) republished with permission from Nature, from T.-W. Chen et al., 2013, "Ultrasensitive Fluorescent Proteins for Imaging Neuronal Activity," *Nature* **499**(7458):295–300; permission conveyed through Copyright Clearance Center, Inc.]

A Calcium-Binding Protein Regulates Fusion of Synaptic Vesicles with the Plasma Membrane

Fusion of synaptic vesicles with the plasma membrane of axon termini depends on **SNAREs**, the same types of proteins that mediate membrane fusion of other regulated secretory vesicles, and **SM proteins** (for Sec1/Munc18-like proteins). The principal v-SNARE in synaptic vesicles (VAMP) tightly binds syntaxin and SNAP-25, the principal t-SNAREs in the plasma membrane of axon termini, to form four-helix SNARE complexes. The assembly of the SNARE complex brings the synaptic vesicle membrane into close proximity to the presynaptic plasma membrane, but the formation of a fusion pore requires an additional step, association of an SM protein with syntaxin. After fusion, proteins within the axon terminus promote disassociation of VAMP from t-SNAREs, as in the fusion of secretory vesicles depicted in Figure 14-11.

Strong evidence for the role of VAMP in neurotransmitter exocytosis is provided by the mechanism of action of botulinum toxin, a bacterial protein that can cause the paralysis and death characteristic of *botulism,* a type of food poisoning. The toxin is composed of two polypeptides, one of which binds to motor neurons that release acetylcholine at synapses with muscle cells, facilitating entry of the other polypeptide, a protease, into the cytosol of the axon terminus. The only protein this protease cleaves is VAMP (see Figure 23-27, step **3**). After the botulinum protease enters an axon terminus, synaptic vesicles that are not already docked rapidly lose their ability to fuse with the plasma membrane because cleavage of VAMP prevents assembly of SNARE complexes. The resulting block in acetylcholine release at neuromuscular synapses causes paralysis. However, vesicles that are already docked exhibit remarkable resistance to the toxin, indicating that SNARE complexes may already be in a partially assembled, protease-resistant state when vesicles are docked on the presynaptic membrane.

The signal that triggers exocytosis of docked synaptic vesicles is a very localized rise in the Ca^{2+} concentration in the cytosol near vesicles from 0.1 μM, characteristic of resting cells, to 1–100 μM following arrival of an action potential in stimulated cells. The speed with which synaptic vesicles fuse with the presynaptic membrane after a rise in cytosolic Ca^{2+} (less than 1 ms) indicates that the fusion machinery is entirely assembled in the resting state and can

rapidly undergo a conformational change leading to exocytosis of neurotransmitter (Figure 23-29). A Ca^{2+}-binding protein called *synaptotagmin,* located in the membrane of synaptic vesicles, is a key component of the vesicle-fusion machinery that triggers exocytosis in response to Ca^{2+}. A protein called *complexin* is thought to bind to the α-helical bundle of an assembled v-SNARE/t-SNARE complex that bridges the synaptic vesicle and plasma membranes, preventing the final fusion step. Binding of Ca^{2+} to synaptotagmin relieves this inhibition, releasing complexin and allowing the fusion event to occur very rapidly. Mammals express multiple different synaptotagmin isoforms, each of which has a different binding affinity for Ca^{2+}, and as a result the kinetics of exocytosis depend on the particular synaptotagmin isoform expressed in the neuron.

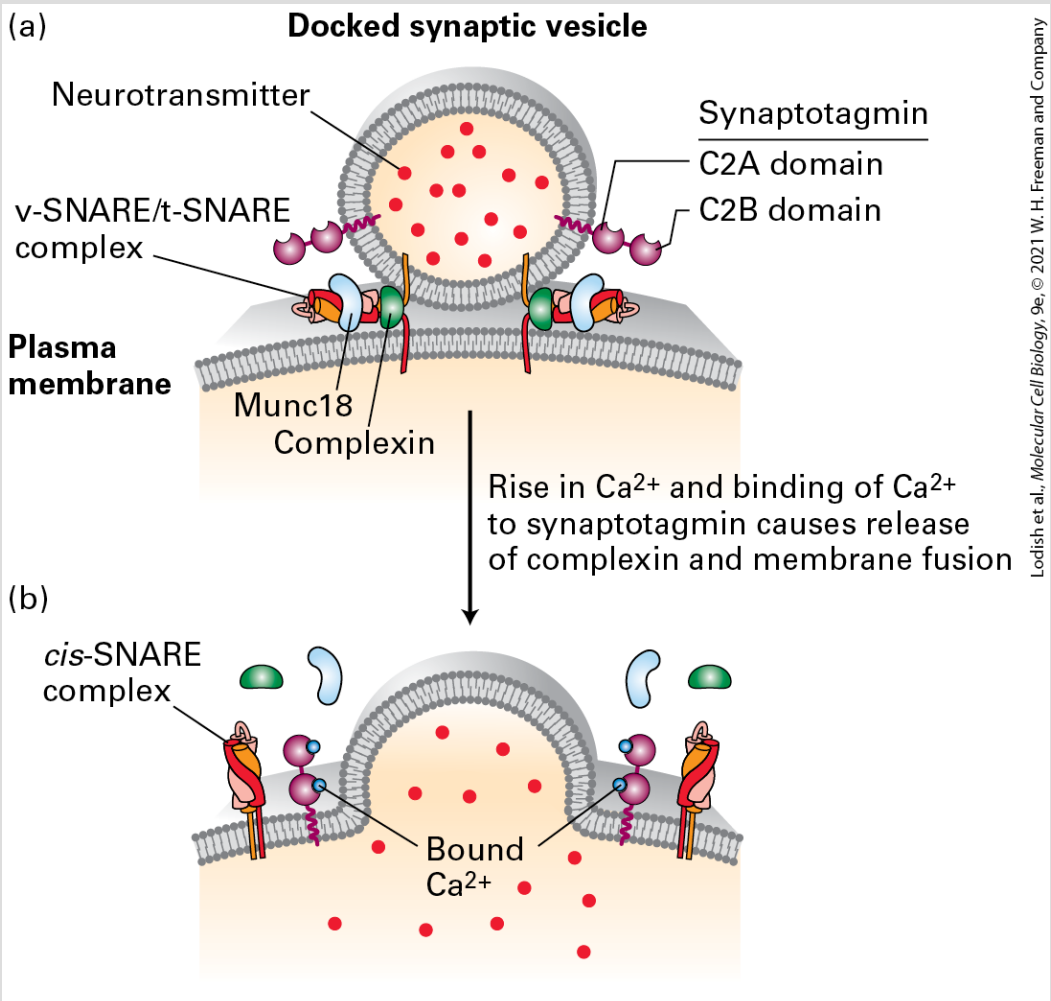


FIGURE 23-29 Synaptotagmin- and SNARE-mediated fusion of synaptic vesicles with the plasma membrane. Only a few synaptic vesicles are docked at the presynaptic plasma membrane; these are primed for fusion with the plasma membrane. (a) The tight interconnections between the synaptic vesicle and plasma membrane are mediated in part by bundles of four α helices derived from complexes of vesicle v-SNARE (e.g., VAMP) and plasma membrane t-SNARE proteins (e.g., SNAP-25 and syntaxin; see Figure 14-11). The fusion of the two membranes is prevented by binding of complexin protein to the v-SNARE/t-SNARE complex. Synaptotagmin is composed of a short intraluminal sequence, a single transmembrane α helix that anchors it in the synaptic vesicle membrane, a linker, and two Ca^{2+}-binding domains termed *C2A* and *C2B*. Synaptotagmin without bound Ca^{2+} may also bind to the v-SNARE/t-SNARE complex and prevent membrane fusion. (b) A localized rise in Ca^{2+} allows Ca^{2+} ions to bind to synaptotagmin, altering its three-dimensional

conformation. This triggers release of the complexin fusion inhibitor, binding (or altered binding) of synaptotagmin to the v-SNARE/t-SNARE complex, instantaneous membrane fusion, and release of neurotransmitters into the extracellular space. The SM protein Munc18, which binds to syntaxin, is required for SNARE-mediated fusion, although its precise mechanisms of action are not known. See T. C. Südhof and J. E. Rothman, 2009, *Science* **323**:474; and T. C. Südhof, 2013, *Neuron* **80**:675–690.

Several lines of evidence support a role for synaptotagmin as the Ca^{2+} sensor for exocytosis of neurotransmitters. Mutant embryos of *Drosophila* and *Caenorhabditis elegans* that completely lack synaptotagmin fail to hatch and exhibit very reduced, uncoordinated muscle contractions. Larvae with partial loss-of-function mutations of synaptotagmin survive, but their neurons are defective in Ca^{2+}-stimulated vesicle exocytosis. Moreover, in mice, mutations in synaptotagmin that decrease its affinity for Ca^{2+} cause a corresponding increase in the amount of cytosolic Ca^{2+} needed to trigger rapid exocytosis.

An important characteristic of synaptic vesicle exocytosis is its speed. Synaptic vesicle fusion occurs within a few hundred microseconds after the arrival of an action potential, which is not very different from the timescale of Ca^{2+} influx through the voltage-gated Ca^{2+} channel. What makes this speed possible is the proximity of the release machinery to the voltage-gated Ca^{2+} channels. This proximity is mediated by two scaffolding proteins called *RIM* (for *R*ab3-*i*nteracting *m*olecule protein) and *RIM-BP* (for

*RIM b*inding *p*rotein), which form a complex between Rab3-containing synaptic vesicles and voltage-gated Ca^{2+} channels. In mice lacking RIM, and flies lacking RIM-BP, active zones lack voltage-gated Ca^{2+} channels, which leads to a dramatic decrease and desynchronization of neurotransmitter release.

Fly Mutants Lacking Dynamin Cannot Recycle Synaptic Vesicles

Synaptic vesicles are formed primarily by endocytic budding from the plasma membrane of axon termini (see Figure 23-27). Endocytosis usually involves clathrin-coated pits and is quite specific, in that several membrane proteins unique to the synaptic vesicles (e.g., neurotransmitter transporters) are specifically incorporated into the endocytosed vesicles and resident plasma membrane proteins (e.g., the voltage-sensitive Ca^{2+} channel) remain. In this way, synaptic-vesicle membrane proteins can be reused and the recycled vesicles refilled with neurotransmitter (see Figure 23-27, step **6**).

As in the formation of other clathrin/AP-coated vesicles, pinching off of endocytosed synaptic vesicles requires the GTP-binding protein *dynamin* (see Figure 14-20). Indeed, analysis of a temperature-sensitive *Drosophila* mutant called *shibire* (shi^{ts}), which encodes the fly dynamin protein, provided early evidence for the

role of dynamin in endocytosis. At the permissive temperature of 20 °C, the mutant flies are normal, but at the nonpermissive temperature of 30 °C, they are paralyzed (*shibire* means "paralyzed" in Japanese) because pinching off of clathrin-coated pits in neurons and other cells is blocked. When viewed in the electron microscope, the shi^{ts} neurons at 30 °C show abundant clathrin-coated pits with long necks but few clathrin-coated vesicles. Because of their inability to pinch off new synaptic vesicles, the neurons in shi^{ts} mutants eventually become depleted of synaptic vesicles when flies are shifted to the nonpermissive temperature, leading to a cessation of synaptic signaling and to paralysis.

Signaling at Synapses Is Terminated by Degradation or Reuptake of Neurotransmitters

Following neurotransmitter release from a presynaptic cell, signaling can be terminated by diffusion of a transmitter away from the synaptic cleft, but this is a slow process. Instead, one of two more rapid mechanisms terminates the action of neurotransmitters at most synapses: neurotransmitters are actively removed from the synapse by neurotransmitter transporters or destroyed by enzymes present in the synaptic cleft to prevent continued stimulation of the postsynaptic cell.

Signaling by acetylcholine is terminated when it is hydrolyzed to acetate and choline by *acetylcholinesterase,* an enzyme localized to the synaptic cleft. Choline released in this reaction is transported back into the presynaptic axon terminus by a Na^+/choline symporter and used in synthesis of more acetylcholine. The operation of this transporter is similar to that of the Na^+-linked symporters used to transport glucose into cells against a concentration gradient (see Figure 11-26).

With the exception of acetylcholine, all the neurotransmitters shown in Figure 23-26 are removed from the synaptic cleft by transport back into the axon termini that released them. These transmitters are recycled intact, as depicted in Figure 23-27 (step **5**). Transporters for GABA, norepinephrine, dopamine, and serotonin were the first to be cloned and studied. These four transport proteins are all Na^+-linked symporters. They are 60–70 percent identical in their amino acid sequences, and each is thought to contain 12 transmembrane α helices. As with other Na^+ symporters, the movement of Na^+ into the cell down its electrochemical gradient provides the energy for uptake of the neurotransmitter. To maintain electroneutrality, Cl^- often is transported via an ion channel along with the Na^+ and neurotransmitter.

Neurotransmitter transporters are targets of a variety of drugs of abuse as well as many therapeutic drugs commonly used in

psychiatry. Cocaine binds to and inhibits the transporters for norepinephrine, serotonin, and dopamine. In particular, binding of cocaine to the dopamine transporter inhibits reuptake of dopamine, causing a higher than normal concentration of dopamine to remain in the synaptic cleft and prolonging the stimulation of postsynaptic neurons. Long-lasting exposure to cocaine, as occurs with habitual use, leads to down-regulation of dopamine receptors and thus altered regulation of dopaminergic signaling. It is thought that decreased dopaminergic signaling after chronic cocaine use may contribute to depressive mood disorders and sensitize important brain reward circuits to the reinforcing effects of cocaine, leading to addiction. Similarly, therapeutic agents such as the antidepressant drugs fluoxetine (Prozac) and imipramine block serotonin reuptake, and the tricyclic antidepressant desipramine blocks norepinephrine reuptake. [Fluoxetine and similarly acting drugs such as paroxetine (Paxil) and sertraline (Zoloft) are often referred to collectively as *s*elective *s*erotonin *r*euptake *i*nhibitors (SSRIs).] These drugs also cause a higher than normal concentration of neurotransmitter to remain in the synaptic cleft and prolong the stimulation of postsynaptic neurons. ■

Exciting recent single-molecule cryoelectron microscopy studies have solved the structure of several neurotransmitter transporters, including the dopamine transporter and the serotonin transporter, in the presence and absence of drugs such as the hallucinogen ibogaine. These studies have confirmed the hypothesis that

neurotransmitter transporters share structural and functional features with the bacterial two-Na^+/one-leucine symporter (see Figures 11-25 and 11-26). This is important because past research on drugs that target neurotransmitter transporters, including antidepressants, were based on the structure of the bacterial two-Na^+/one-leucine symporter. Determination of the ultrastructure of the neurotransmitter transporters themselves not only validates these models but also provides a powerful means of developing more refined and targeted therapeutic drugs.

Opening of Acetylcholine-Gated Cation Channels Leads to Muscle Contraction

Now we look at how binding of neurotransmitters by receptors on postsynaptic cells leads to changes in the cells' membrane potential, using the communication between motor neurons and muscles as an example. At neuromuscular junctions, acetylcholine is the neurotransmitter. A single axon terminus of a frog motor neuron may contain a million or more synaptic vesicles, each containing 1000–10,000 molecules of acetylcholine; these vesicles often accumulate in rows in the active zone (see Figures 23-23 and 23-24). Such a neuron can form synapses with a single skeletal muscle cell at several hundred points.

The *nicotinic acetylcholine receptor*, which is expressed in muscle cells, is a *ligand-gated channel* that admits both K^+ and Na^+. These receptors are also produced in brain neurons and are important in learning and memory; loss of these receptors is observed in schizophrenia, epilepsy, drug addiction, and Alzheimer's disease. Antibodies against acetylcholine receptors constitute a major part of the autoimmune reactivity in the disease *myasthenia gravis*. The receptor is so named because it is bound by nicotine; it has been implicated in nicotine addiction in tobacco smokers. There are at least 14 different isoforms of the receptor, which assemble into homo- and heteropentamers with varied properties. Given their many physiological functions and their role in disease, these various isoforms are important targets for new drug development. ■

The effect of acetylcholine on this receptor can be determined by patch-clamp recording from isolated outside-out patches of muscle plasma membranes. Outside-out patch-clamp recording is a technique that measures the effects of extracellular solutes on channel receptors within the isolated patch (see Figure 11-22c). Such measurements have shown that acetylcholine causes opening of a cation channel in the receptor capable of transmitting 15,000–30,000 Na^+ and K^+ ions per millisecond. However, since the resting potential of the muscle plasma membrane is near E_K, the potassium equilibrium potential, opening of acetylcholine receptor channels causes little increase in the efflux of K^+ ions; Na^+ ions, on the other

hand, flow into the muscle cell, driven by the Na^+ electrochemical gradient.

The simultaneous increase in permeability to Na^+ and K^+ ions following binding of acetylcholine produces a net depolarization to about −15 mV from the muscle resting potential of −85 to −90 mV. As shown in Figure 23-30, this localized depolarization of the muscle plasma membrane triggers opening of voltage-gated Na^+ channels, leading to generation and conduction of an action potential in the muscle cell-surface membrane by the same mechanisms described previously for neurons. When the membrane depolarization reaches transverse tubules, specialized invaginations of the plasma membrane (see Figure 17-32b), it acts on Ca^{2+} channels in the plasma membrane apparently without causing them to open. This in turn triggers the opening of adjacent Ca^{2+}-release channels in the sarcoplasmic reticulum membrane. The subsequent flow of stored Ca^{2+} ions from the sarcoplasmic reticulum into the cytosol raises the cytosolic Ca^{2+} concentration sufficiently to induce muscle contraction.

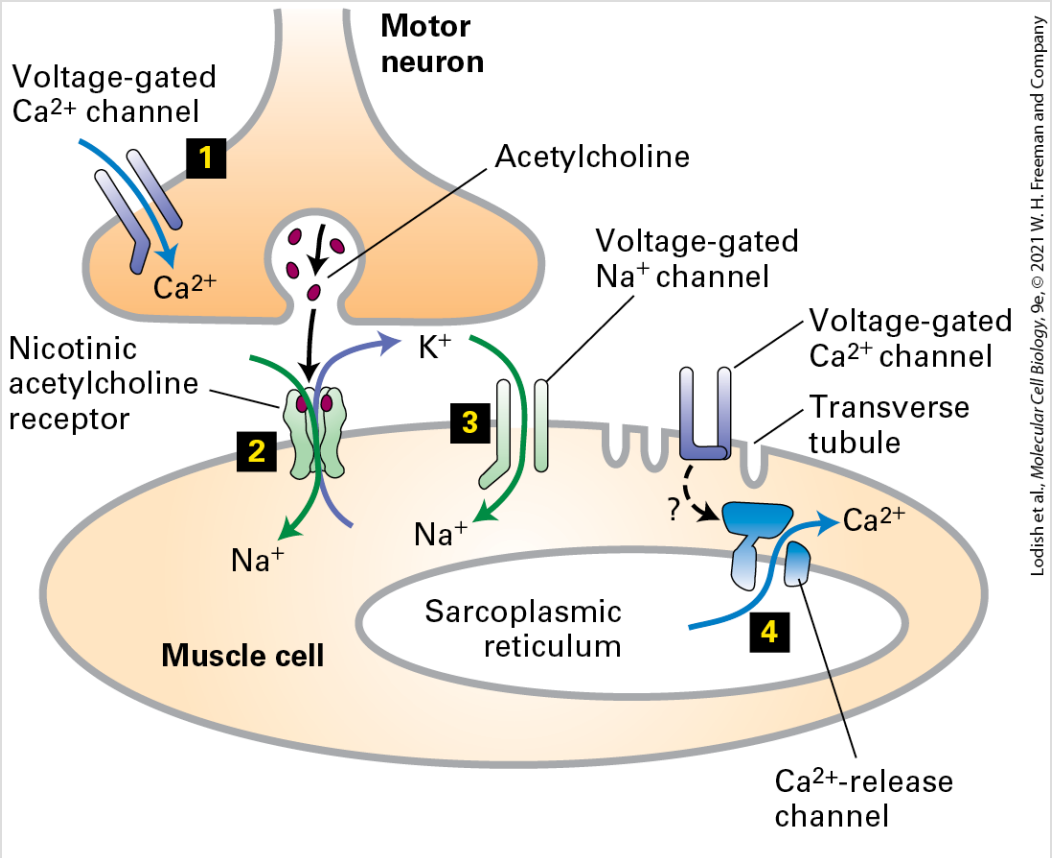


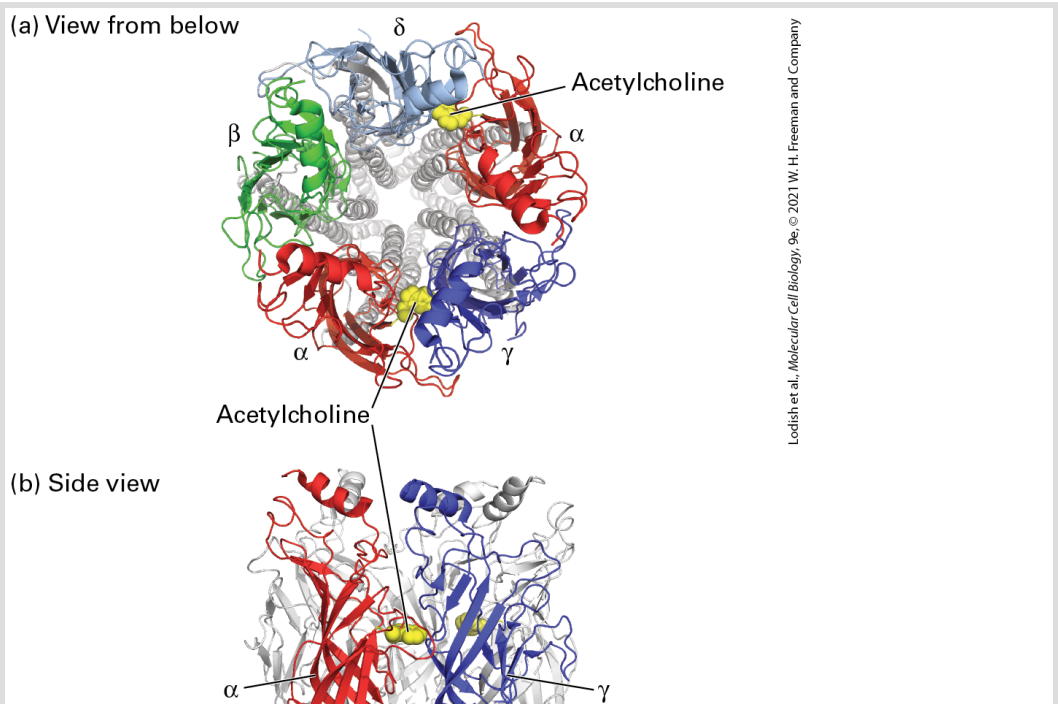
FIGURE 23-30 Sequential activation of gated ion channels at a neuromuscular junction. Arrival of an action potential at the terminus of a presynaptic motor neuron induces (step 1) opening of voltage-gated Ca^{2+} channels in the neuron and subsequent release of acetylcholine, which (step 2) triggers opening of the ligand-gated acetylcholine receptors in the muscle plasma membrane. The open receptor channel allows an influx of Na^+ and an efflux of K^+ from the muscle cell. The Na^+ influx produces a localized depolarization of the membrane, leading to (step 3) opening of voltage-gated Na^+ channels and generation of an action potential. When the spreading depolarization reaches transverse tubules, it is sensed by voltage-gated Ca^{2+} channels in the plasma membrane. Through an unknown mechanism (indicated as a "?") these channels remain closed but influence Ca^{2+} channels in the sarcoplasmic reticulum membrane (a network of membrane-bound compartments in muscle), which release stored Ca^{2+} into the cytosol (step 4). The resulting rise in cytosolic Ca^{2+} causes muscle contraction by mechanisms discussed in Chapter 17.

Careful monitoring of the membrane potential of the muscle membrane at a synapse with a cholinergic motor neuron has demonstrated spontaneous, intermittent, and random ~2 ms depolarizations of about 0.5–1.0 mV in the absence of stimulation of the motor neuron. Each of these depolarizations is caused by the spontaneous release of acetylcholine from a single synaptic vesicle in the neuron. The release of one acetylcholine-containing synaptic vesicle results in the opening of about 3000 ion channels in the postsynaptic membrane, far short of the number needed to reach the threshold depolarization that induces an action potential. Clearly stimulation of muscle contraction by a motor neuron requires the nearly simultaneous release of acetylcholine from numerous synaptic vesicles.

All Five Subunits in the Nicotinic Acetylcholine Receptor Contribute to the Ion Channel

The excitatory nicotinic acetylcholine receptor, found at many nerve-muscle synapses, was the first ligand-gated ion channel to be purified, cloned, and characterized at the molecular level and provides a paradigm for other neurotransmitter-gated ion channels. The acetylcholine receptor from skeletal muscle is a pentameric protein with a subunit composition of $\alpha_2\beta\gamma\delta$. These four different subunit types have considerable sequence homology with each

other; on average, about 35–40 percent of the residues in any two subunits are similar, suggesting that they all derived from a common ancestral gene. The complete receptor has fivefold symmetry, and the actual cation channel is a tapered central pore lined by homologous segments from each of the five subunits (Figure 23-31). The channel opens when the receptor cooperatively binds two acetylcholine molecules to sites located at the interfaces of the αδ and αγ subunits, as shown in Figure 23-31a. Once acetylcholine is bound to a receptor, the channel is opened within a few microseconds. Studies measuring the receptor's permeability to different small cations suggest that the open ion channel is about 0.65–0.80 nm in diameter at its narrowest, in agreement with estimates from electron micrographs. This would be sufficient to allow passage of both Na^+ and K^+ ions with their shell of bound water molecules.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

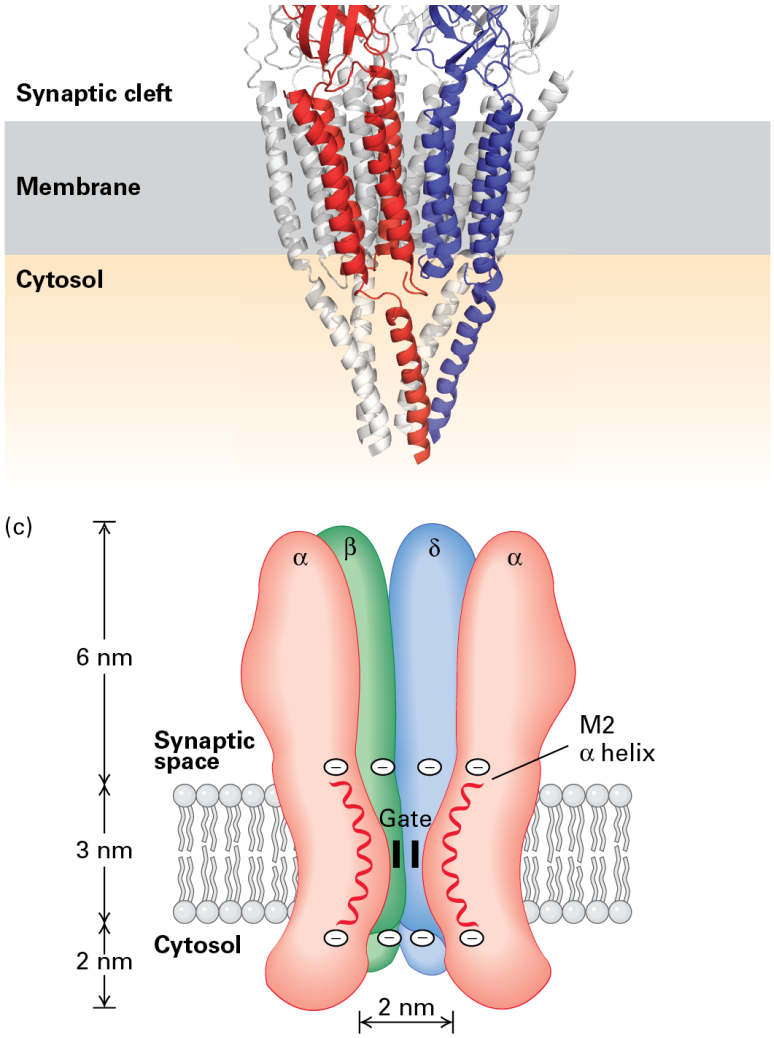


FIGURE 23-31 Three-dimensional structure of the nicotinic acetylcholine receptor. Three-dimensional molecular structure of the *Torpedo* nicotinic acetylcholine receptor as viewed (a) from the synaptic cleft and (b) parallel to the plane of the membrane. For clarity, only the front two subunits, α and γ, are highlighted in (b) (colors: α, red; β, green; γ, blue; δ, light blue). The two acetylcholine-binding sites are located about 3 nm from the membrane surface and are highlighted in yellow; only the one at the α-γ interface is shown in panel (b). (c) Schematic cutaway model of the pentameric receptor in the membrane. Each subunit has four membrane-spanning α helices, M1–M4; the M2 α helix (red) faces the central pore. Aspartate and glutamate side chains form two rings of negative charges, one at each end of the M2 helices, that help exclude anions from and attract cations to the channel. The gate, which is opened by binding of acetylcholine, lies within the pore.

[Data from N. Unwin, 2005, *J. Mol. Biol.* **346**:967–989, PDB ID 2bg9.]

We have described the neuromuscular junction as an excellent example of how neurotransmitters and their receptors work. Like acetylcholine, glutamate, a principal neurotransmitter used for signaling between neurons in the vertebrate brain, uses two main types of receptors. One class, termed *ionotropic glutamate receptors*, are ligand-gated ion channels that allow the flow of K^+, Na^+, and sometimes Ca^{2+} in response to glutamate binding and that work along the same principles as AChR. Glutamate also binds to a second class of receptors, coupled to G proteins. Later in this chapter, we will see how such G protein–coupled receptors (GPCRs) and ion channels function as receptors for **odorants** (molecules that activate odor receptors) and **tastants** (molecules that activate taste receptors).

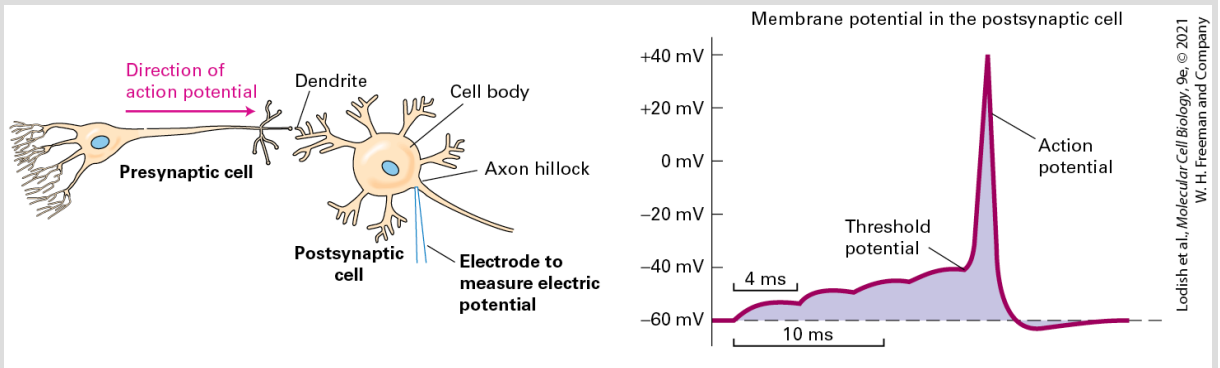
Nerve Cells Integrate Many Inputs to Make an All-or-None Decision to Generate an Action Potential

At the neuromuscular junction, virtually every action potential in the presynaptic motor neuron triggers an action potential in the postsynaptic muscle cell that propagates along the muscle fiber. The situation at synapses between neurons, especially those in the brain, is much more complex because the postsynaptic neuron commonly

receives signals from many presynaptic neurons. The neurotransmitters released from presynaptic neurons may bind to an **excitatory receptor** on the postsynaptic neuron, thereby opening a channel that admits Na^+ ions or both Na^+ and K^+ ions. The acetylcholine and glutamate receptors just discussed are examples of excitatory receptors, and opening of such ion channels leads to depolarization of the postsynaptic plasma membrane, promoting generation of an action potential. In contrast, binding of a neurotransmitter to an **inhibitory receptor** (inhibiting action potential firing) on the postsynaptic cell causes opening of K^+ or Cl^- channels, leading to an efflux of additional K^+ ions or an influx of Cl^- ions. In either case, the ion flow tends to hyperpolarize the plasma membrane, which inhibits generation of an action potential in the postsynaptic cell.

A single neuron can be affected simultaneously by signals received at multiple excitatory and inhibitory synapses. The neuron continuously integrates these signals and determines whether or not to generate an action potential. In this process, the various small depolarizations and hyperpolarizations generated at synapses move along the plasma membrane from the dendrites to the cell body and then to the axon hillock, where they are integrated. An action potential is generated whenever the membrane at the axon hillock becomes depolarized to a certain voltage, which can be different for different neurons, called the *threshold potential* (Figure 23-32). Thus an action potential is generated in an all-or-nothing fashion:

depolarization to the threshold always leads to an action potential, whereas any depolarization that does not reach the threshold potential never induces it.



EXPERIMENTAL FIGURE 23-32 Incoming signals must reach the threshold potential to trigger an action potential in a postsynaptic neuron. In this example, the presynaptic neuron is generating about one action potential every 4 ms. Arrival of each action potential at the synapse causes a small change in the membrane potential at the axon hillock of the postsynaptic neuron, in this example a depolarization of ~5 mV. When multiple stimuli cause the membrane of this postsynaptic cell to become depolarized to the threshold potential, here approximately 40 mV, an action potential is induced.

Whether a neuron generates an action potential in the axon hillock depends on the balance of the timing, amplitude, and localization of all the various inputs it receives; this signal computation differs for each type of neuron. In a sense, each neuron is a tiny analog-to-digital computer that averages all the receptor activations and electrical disturbances on its membrane (analog) and makes a decision whether or not to trigger an action potential (digital) and conduct it down the axon. An action potential will always have the same magnitude in any particular neuron. As we have noted, the

frequency with which action potentials are generated in a particular neuron is the important parameter in its ability to signal other cells.

Gap Junctions Allow Direct Communication Between Neurons and Between Glia

Chemical synapses employing neurotransmitters allow one-way communication at reasonably high speed. However, sometimes signals go from cell to cell electrically, without the intervention of chemical synapses. *Electrical synapses* depend on gap-junction channels that link two cells together (see Chapter 20). The effect of gap-junction connections is to perfectly coordinate the activities of joined cells. An electrical synapse is usually *bidirectional*; either neuron can excite the other. Electrical synapses are common in the neocortex and thalamus, among others. The key feature of electrical synapses is their speed. While it takes about 0.5–5 ms for a signal to cross a chemical synapse, transmission across an electrical synapse is almost instantaneous, on the order of a fraction of a millisecond, since the cytoplasm is continuous between the cells. In addition, the presynaptic cell (the one sending the signal) does not have to reach a threshold at which it can cause an action potential in the postsynaptic cell. Instead, any electrical current continues into the next cell and causes depolarization in proportion to the current.

Gap junctions can also form between glial cells and between neurons and glia. Astrocytes in the brain are connected to one another through gap junctions, which gives rise to the generation of waves of Ca^{2+} that propagate through networks of astrocytes in the brain at a speed of 1 μm/sec. Gap junctions also form within individual Schwann cells, forming connections between the layers of myelin formed by a single Schwann cell. These gap junctions are thought to facilitate the passage of metabolites and ions between myelin layers.

An electrical synapse may contain thousands of gap-junction channels, each composed of two hemichannels, one in each apposed cell. Gap-junction channels in the neuron have a structure similar to conventional gap junctions (see Figure 20-21). Each hemichannel is an assembly of six copies of the connexin protein. Since there are about 20 mammalian connexin genes, diversity in channel structure and function can arise from the different protein components. The 1.6–2.0 nm channel itself allows the diffusion of molecules up to about 1000 Da in size and has no trouble at all accommodating ions.

KEY CONCEPTS OF SECTION 23.3

Communication at Synapses

- Synapses are the junctions between a presynaptic cell and a postsynaptic cell and are the site of communication between neurons (see Figure 23-3).

- Synapse formation is mediated by interactions between presynaptic axonal compartments and postsynaptic dendritic compartments. Cell-adhesion molecules keep the cells aligned. At the neuromuscular junction, motor neurons induce the accumulation of acetylcholine receptors in the postsynaptic muscle plasma membrane close to the forming axon terminus (see Figure 23-23).
- In presynaptic cells, low-molecular-weight neurotransmitters (e.g., acetylcholine, dopamine, epinephrine) are imported from the cytosol into synaptic vesicles by H^+-linked antiporters. V-class proton pumps maintain the low intravesicular pH that drives neurotransmitter import against a concentration gradient.
- Neurotransmitters are stored in hundreds to thousands of synaptic vesicles in the axon termini of the presynaptic cell. When an action potential arrives there, voltage-sensitive Ca^{2+} channels open and the calcium causes synaptic vesicles to fuse with the plasma membrane, releasing neurotransmitter molecules into the synapse (see Figure 23-27, step **4**).
- Neurotransmitters diffuse across the synapse and bind to receptors on the postsynaptic cell, which can be a neuron or a muscle cell. Chemical synapses of this sort are unidirectional (see Figure 23-3).
- Synaptic vesicles fuse with the plasma membrane using cellular machinery that is standard for exocytosis, including SNAREs and SM proteins. Synaptotagmin protein is the calcium sensor that detects the action potential–stimulated rise in calcium that leads to the fusion (see Figure 23-29). RIM and RIM-BP tether voltage-gated Ca^{2+} channels to the release machinery, ensuring fast coupling between action potentials and neurotransmitter release.
- Following neurotransmitter release from the presynaptic cell, vesicles are re-formed by endocytosis and recycled (see Figure 23-27, step **6**).
- Coordinated operation of four gated ion channels at the synapse of a motor neuron and a striated muscle cell leads to release of acetylcholine from the axon terminus, depolarization of the muscle membrane, generation of an action potential, and subsequent muscle contraction (see Figure 23-30).
- Neurotransmitter receptors fall into two classes: ligand-gated ion channels, which permit ion passage when open, and G protein–coupled receptors (GPCRs), which are linked to separate ion channels.
- A postsynaptic neuron generates an action potential only when the plasma membrane at the axon hillock is depolarized to the threshold potential by the summation of small depolarizations and hyperpolarizations caused by activation of multiple neuronal receptors (see Figure 23-32).
- Electrical synapses are direct gap-junction connections between neurons and between glia. Electrical synapses are extremely fast in signal transmission and are usually bidirectional.

23.4 Sensing the Environment: Touch, Pain, Taste, and Smell

Our bodies are constantly receiving signals from our environment: light, sound, smells, tastes, mechanical stimulation, heat, and cold. Our perception of these signals is mediated by the brain. In recent years dramatic progress has been made in understanding how our senses record impressions of the outside world and how the brain processes that information. For example, in Chapter 15 we described the functions of one of the two types of photoreceptors in the human retina, the *rods*, and learned how they serve as primary recipients of visual stimulation. Rods are stimulated by weak light, like moonlight, over a range of wavelengths, while the other photoreceptors, the *cones*, mediate color vision in the presence of bright light. Their photoreceptors synapse on layer upon layer of interneurons that are innervated by different combinations of photoreceptor cells. These signals are processed and interpreted by the part of the brain called the *visual cortex,* where these nerve impulses are translated into an image of the world around us.

In this section, we discuss the cellular and molecular mechanisms and specialized nerve cells underlying several of our other senses: touch, pain, taste, and smell. We see how two broad classes of receptors, ion channels and G protein–coupled receptors, function in these sensing processes. As with vision, multiple interneurons connect these sensory cells with the brain, where relayed signals are

converted into perceptions of the environment. New technologies, such as optogenetics, have been especially powerful tools in mapping these circuits. In the case of smell, each sensory neuron expresses a single odorant receptor, and we shall see how multiple sensory neurons that express the same receptor activate the same brain center. The connections between odorant binding and perception by the brain are thus direct and fairly well understood.

Mechanoreceptors Are Gated Cation Channels

Our skin, especially the skin of our fingers, is highly specialized for collecting sensory information. Our whole body, in fact, has numerous **mechanosensors** embedded in its various tissues. These sensors frequently make us aware of touch, the positions and movements of our limbs or head (proprioception), pain, and temperature, though we often go through periods when we ignore these inputs. Mammals use different sets of receptor cells to report on touch, temperature, and pain. These mechanosensory receptors are located at the terminals of a class of bipolar sensory neurons called *dorsal root ganglion cells*. The cell bodies of dorsal root ganglion cells are located in the dorsal root ganglion, adjacent to the spinal cord (see Figure 23-4), and the neurons extend an axon that bifurcates into a peripheral branch that innervates the skin and contains the mechanosensory receptors, and a central branch that

projects to the spinal cord or brain stem to relay sensory signals for processing.

Many mechanosensory receptors are Na^+ or Na^+/Ca^{2+} channels that are gated, or opened, in response to specific stimuli; activation of such receptors causes an influx of Na^+ or both Na^+ and Ca^{2+} ions, leading to membrane depolarization. Examples include the stretch and touch receptors that are activated by stretching of the cell membrane; these have been identified in a wide array of cells, ranging from vertebrate muscle and epithelial cells to yeast, plants, and even bacteria.

The cloning of genes encoding touch receptors began with the isolation of mutant strains of *C. elegans* that were insensitive to gentle body touching. Four of the genes in which mutations were isolated, *MEC-2, MEC-4, MEC-6*, and *MEC-10*, encode two subunits (MEC-4 and MEC-10) and two accessory proteins (MEC-2 and MEC-6) of a Na^+ channel in the touch-receptor cells. Studies on worms with mutations in these genes showed that these channels are necessary for transduction of a gentle body touch; biophysical studies indicated that these channels likely open directly in response to mechanical stimulation (Figure 23-33). The touch-sensitive complexes contain several other proteins essential for touch sensitivity, including subunits of novel 15-protofilament microtubules in the cytosol and specific proteins in the extracellular matrix. Similar kinds of channels are found in bacteria and lower

eukaryotes; by opening in response to membrane stretching, these channels may play a role in osmoregulation and the control of a constant cell volume.

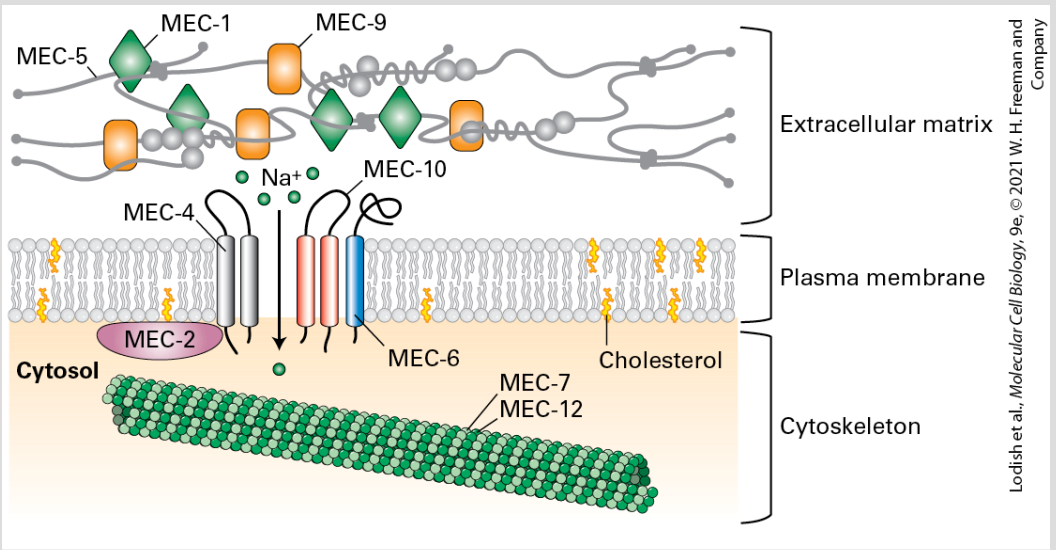


FIGURE 23-33 The MEC-4 touch-receptor complex in *C. elegans* consists of a Na^+ channel and extracellular matrix proteins. The MEC-4 and MEC-10 proteins are the pore-forming subunits of the Na^+ channel; MEC-2 and MEC-6 are accessory subunits that enable channel activity. Mechanotransduction also requires a specialized extracellular matrix, consisting of MEC-5, a collagen isoform, and MEC-1 and MEC-9, both proteins with multiple EGF repeats. MEC-7 and MEC-12 are tubulin monomers that form novel 15-protofilament microtubules that are also required for touch sensitivity. See E. A. Lumpkin, K. L. Marshall, and A. M. Nelson, 2010, *J. Cell Biol.* **191**:237.

Two novel channels were discovered in 2010 that directly convert mechanical stimuli into cation conductance in mammalian cells, called *Piezo1* and *Piezo2* (from the Greek word *piesi*, which means "pressure"). Both form large cation-selective channels composed of three identical subunits, with each subunit containing 38

membrane-spanning domains, creating a channel that has 114 transmembrane segments! Expression of Piezo1 or Piezo2 induces mechanosensitive-cation currents in these cells. This can be assayed by expressing the channels in cell culture, and using calcium imaging to monitor the response of the cells to stretch induced by poking the cells with a small glass pipette (Figure 23-34). Reduction of Piezo2 expression in dorsal root ganglion sensory neurons in mice reduced their mechanosensitivity, and knockout of the single Piezo homolog in *Drosophila* resulted in flies with severely reduced behavioral responses to noxious mechanical stimuli. Together, these experiments show that Piezo channels mediate mechanical signal transduction.

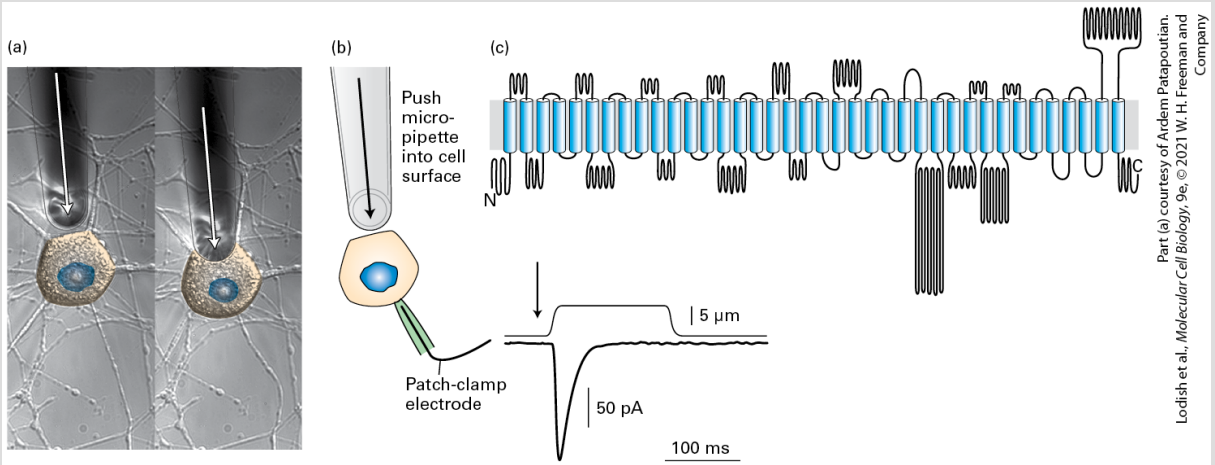


FIGURE 23-34 Piezo channels are mechanotransducers. (a) To identify channels that transduced mechanical information, cDNAs encoding transmembrane proteins were expressed in heterologous cells, and the response to mechanical perturbation with a glass pipette was determined by patch-clamp recording (shown) or by calcium imaging (not shown). (b) When Piezo1 or Piezo2 cDNAs are expressed in cultured cells, poking the cell with a glass pipette elicits a strong inward current. (c) Piezo1 and 2 form homotrimeric cation channels. Each subunit is extremely large, with over 2000 amino acid residues and 38 transmembrane domains. The assembled channel thus contains 114 membrane spanning

domains and has a mass of over 1.2 million daltons (about the same mass as the small subunit of the ribosome)!

Pain Receptors Are Also Gated Cation Channels

Animals as diverse as snails and humans sense noxious events (the process termed *nociception*); pain receptors, called **nociceptors**, respond to mechanical change, heat, and certain toxic chemicals. Pain serves to alert us to events such as tissue damage that are capable of producing injury and evokes behaviors that promote tissue healing. Persistent pain in response to tissue injury is common, and many individuals suffer from chronic pain. Thus understanding both acute and chronic pain is a major research goal, as is the development of new types of drugs to treat pain.

One of the first mammalian pain receptors to be cloned and identified was TRPV1, a Na^+/Ca^{2+} channel that is found in many sensory pain neurons of the peripheral nervous system and is activated by a wide variety of exogenous and endogenous physical and chemical stimuli. The best known activators of TRPV1 are heat greater than 43 °C, acidic pH, and capsaicin, the molecule that makes chili peppers hot. Activation of TRPV1 receptors leads to painful, burning sensations. Numerous TRPV1 antagonists have been developed by pharmaceutical companies as possible pain

medications. However, a major side effect that has limited the utility of these drugs is that they result in an elevation in body temperature; this suggests that one function of TRPV1 is to sense and regulate body temperature and that the drugs inhibit this function. ■

In a landmark study, scientists used single-particle cryoelectron microscopy (cryoEM, see Chapter 3 and 11) to obtain a high-resolution (0.34 nm) model of the rat TRPV1 channel in the closed configuration and in two open configurations, one bound to capsaicin and the other bound to two potent TRPV1 activators, one from plant and the other from spider venom. As shown in Figure 23-35, these studies revealed that the TRPV1 channel structure is similar to that of voltage-gated ion channels (see Figure 23-13), composed of four symmetrical subunits with six transmembrane helices (S1–S6) each. However, the charged amino acids in S1–S4 that function as voltage sensors in voltage-gated ion channels are replaced by aromatic residues in TRPV1. This stabilizes the channel core so that instead of moving like voltage sensors upon depolarization, the TRPV S1–S4 helices provide an anchor for movements within the pore that are triggered by ligand binding. Two constrictions, or gates, were identified in the pore region. The spider toxin bound to the extracellular surface of the channel, near the pore helix, and locked open the extracellular end of the channel. Capsaicin and the plant toxin bound to a site deep within the membrane toward the cytoplasmic end of the pore, with binding increasing the diameter of the pore. These findings indicate that the

TRPV channel undergoes dual gating and is able to respond to diverse ligands.

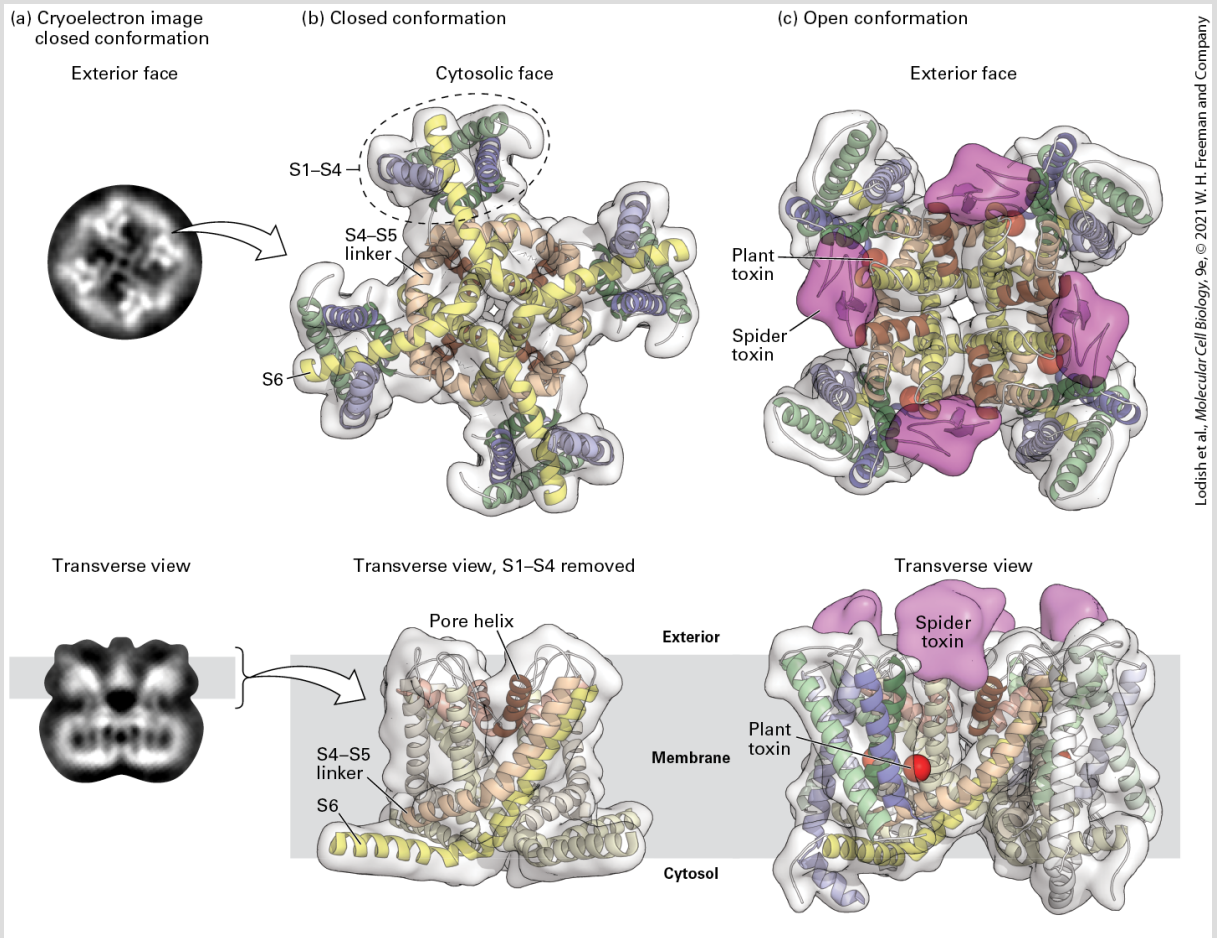


FIGURE 23-35 Single-particle cyroelectron microscopy high-resolution structure of the TRPV1 channel demonstrates how this channel is activated by various chemical stimuli. The high-resolution structure of the rat TRPV1 channel was obtained by single-particle cryoelectron microscopy at 0.34 nm resolution. (a) Photomicrographs of the two-dimensional structure of the tetrameric TRPV1 channel embedded in a thin layer of vitreous ice, with a face view of the channel in the top panel, and a side view in the bottom panel. (b, *top*) Ribbon diagram of a bottom view of the channel that focuses on the S1−S4 transmembrane domains, and the pore domain formed by S5 and S6, together with linking pore (P) loops. The S1−S4 domain is similar in structure to the voltage-sensing domains in the voltage-gated K^+ and Na^+ channels (see Figure 23-14) but differ in that they do not move. (b, *bottom*) Ribbon diagram of the side view of the channel in the closed conformation, focusing on the pore domain that is formed by S5-P-S6. (c) The open

conformation was stabilized by incubating the channel with two agonists, a spider toxin (in magenta) and a plant toxin (in red). Cryoelectron density maps reveal that the spider toxin (magenta) binds to external domains of the channel, linking two subunits of the channel together via its two globular cysteine-knot domains, while the plant toxin (red) binds to a region deep within the pore. Capsaicin binds to the same sites as the plant toxin (not shown). Binding of agonists to two distinct sites indicates that the TRPV1 channel is dually gated, allowing for significant modulation of channel function.

[Part (a) republished with permission from Nature, from M. Liao et al., 2013, "Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy," *Nature* **504**:107–112; permission conveyed through Copyright Clearance Center, Inc. Part (b) data from M. Liao et al., 2013, *Nature* **504**:107–112, PDB ID 3j5p. Part (c) data from E. Cao et al., 2013, *Nature* **504**:113–118, PDB ID 3j5q.]

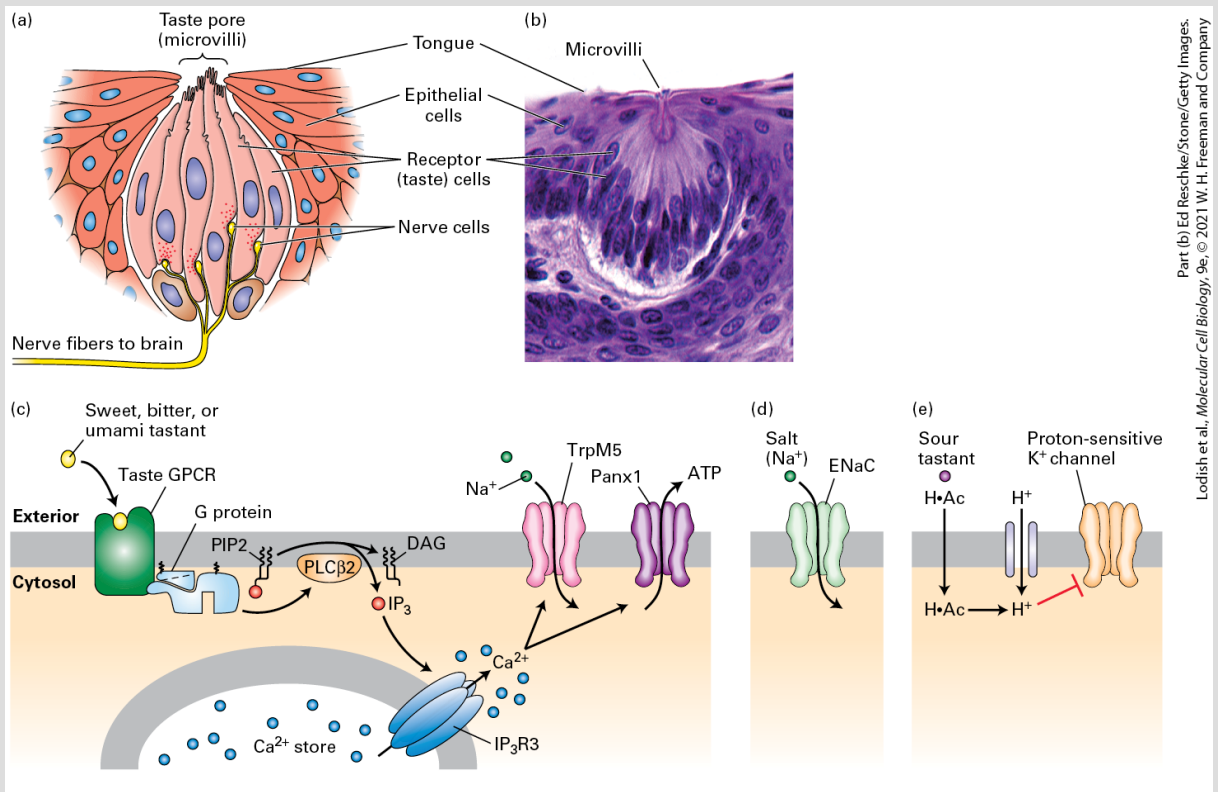
Five Primary Tastes Are Sensed by Subsets of Cells in Each Taste Bud

We taste many chemicals, all of which are hydrophilic and nonvolatile molecules dissolved in saliva. Like the other senses, that of taste likely evolved to increase an animal's chance of survival. Many toxic substances taste bitter or acidic, and nourishing foods are broken down into molecules that taste sweet (e.g., sugars), salty, or umami (e.g., the meaty or savory taste of monosodium glutamate and other amino acids). Animals (including humans) can never be certain exactly what enters their mouth; the sense of taste enables an animal to make a quick decision — eat it, or get rid of it. Taste is less demanding of the nervous system than olfaction, because fewer types of molecules are monitored. What is impressive is the

sensitivity of taste; bitter molecules can be detected at concentrations as low as 10^{-12} M.

All tastes are sensed on all areas of the tongue, and selective cells respond preferentially to certain tastes. There are receptors for salty, sweet, sour, umami, and bitter tastes in all parts of the tongue. The receptors are of two different types: channel proteins for salty and sour tastes and seven-transmembrane-domain proteins (G protein–coupled receptors) for sweetness, umami, and bitterness. Specific membrane receptors that detect fatty acids are present on taste bud cells, and fatty taste is considered by some to be a sixth basic taste quality.

Taste buds are located in bumps in the tongue called *papillae*; each bud has a pore through which fluid carries solutes inside. Each taste bud has about 50–100 taste cells (Figure 23-36a, b), which are epithelial cells but with some of the functions of neurons. Microvilli on the taste cells' apical tips bear the **taste receptors**, directly contacting the external environment in the oral cavity and thus experiencing wide fluctuations in food-derived molecules as well as the presence of potentially harmful compounds. Cells in the tongue and other parts of the mouth are subjected to a lot of wear and tear, and taste bud cells are continuously replaced by cell divisions in the underlying epithelium.



Part (b) Ed Reschke/Stone/Getty Images.
Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 23-36 The sense of taste is mediated by chemical signals. (a) The taste cells (pink) in a mammalian taste bud contact the nerve cells (yellow). The chemical signals arrive at the microvilli seen at the top. (b) Micrograph of a mammalian taste bud, showing the receptor cells. The microvilli are barely visible at the top of the taste bud, indicated by the label. (c) Sweet, bitter, and umami ligands bind specific taste GPCRs expressed in type II receptor cells, activating a phosphoinositide pathway that elevates cytosolic Ca^{2+}. Ca^{2+} in turn binds to and opens a Ca^{2+}-gated Na^+ channel, TrpM5, leading to an influx of Na^+ and membrane depolarization. The combined action of elevated Ca^{2+} and membrane depolarization opens the large pores of a membrane channel termed *Panx1*, resulting in release of ATP and probably other signaling molecules into the extracellular space. ATP and these other molecules stimulate the nerve cells that will ultimately carry the information to the brain. (d) Salt is detected by direct permeation of Na^+ ions through membrane ion channels, including the ENaC channel, directly depolarizing the plasma membrane. (e) Organic acids like acetic acid diffuse in their protonated form (H·Ac) through the plasma membrane and dissociate into an anion and proton, acidifying the cytosol. Entry of strong acids like HCl is facilitated by a proton channel in the apical membrane of the sour-sensing cells that enables protons to reach the cytosol. Intracellular H^+ is believed to block a proton-sensitive K^+ channel (as yet unidentified) and thus depolarize the membrane. Voltage-gated Ca^{2+} channels would open, leading to an elevation in cytosolic Ca^{2+} that

triggers exocytosis of synaptic vesicles that are not depicted. See N. Chaudhari and S. D. Roper, 2010, *J. Cell Biol.* **190**:285; and S. Frings, 2010, *Proc. Nat'l Acad. Sci. USA* **107**:21955.

Reception of a taste signal causes cell depolarization that triggers action potentials; these in turn cause Ca^{2+} uptake through voltage-dependent Ca^{2+} channels and release of neurotransmitters (Figure 23-36 c–e). Taste cells do not grow axons; instead, they signal over short distances to adjacent neurons. These neurons convey the information about taste through multiple connections to a region of the cortex that is specialized for taste, called the *insula*. To determine how the insula knows that a salty taste receptor as opposed to a sweet taste receptor has been activated, scientists performed two-photon imaging (see Chapter 4) of insula in mice after presentation of a specific tastant. They used calcium indicators to detect the neurons that were activated and in this way were able to monitor the activation of large numbers of neurons as calcium-dependent increases in fluorescence. These experiments revealed that four of the tastes — sweet, bitter, umami, and salty — are represented in separate, nonoverlapping regions within the insula, thereby demonstrating the existence of a gustotopic map in the brain that mediates our representation of taste.

Bitter Taste

Bitter tastants are diverse and are detected by a family of about 25–30 different G protein–coupled receptors (GPCRs) known as

T2Rs. As depicted in Figure 23-36c, all of these GPCRs activate a particular Gα isoform, called *gustducin*, which is expressed only in taste cells. However, it is the released ubiquitous Gβγ subunit of the heterotrimeric G protein that binds to and activates a specific isoform of phospholipase Cβ, which in turn generates IP_3. IP_3 triggers Ca^{2+} release from the endoplasmic reticulum (see Figure 15-28). Ca^{2+} in turn binds to and opens a Ca^{2+}-gated Na^+ channel, TRPM5, a member of the TRP family of ion channels, leading to an influx of Na^+ and membrane depolarization. The combined action of elevated Ca^{2+} and membrane depolarization opens the large pores of a membrane channel termed *Panx1*, resulting in release of ATP and probably other signaling molecules into the extracellular space. ATP is then thought to act as a neurotransmitter to stimulate the nerve cells that will ultimately carry the taste information to the brain.

Different bitter taste molecules are quite distinct in structure, which probably accounts for the need for the diverse family of T2Rs. Some T2Rs bind only 2–4 bitter-tasting compounds, whereas others bind a wider variety of bitter compounds. The first member of the T2R family to be identified came from human genetics studies that showed an important bitterness-detection gene, the *TAS2R38* gene, on chromosome 5. Mice that have five amino acid changes in the T2R protein T2R5 are unable to detect the bitter taste of cycloheximide (a protein synthesis inhibitor). Multiple T2R types are

often expressed in the same taste cell, and about 15 percent of all taste cells express T2Rs.

A clever gene regulation swap experiment was done to demonstrate the role of T2R proteins. Mice were engineered to express a bitter-taste receptor, a T2R protein, in cells that normally detect sweet tastants that attract mice. The mice developed a strong attraction for bitter tastes, evidently because the cells continued to send a "go and eat this" signal even though they were detecting bitter tastants. This experiment demonstrates that the specificity of taste cells is determined within the cells themselves and that the signals they send are interpreted according to the neural connections made by that class of cells. It implies that the T2R-expressing sweet taste receptors were wired to the region of the insula that receives bitter signals and that it was thus represented as a sweet taste.

Sweet and Umami Tastes

Sweet and umami tastants are detected by a GPCR family called the *T1Rs*, which are related to the T2Rs and that also transduce signals through a phosphoinositide signaling pathway. The three mammalian T1Rs differ from one another in a small number of amino acids. The T1Rs have very large extracellular domains that comprise the taste-binding domain of the protein. In the taste-sensing glutamate receptor, the extracellular domain closes around glutamate in a way that is described as analogous to a Venus flytrap

(see Chapter 15). Unlike most GPCRs, which generally function as monomers, T1Rs form homodimers and heterodimers, which is thought to increase the repertoire of molecules that can act as signals. However, the code of responses to different molecules is still under investigation. Mice lacking T1R2 or T1R3 fail to detect sugar; it is thought that the actual receptor is a heterodimer of the two. T1R3 appears to be a receptor for both sweet tastes and umami and that is because it detects sweets when combined with T1R2 and umami when combined with T1R1. Accordingly, taste cells express T1R1 or T1R2 but not both, as otherwise they would send an ambiguous message to the brain.

Interestingly, sweet-taste receptors are also found on the surface of certain endocrine cells in the gut; these cells also express gustducin and several other taste transduction proteins. The presence of glucose in the gut causes these cells to secrete the hormone *g*lucagon-*l*ike *p*eptide-1 (GLP-1), which in turn regulates appetite, and enhances insulin secretion and gut motility. Thus certain cells of the gut "taste" glucose through the same mechanisms used by taste cells of the tongue.

Salty Taste

The taste of salt is elicited by a wide range of Na^+ concentrations, from 10 mM to 500 mM. Salt is sensed by a member of a family of Na^+ channels called *ENaC channels* (Figure 23-36d). Indeed,

knocking out a critical ENaC subunit in taste buds impaired salty-taste detection in mice. The influx of Na^+ through the channel depolarizes the taste cell, leading to neurotransmitter release. The role of ENaC channels as salt sensors is evolutionarily ancient; ENaC proteins also detect salt when expressed in insects. In *Drosophila*, taste sensors are located in multiple places including the legs, so when the fly steps on something tasty, the proboscis extends to explore it further.

Sour Taste

Perception of sourness is due to the detection of H^+ ions. One of the sour taste receptors was recently identified as the otopetrin-1 proton channel, a member of a family of proteins originally discovered for their importance in the vestibular system. Acids are detected by the otopetrin-1 proton channel in the apical membrane of the sour-sensing cells, triggering an increase in the intracellular H^+ concentration. Protons are believed to block a proton-sensitive K^+ channel in mammals and thus depolarize the membrane (Figure 23-36e). In both salt and sour taste detection, voltage-gated Ca^{2+} channels would then open, elevate cytosolic Ca^{2+} and thus trigger exocytosis of neurotransmitter-filled synaptic vesicles.

The Largest Group of G Protein–Coupled Receptors Detect Odors

The perception of volatile airborne chemicals imposes different demands than the perception of light, sound, touch, or taste. Light is sensed by only four rhodopsin molecules, tuned to different wavelengths. Sound is detected by mechanical effects through hairs that are tuned to different wavelengths. Touch and pain requires a small number of different gated ion channels. The sense of taste measures a small number of substances dissolved in water. In contrast to all these other senses, olfactory systems can discriminate between many hundreds of volatile molecules moving through air. Discrimination between a large number of chemicals is useful in finding food or a mate, sensing pheromones, and avoiding predators, toxins, and fires. **Olfactory receptors** work with enormous sensitivity. Male moths, for example, can detect single molecules of the signals sent drifting through the air by females. In order to cope with so many signals, the olfactory system employs a large family of olfactory receptor proteins. Humans have about 700 olfactory receptor genes, of which about half are functional (the rest are unproductive pseudogenes), a remarkably large fraction of the estimated 20,000 human genes. Mice are more efficient, with more than 1200 olfactory receptor genes, of which about 800 are functional. That means 3 percent of the mouse genome is composed of olfactory receptor genes. *Drosophila* has about 60 olfactory receptor genes. Odor molecules are called *odorants*. They have diverse chemical structures, so olfactory receptors face some of the same challenges faced by antibodies and hormone receptors, specifically, the need to bind and distinguish many variants of

relatively small molecules. In this section, we will examine how olfactory receptor genes are employed, and how the brain can recognize which odor has been sensed, which are the initial stages of interpretation of our chemical world.

Olfactory receptors are seven-transmembrane-domain proteins (Figure 23-37). In mammals, olfactory receptors are produced by cells of the nasal epithelium. These cells, called *olfactory receptor neurons (ORNs),* transduce the chemical signal into action potentials. Each ORN extends a single dendrite to the luminal surface of the epithelium, from which immotile cilia extend to bind inhaled odorants from the air (Figure 23-38a). These olfactory sensory cilia are enriched in the odorant receptors and signal transduction proteins that mediate the initial transduction events. In *Drosophila,* ORNs have similar structures and are located in the antennae (Figure 23-38b).

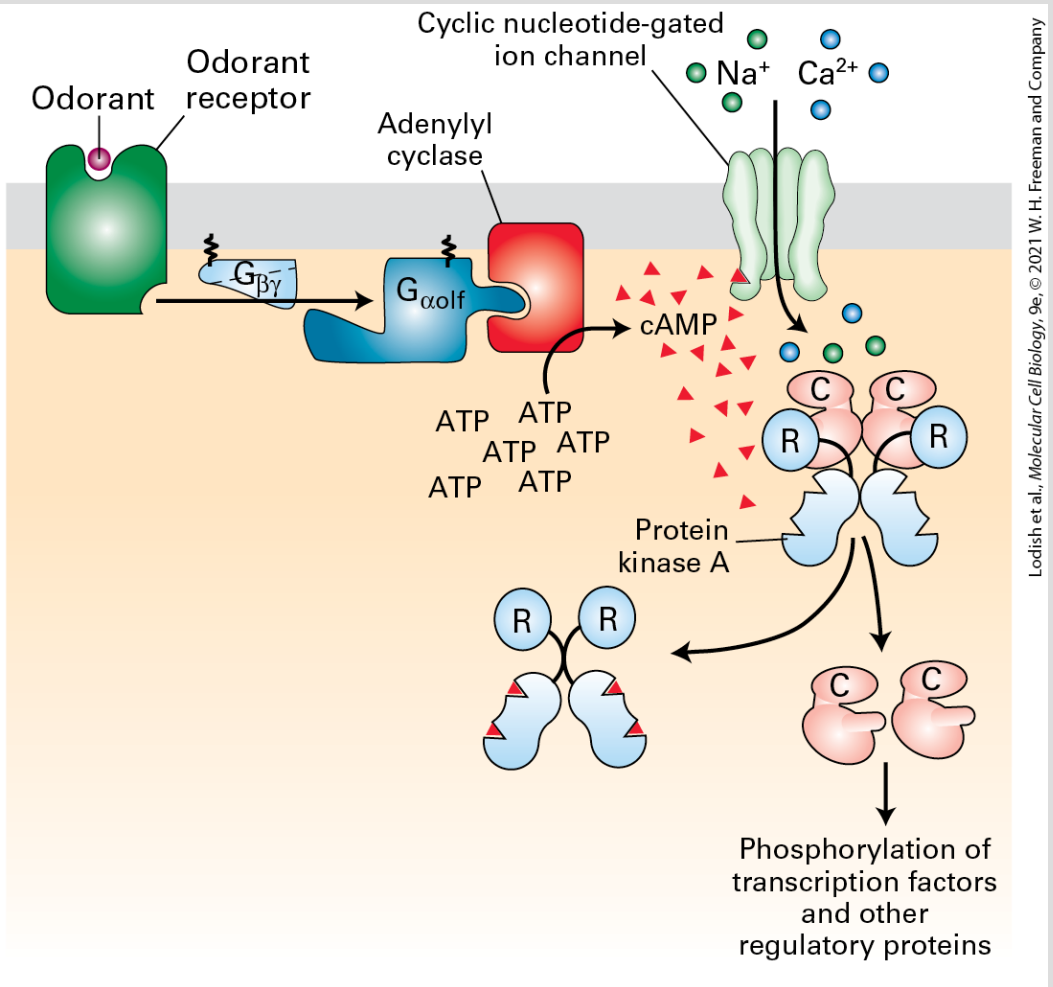


FIGURE 23-37 Signal transduction from the olfactory GPCRs. Binding of an odorant to its cognate odorant receptor (OR) triggers activation of the trimeric G protein $G_{\alpha olf}{\cdot}G_{\beta\gamma}$, releasing the active $G_{\alpha olf}{\cdot}GTP$. Activated $G_{\alpha olf}{\cdot}GTP$ in turn activates type III adenylyl cyclase (AC3), leading to the production of cyclic AMP (cAMP) from ATP. Molecules of cAMP bind to and open the cyclic nucleotide–gated (CNG) ion channel, leading to the influx of Na^+ and Ca^{2+} and depolarizing the cell. cAMP also activates protein kinase A (PKA), which phosphorylates and thus regulates transcription factors and other intracellular proteins.

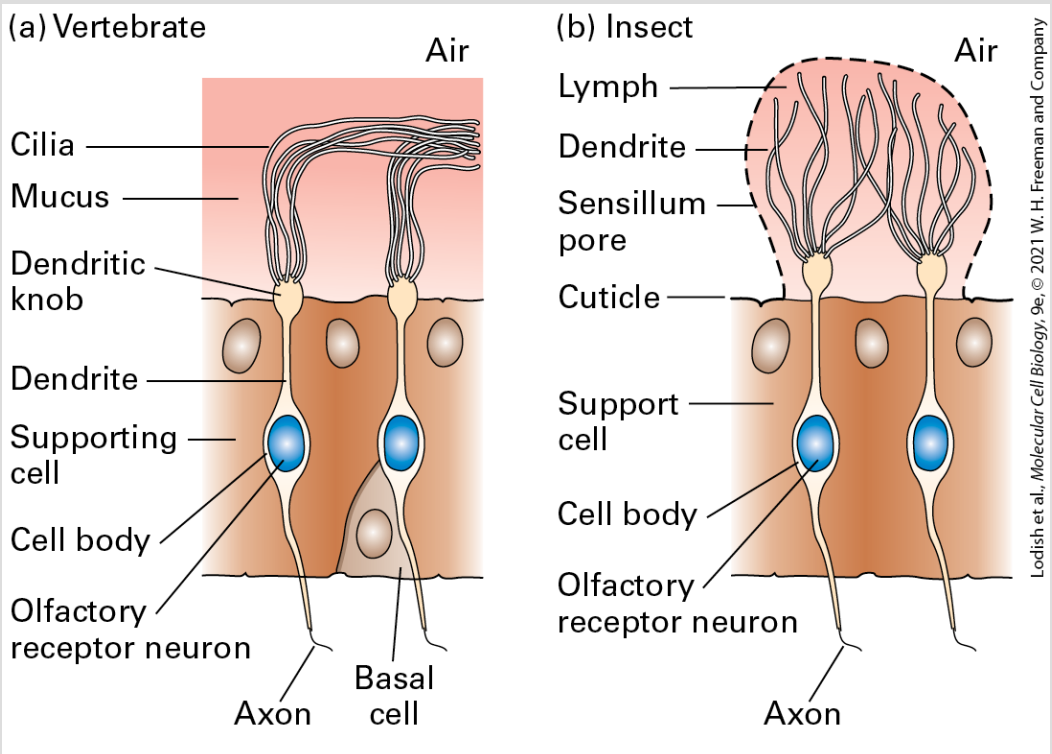
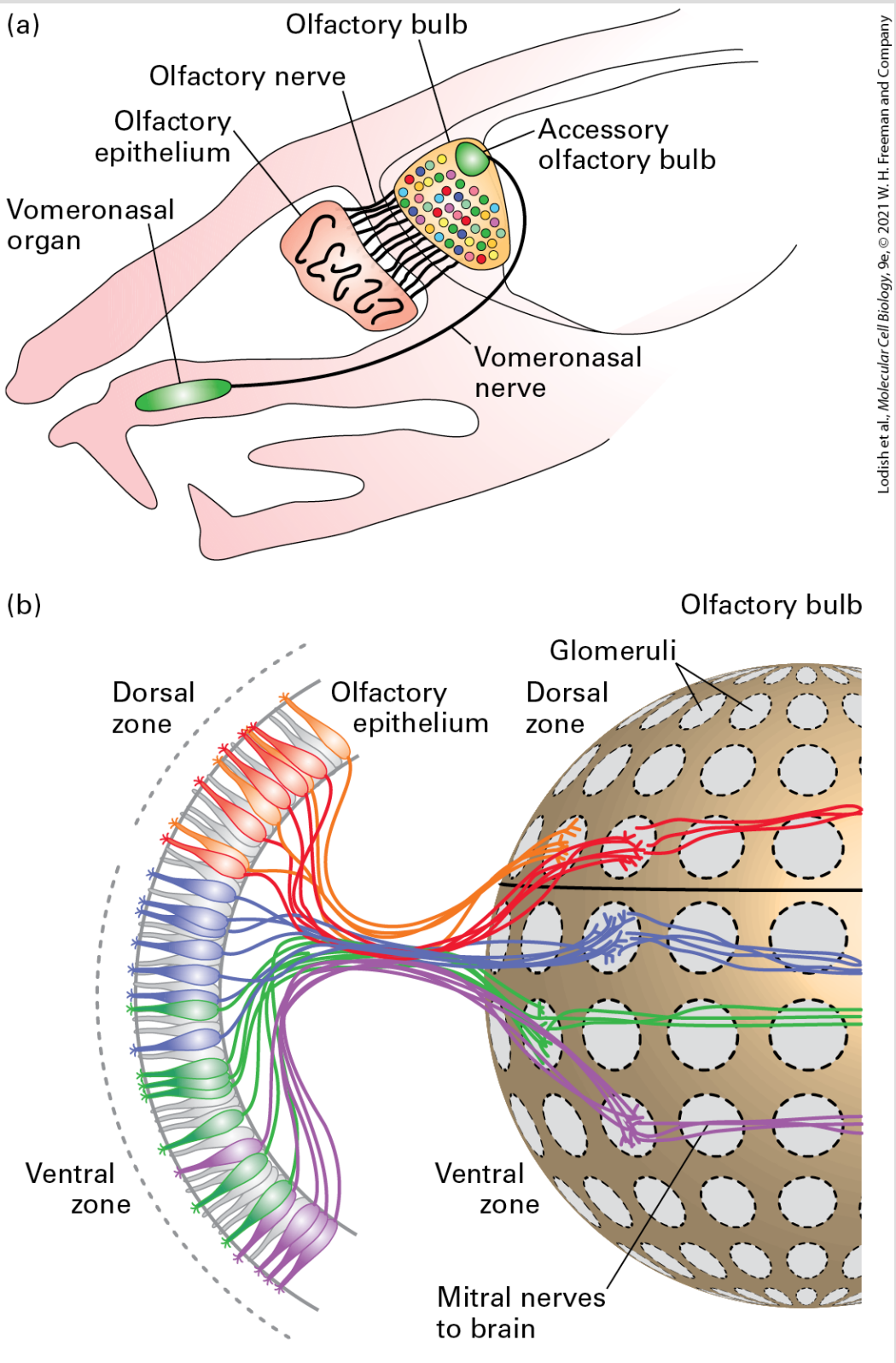


FIGURE 23-38 The structure of olfactory receptor neurons is evolutionarily conserved. (a) Vertebrate olfactory receptor neurons have one dendrite, which ends in a dendritic knob; from each dendritic knob, approximately 15 cilia extend into the nasal mucus. (b) Insect olfactory receptor neurons are morphologically similar: the bipolar neuron gives rise to a single basal axon that projects to an olfactory glomerulus in the antennal lobe. At its apical side it has a single dendritic process, from which sensory cilia extend.

[Data from U. B. Kaupp, 2010, *Nat. Rev. Neurosci*. **11**:188–200.]

In both mammals and *Drosophila*, the ORNs project their axons to the next higher level of the nervous system, which in mammals is located in the olfactory bulb of the brain. The ORN axons synapse with dendrites from *mitral neurons* in mammals (called *projection neurons* in insects); these synapses occur in the clusters of synaptic

structures called **glomeruli**. The mitral neurons connect to higher olfactory centers in the brain (Figure 23-39).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 23-39 The anatomy of olfaction in the mouse. (a) Schematic representation of a sagittal section through an adult mouse head. Axons of the

olfactory receptor neurons (ORNs) in the main olfactory epithelium bundle to form the olfactory nerve and innervate the olfactory bulb. Each ORN of the main olfactory epithelium expresses only one odorant receptor gene. The vomeronasal organ and the accessory olfactory bulb are involved in pheromone sensing. (b) All of the ORNs that express a single type of receptor send their axons to the same glomerulus. In this figure, each color represents the neural connections for each distinct expressed receptor. The glomeruli are located in the olfactory bulb near the brain; in the glomeruli, the ORNs synapse with mitral neurons; each mitral neuron has its dendrites localized to a single glomerulus and its corresponding ORNs, thus carrying information about a particular odorant to higher centers of the brain. Each glomerulus thus receives innervation from sensory neurons expressing a single odorant receptor, providing the anatomical basis of the olfactory sensory map. See T. Komiyama and L. Luo, 2006, *Curr. Opin. Neurobiol.* **16**:67–73; and S. DeMaria and J. Ngai, 2010, *J. Cell Biol.* **191**:443.

Humans vary markedly in their ability to detect certain odors. For example, some cannot detect the steroid androstenone, a compound derived from testosterone and found in human sweat. Some describe the odor as pleasant and musky, while others compare it to the smell of dirty socks. These differences are all ascribed to inactivating missense mutations in the gene encoding the single androstenone GPCR. Individuals with two copies of the wild-type allele perceive androstenone as unpleasant, whereas those possessing one or no functional alleles perceive androstenone as less unpleasant or undetectable.

Despite the vast number of olfactory receptors, all generate the same intracellular signals through activation of the same trimeric G protein: $G_{\alpha olf} \cdot G_{\beta\gamma}$ (see Figure 23-37). $G_{\alpha olf}$ is expressed mainly in

olfactory neurons. Like $G_{\alpha}s$, the active $G_{\alpha olf}$·GTP formed after ligand binding activates an adenylyl cyclase that leads to the production of cyclic AMP (cAMP; see Figure 15-20). Two downstream signaling pathways are activated by cAMP. It binds to a site on the cytosolic face of a cyclic nucleotide–gated (CNG) Na^+/Ca^{2+} channel, opening the channel and leading to an influx of Na^+ and Ca^{2+} and local depolarization of the cell membrane. This odorant-induced depolarization in the olfactory dendrites spreads throughout the neuronal membrane, resulting in opening of voltage-gated Na^+ channels in the axon hillock and the generation of action potentials. Molecules of cAMP also activate protein kinase A (PKA), which phosphorylates and thus regulates transcription factors and other intracellular proteins.

Each Olfactory Receptor Neuron Expresses a Single Type of Odorant Receptor

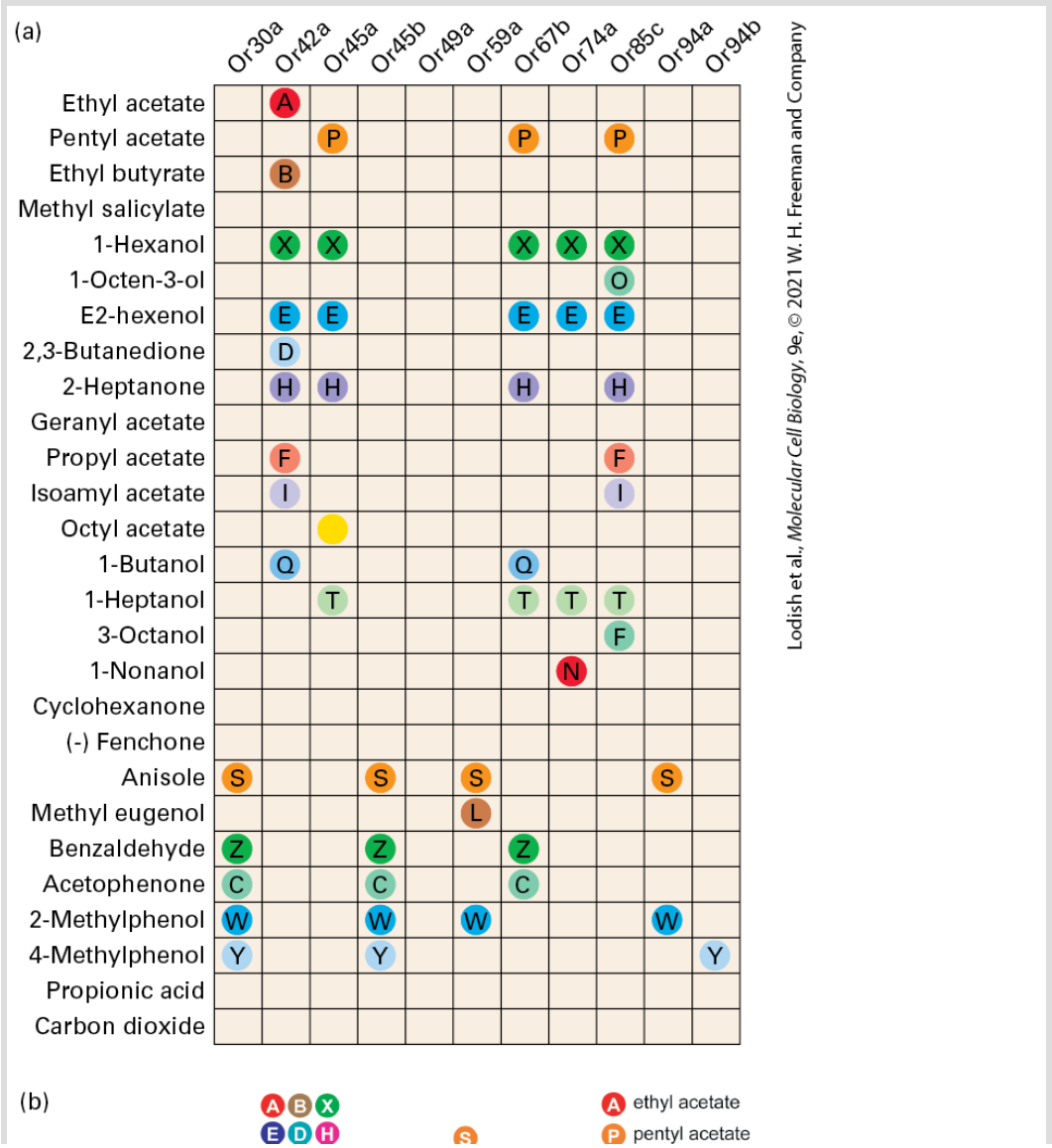
The key to understanding the specificity of the olfactory system is that in both mammals and insects each ORN produces only a single type of odorant receptor. Any electrical signal from that cell will convey to the brain a simple message: “an odorant is binding to my receptors.” Receptors are not always completely monospecific for odorants. Some receptors can bind more than one kind of molecule,

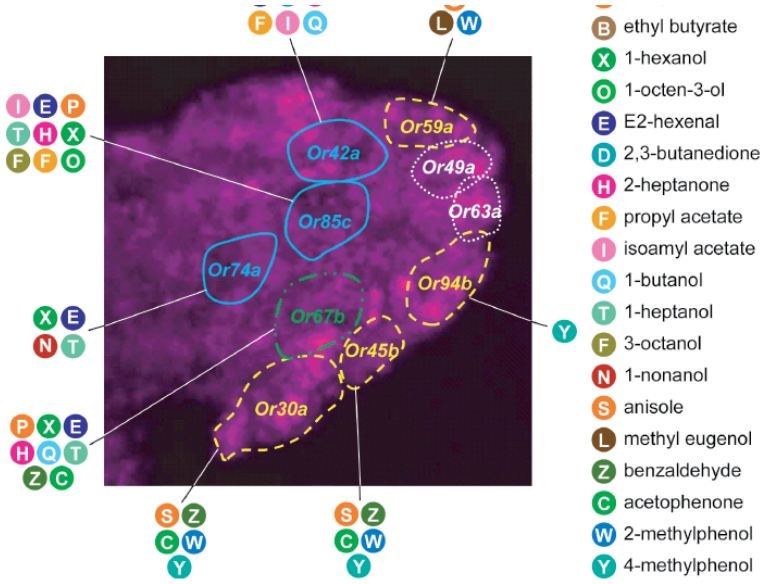
but the molecules detected are usually closely related in structure. Conversely, some odorants bind to multiple receptors.

There are about 5 million ORNs in the mouse, so on average each of the 800 or so olfactory receptor genes is active in approximately 6000 cells. There are about 2000 glomeruli (roughly 2 for each odorant receptor gene), so on average the axons from a few thousand ORNs converge on each glomerulus (see Figure 23-39). From there, about 25 mitral axons per glomerulus, or a total of 50,000 mitral neurons, connect to higher brain centers. Thus the initial odorant-sensing information is carried directly to higher parts of the brain without processing, a simple report of what odorant has been detected.

The one neuron–one receptor rule extends to *Drosophila*. Detailed studies have been done in larvae, where a simple olfactory system with only 21 ORNs uses about 10–20 olfactory receptor genes. It appears that a unique receptor is expressed in each ORN, which sends its projections to one glomerulus. ORNs can send either excitatory or inhibitory signals from their axon termini, probably in order to distinguish attractive versus repulsive odors. Similar to mammals, the axons from the ORNs end in the glomeruli, which in flies are located in the antennal lobe of the larval brain. The research in *Drosophila* began with tests of which odorants bind to which receptors (Figure 23-40a). Some odorants are detected by a single receptor, some by several, so the combinatorial pattern allows many more odorants to be distinguished than just the

number of different olfactory receptors. The small total number of neurons has allowed a map to be constructed showing which odorants are detected by every glomerulus (Figure 23-40b). One striking finding was that glomeruli located near each other respond to odorants with related chemical structures: for example, linear aliphatic compounds or aromatic compounds. The arrangement may reflect evolution of new receptors concomitant with a process of subdivision of the olfactory part of the brain.





EXPERIMENTAL FIGURE 23-40 Individual olfactory receptor types can be experimentally linked to various odorants and traced to specific glomeruli in the *Drosophila* larval olfactory system. (a) The different olfactory receptor proteins are listed across the top, and the 27 odorants tested are shown down the left side. Colored dots indicate strong odor responses. Note that some odorants stimulate multiple receptors (e.g., pentyl acetate), while others (e.g., ethyl butyrate) act on only a single receptor. Note that many receptors, such as Or42a or Or67b, respond primarily to aliphatic compounds, whereas others, such as Or30a and Or59a, respond to aromatic compounds. (b) Spatial map of olfactory information in glomeruli of the *Drosophila* larval brain. The mapping was done by expressing a reporter gene under the control of each of the selected ORNs. The micrograph indicates the glomeruli that receive projections from ORNs producing each of the 10 indicated receptor protein types (Or42a, etc.). Also indicated are the odorants to which each receptor responds strongly. Note that, in this study, with one exception (Or30a and Or45b) each glomerulus has unique sensory capacities. Glomeruli-sensing odorants that are chemically similar tend to be situated next to one another. For example, the three glomeruli indicated by a blue solid line sense linear aliphatic compounds; those with yellow dashed lines, aromatic compounds.

[Part (b) republished with permission from Elsevier, from S. A. Kreher, J. Y. Kwon, and J. R. Carlson, 2005, “The Molecular Basis of Odor Coding in the

Drosophila Larva," *Neuron* **46**(3):445–456; permission conveyed through Copyright Clearance Center, Inc.]

The simple system of having each cell make only one receptor type also has some impressive difficulties to overcome: (1) Each receptor must be able to distinguish a type of odorant molecule or a set of molecules with specificity adequate to the needs of the organism. A receptor stimulated too frequently would not be useful. (2) Each cell must express one and only one receptor gene product. All the other receptor genes must be turned off. At the same time, the collective efforts of all the cells in the nasal epithelium must allow the production of enough different receptors to give the animal adequate sensory versatility. It does little good to have genes for hundreds of receptors if most of them are never expressed, but it is a regulatory challenge to turn on one and only one gene in each cell and at the same time express all the receptor genes across the complete population of cells. (3) The neuronal wiring of the olfactory system must make discrimination among odorants possible so that the brain can determine which odorants are present. Otherwise the animal might be feeling at ease and relaxed when it should be running away as fast as possible.

The solution to the first problem is the great variability of the olfactory receptor proteins, both within and between species. The solution to the second problem, the expression of a single olfactory receptor gene per cell, has been shown to involve a remarkable form of epigenetic silencing that assures that thousands of olfactory

receptor alleles remain inactive in each ORN. These studies have shown that receptor choice relies on the selective activation of a single olfactory receptor gene from a developmental state in which all olfactory receptor genes are silenced. Activation is triggered by a histone demethylase and a specific adenylate cyclase, both of which are required to derepress the single olfactory receptor locus. The active and inactive genes are spatially segregated within the nucleus, with the inactive genes being buried in heterochromatic foci while the active genes are located in euchromatic domains (see Chapter 8).

The third problem, how the system is wired so the brain can understand which odor has been detected, has been partly answered. First, ORNs that express the same receptor send their axons to the same glomerulus. Thus all cells that respond to the same odorant send processes to the same destination. In mice, a crucial clue about the patterning of the olfactory system came from the discovery that olfactory receptors play two roles in ORNs: odorant binding and, during development, axon guidance. Multiple ORN axons expressing the same receptor are guided to the same glomerulus destination. Each olfactory receptor has a distinct, odorant-independent level of activation that turns on adenylate cyclase, with the varying levels of cAMP turning on CREB-dependent expression of standard axon-guidance molecules, whose graded activity is used to specify targeting to a specific glomerulus.

KEY CONCEPTS OF SECTION 23.4

Sensing the Environment: Touch, Pain, Taste, and Smell

- Mechanoreceptors and pain receptors are gated Na^+ or Na^+/Ca^{2+} channels. They are present on specialized terminals at the end of sensory neurons and convey afferent information to the spinal cord and brain stem.
- Touch sensitivity requires several cytoskeletal and extracellular matrix proteins as well as a gated Na^+ channel (see Figure 23-33).
- Piezo 1 and 2 are large-channel proteins that convert mechanical stimulus directly into cation conductance (see Figure 23-34).
- TRPV channels are nociceptors that are activated by a variety of stimuli, including heat and capsaicin. Their molecular structure is similar to the structure of voltage-gated ion channels and was recently solved by single particle cryoEM (see Figure 23-35).
- Five primary tastes are sensed by subsets of cells in each taste bud. Salty and sour tastes are detected by specific ion channel proteins, and G protein–coupled receptors detect sweetness, umami, and bitterness.
- In all cases, tastants lead to membrane depolarization and secretion of small molecules such as ATP that stimulate the adjacent neurons.
- Taste is represented in a topographic map in a part of the cortex called the *insula*. Activation of taste receptors by a specific type of taste (e.g., sweet or salty) activates neurons in specific, nonoverlapping regions of the insula.
- Odorant receptors, which are seven-transmembrane G protein–coupled receptors, are encoded by a very large set of genes. Any one olfactory receptor neuron expresses one and only one olfactory receptor gene, so a signal from that cell to the brain unambiguously conveys the nature of the chemical sensed.
- Olfactory receptor neurons (ORNs) that express the same receptor gene send their axons to the same glomerulus, and projection nerves (mitral neurons in mammals) carry odorant-specific information from the glomeruli to the brain (see Figures 23-38, 23-39, and 23-40).

23.5 Forming and Storing Memories

One of the most remarkable features of the brain is its capacity to form and store memories. Decades of research have revealed that memories are stored as changes in the strength and number of connections that form between neurons. While the overall structure of the nervous system is genetically hardwired, neural circuits undergo extensive sculpting and rewiring in response to a variety of stimuli. This process of experience-dependent changes in synaptic connectivity is called **synaptic plasticity**. By modifying brain wiring in response to experiences, synaptic plasticity provides a biological means of integrating nature and nurture to determine our identities.

Memories Are Formed by Changing the Number or Strength of Synapses Between Neurons

The concept of synaptic plasticity has a long history, beginning with the neuroanatomical studies of Santiago Ramón y Cajal at the turn of the nineteenth century. He used a method called the *Golgi stain* to visualize individual neurons in the brains of humans and other animals (Figure 23-41a). The Golgi stain was developed by the Italian scientist Camillo Golgi, with whom Ramón y Cajal shared the 1906 Nobel Prize in Physiology or Medicine for their work on the structure of the nervous system. While Golgi believed that the brain

consisted of a “reticular network,” a large syncytium of interconnected nerve cells, Ramón y Cajal recognized that the brain consisted of individual neurons that interacted with one another at sites of contact — what we now know of as synapses. Ramón y Cajal detected synapses as small dendritic protuberances. These protuberances are the postsynaptic compartments of excitatory synapses and can be visualized not only with the Golgi stain but also with more modern methods based on genetic expression of fluorescent proteins (Figure 23-41b). Based on his histological data, Ramón y Cajal hypothesized that memories were stored in the brain by changing the structure of the neuronal arbor and by changing the structure and number of synapses that formed between neurons. In poetic terms, Ramón y Cajal speculated that: “the cerebral cortex is like a garden full of innumerable trees, the pyramidal cells, which in response to intelligent cultivation can increase the number of their branches … and produce ever more varied flowers and fruit.”

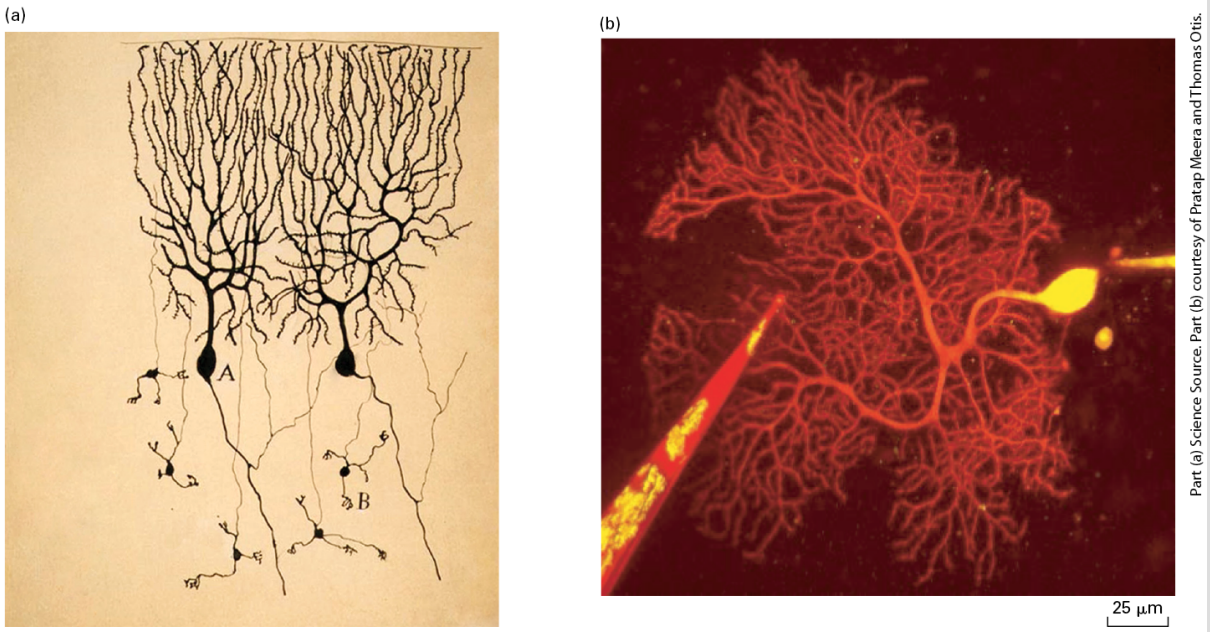


FIGURE 23-41 Visualizing dendritic spines. (a) Santiago Ramón y Cajal used the Golgi staining method to visualize individual neurons in the cerebellum of a pigeon in 1899. This method permitted Ramón y Cajal to visualize individual neurons in the brain; the tissue is densely packed with neurons but the Golgi stain only labels sparse neurons in the tissue. The postsynaptic compartment of excitatory synapses consists of a spiny protuberance from the dendrite, called a *spine*. Ramón y Cajal detected these spines in neurons (here in the Purkinje neurons of the cerebellum). In modern-day approaches, fluorescent proteins can be delivered using a microelectrode or expressed genetically to allow visualization of a single neuron in tissue. (b) A fluorescent dye is delivered to the cell body of a single Purkinje neuron in mouse cerebellum by a microelectrode (*right*) and is visualized by two-photon microscopy. The fluorescence is yellow in the cell body and orange-red in the dendrites. At higher resolution, we can image spine dynamics using time-lapse microscopy and in this way directly demonstrate changes in synaptic connectivity with experience. In this image, a second, larger electrode (*left*) filled with a red fluorescent dye (with yellow precipitates within the electrode) is used to electrically stimulate synapses forming onto the labeled neurons.

Decades of research have largely validated Ramón y Cajal's predictions, although memories are now thought to be stored

primarily as changes in the synapses ("flowers and fruit") rather than by changes in dendrites and axons ("branches"). Studies of the gill-withdrawal reflex in the sea slug *Aplysia californica* provide a classic demonstration of the structural basis of memory storage (Figure 23-42). *Aplysia californica* is a useful model organism for studying the cell biology of memory because its nervous system is relatively simple and its neurons are very large and identifiable, which means that the same neuron can be identified from one animal to another. These features allowed Nobel laureate Eric Kandel and his colleagues to delineate the neural circuitry underlying specific behaviors in the animal and to then determine how the synaptic connections between neurons in this circuit changed during memory formation. They focused on a simple reflexive behavior, the siphon gill-withdrawal reflex, in which touching the animal's siphon (a tubelike anatomical structure through which water flows) leads to a defensive withdrawal of its respiratory organ, the gill. Sensory neurons from the siphon that synapse onto motor neurons to the gill mediate the reflex. Touching the siphon triggers firing of the sensory neuron, which triggers an action potential in the motor neuron, which in turn synapses on the gill muscle and causes it to contract. The reflex can be either strengthened or weakened by experience. Repeated touching of the siphon leads to a decrease in the amplitude of the gill-withdrawal reflex, called *habituation*. In contrast, presentation of a noxious stimulus like delivery of an electric shock to the tail leads to an increase in the amplitude of the gill-withdrawal reflex, called

sensitization. Sensitization can be thought of as a form of fear learning. Habituation and sensitization can be transient or long lasting, depending on the strength and duration of the stimulus. Long-lasting forms of habituation and sensitization were found to involve dramatic decreases and increases, respectively, in the number of connections that formed between sensory and motor neurons. In this way, just as Ramón y Cajal predicted, the animal's experience changed the wiring of its nervous system, thereby encoding a memory and changing the animal's behavior.

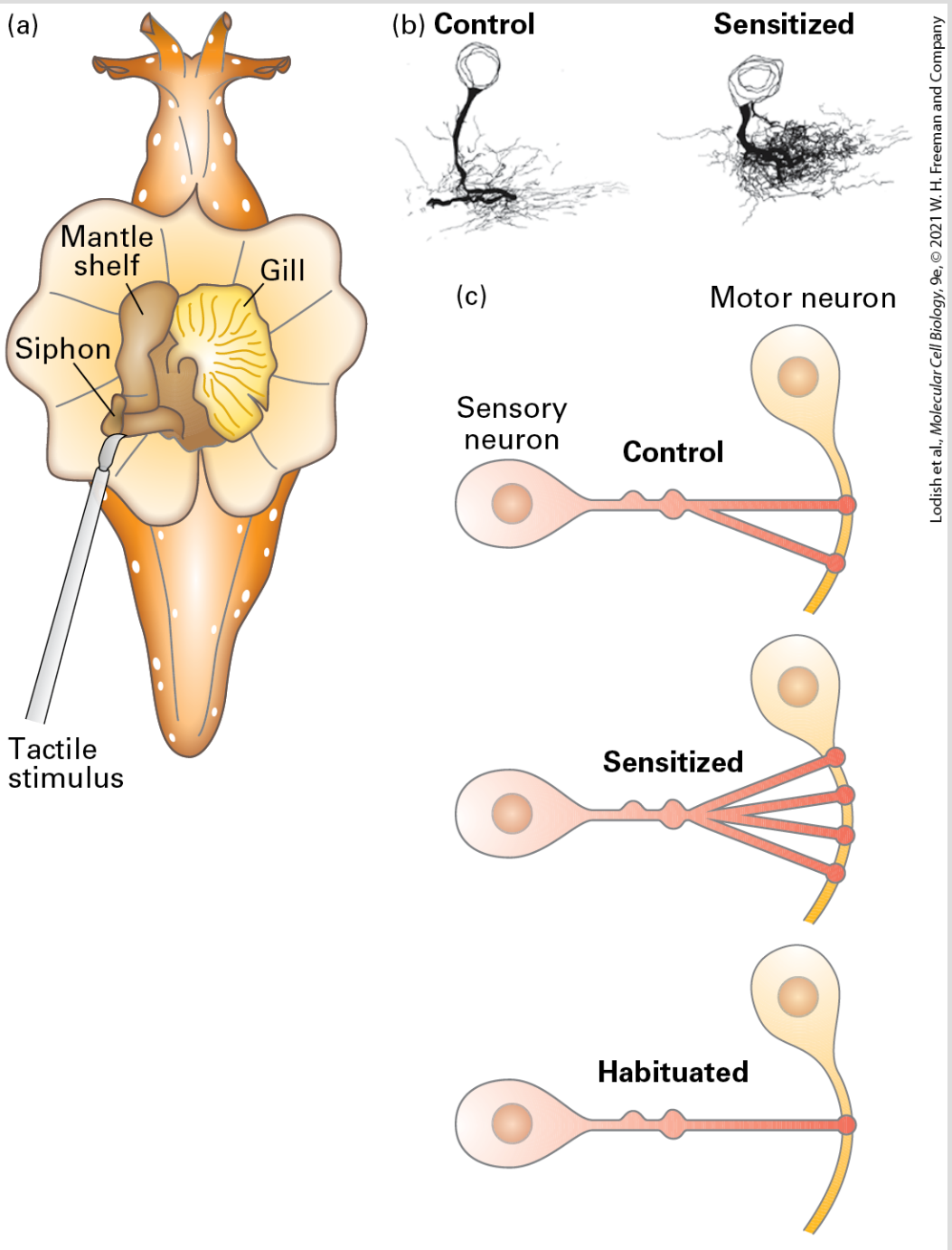


FIGURE 23-42 Long-term memories are stored as changes in synaptic connectivity. (a) The sea slug *Aplysia californica* is a model system for studying the cell biology of synaptic plasticity and memory. Tactile stimulation of the siphon (a tubelike structure through which water flows) stimulates the gill-withdrawal reflex. In habituation, the siphon is repeatedly touched, which habituates the animal to this stimulation and reduces the amplitude of the gill

withdrawal. In sensitization, the animal receives a noxious stimulus like a tail shock, which sensitizes the reflex so that the gill-withdrawal amplitude is enhanced. (b) Three-dimensional reconstructions of siphon sensory neurons from control animals and from animals that have undergone long-term sensitization of the gill-withdrawal reflex. Notice the expansion of the sensory neuron branches after sensitization. The growth of neuronal processes is accompanied by a growth of new synaptic connections between the sensory and motor neurons. (c) Illustrations showing the changes in connectivity that occur during plasticity of the gill-withdrawal reflex. Sensitization is accompanied by the growth of new connections between the sensory and motor neuron, while habituation is accompanied by a decrease in the number of connections between the sensory and motor neuron.

[Part (b) from C. H. Bailey and M. Chen, 1988, "Long-Term Memory in Aplysia Modulates the Total Number of Varicosities of Single Identified Sensory Neurons," *Proc. Nat'l Acad. Sci. USA* **85**:2373–2377.]

The Hippocampus Is Required for Memory Formation

Studies in *Aplysia* and in other model organisms, including *Drosophila* and mice, have begun to reveal many of the molecular mechanisms underlying experience-dependent synaptic plasticity. Clinical studies in humans as well as experimental studies in animals have shown that the hippocampus is required for the formation of long-term memories. Humans and animals with lesions in their hippocampus can form short-term memories and maintain their old memories but are no longer able to form new long-term memories. Not only is the hippocampus critical for long-

term memory formation, but its anatomy also makes it especially suitable for electrophysiological studies of synaptic connectivity. As shown in Figure 23-43, the hippocampus consists of three sequential pathways (perforant, mossy fiber, and Schaffer collateral pathways), each with discrete cell-body layers and axonal and dendritic projections. High-frequency stimulation of the axons of presynaptic neurons in each of these pathways produces a long-lasting strengthening of the connections onto the postsynaptic neurons, called *long-term potentiation (LTP)*, while low-frequency stimulation produces a long-lasting weakening of the connections, called *long-term depression (LTD)*.

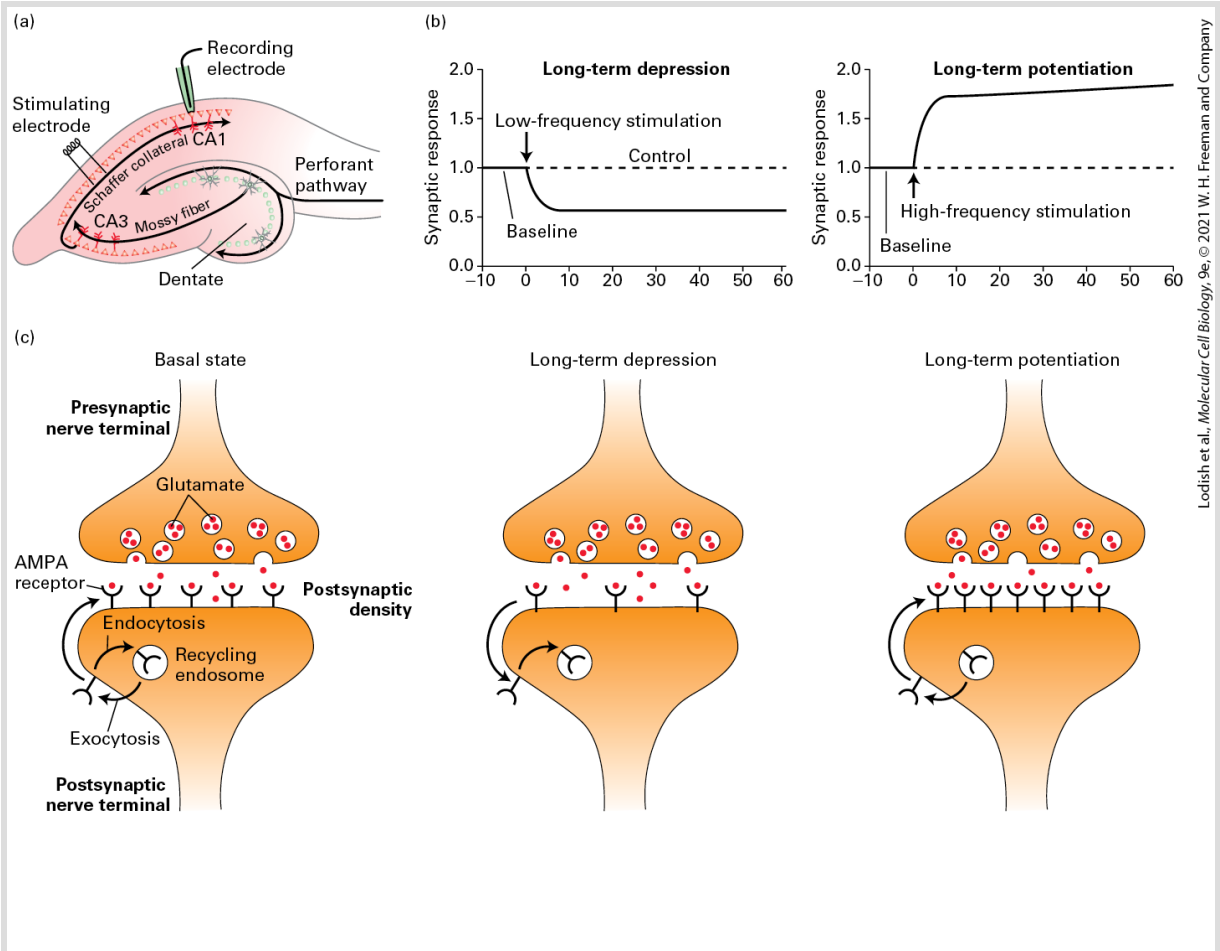


FIGURE 23-43 Synaptic plasticity in the mouse hippocampus: long-term potentiation (LTP) and long-term depression (LTD), and roles for regulated trafficking of AMPA glutamate receptors. (a) The mouse hippocampus can be dissected from mouse brain cut into transverse slices, preserving the three sequential synaptic pathways, from the entorhinal cortex (not shown) to the dentate gyrus (DG), the DG to CA3, and CA3 to CA1. Electrodes can be used to stimulate axonal afferents and record from postsynaptic follower cells, as illustrated for the Schaffer collateral (CA3-CA1) pathway. (b) Trains of low-frequency stimulation or high-frequency stimulation to the axonal fibers produce sustained decreases or increases in synaptic strength, which are measured as the postsynaptic response to a test stimulus. These tracings show the strength of the synaptic connection on the *y* axis, with time on the *x* axis. The dotted line shows the baseline strength of the connections, which are decreased by low-frequency stimulation and increased by high-frequency stimulation. These forms of plasticity are known as *long-term depression (LTD)* and *long-term potentiation (LTP)*. (c) In the basal state, AMPA receptors (black) traffic constitutively to and from the plasma membrane of the postsynaptic compartment via recycling endosomes. Receptors are delivered to the plasma membrane lateral to the postsynaptic density via exocytosis and are internalized by clathrin-mediated endocytosis into recycling endosomes. In the postsynaptic density, the AMPA receptors are stabilized by interactions with proteins, including transmembrane AMPA receptor regulatory proteins (TARPS, not shown). Following induction of LTD of glutamatergic synapses, there is an increase in the diffusion of AMPA receptors out of the postsynaptic density and in their internalization into recycling endosomes. This results in a decrease in the number of AMPA receptors on the postsynaptic membrane and a decrease in the postsynaptic response to a given amount of glutamate release from the presynaptic neuron. Following induction of LTP at glutamatergic synapses, there is an increase in the exocytosis of AMPA receptors and an increase in their diffusion into the postsynaptic density. This results in an increase in the number of AMPA receptors on the postsynaptic membrane and an increase in the postsynaptic response to a given amount of glutamate release from the presynaptic neuron. Regulated trafficking of AMPA receptors provides one molecular mechanism underlying the activity-dependent changes in synaptic strength that accompany synaptic plasticity and memory. See V. M. Ho, J.-A. Lee, and K. C. Martin, 2011, *Science* **334**:623–628; and J. D. Shepherd and R. L. Huganir, 2007, *Annu. Rev. Cell Dev. Bi.* **23**:613–643.

While a multitude of studies have shown correlations between LTP, LTD, and memory, more recent optogenetic studies succeeded in demonstrating a causal role for synaptic plasticity in producing memories. To do this, the investigators expressed channelrhodopsin in hippocampal neurons in mice and stimulated the neurons with light to induce LTP. Induction of LTP caused the mice to acquire a false memory in which they demonstrated fear to an environment even though they had never encountered a frightening stimulus in that environment!

Multiple Molecular Mechanisms Contribute to Synaptic Plasticity

In considering how experience can change synaptic strength, it is useful to think about the structure of the chemical synapse and the process of synaptic transmission described in Section 23.3. Long-lasting changes in plasticity have been shown to involve presynaptic changes in neurotransmitter release, trans-synaptic adhesion, and postsynaptic responses to neurotransmitter. We will briefly touch on pre- and trans-synaptic mechanisms and then delve into slightly more detail on postsynaptic mechanisms, which have been studied in greater depth.

Experiences that stimulate hippocampal neurons trigger elevations in intracellular calcium, which in turn activates kinases that

phosphorylate synapsins, the molecules that organize synaptic vesicles into discrete pools within the presynaptic compartment. This phosphorylation of synapsins increases the number of synaptic vesicles available for release, thereby increasing the amount of neurotransmitter that is released with a given stimulus. Experience also activates kinases that phosphorylate RIM, the molecule that tethers voltage-gated Ca^{2+} channels to the release machinery, and this phosphorylation is required for LTP of hippocampal synapses.

Synaptic plasticity also depends on activation of kinases in the postsynaptic compartment. Influx of Ca^{2+} through voltage-gated Ca^{2+} channels and specific glutamate receptors in the postsynaptic membrane activates one particularly important kinase, the calcium-calmodulin-dependent kinase IIα (CaMKIIα; see Chapter 15 for discussion of calcium calmodulin signaling). This kinase has the special property that once activated, it can remain persistently activated even in the absence of stimulation. This is because once activated, CaMKIIα autophosphorylates itself, which renders it constitutively active for about 30 minutes, during which time the kinase phosphorylates many substrates in the postsynaptic compartment, including glutamate receptors. Mice lacking CaMKIIα have deficits in both LTP in the hippocampus and in memory formation.

As discussed in Chapter 15, the sensitivity of a cell to external signals is determined by the number of surface receptors. In line

with this concept, one of the best characterized mechanisms underlying synaptic plasticity involves activity-dependent changes in the number of glutamate receptors that are present in the postsynaptic membrane. This process has been especially well studied in the context of hippocampal LTP and LTD (Figure 23-43c). One of the major classes of glutamate receptors, called *AMPA receptors*, traffic constitutively to and from the plasma membrane via recycling endosomes. AMPA receptors are delivered by exocytosis at extrasynaptic sites and then laterally diffuse into the postsynaptic density, the protein-dense region of the postsynaptic site that faces the synaptic cleft and thus receives neurotransmitters released by the presynaptic terminal. AMPA receptors are removed by endocytosis, which occurs when the receptors diffuse laterally to extrasynaptic sites and are then internalized by clathrin-mediated, dynamin-dependent endocytosis.

While AMPA receptor trafficking occurs under basal conditions, it is modulated by activity through changes in actin and myosin dynamics as well as AMPA receptor interactions with scaffolding proteins and accessory subunits. One of these accessory subunits, *Stargazin*, mediates the interaction between AMPA receptors and the postsynaptic density protein PSD95. This interaction is critical to the localization of AMPA receptors at the synapse, since interaction with PSD95 stabilizes AMPA receptor localization within the postsynaptic density. Activity causes the phosphorylation of Stargazin, decreasing the mobility of AMPA receptors and increasing their concentrations

at the synapse. Blocking Stargazin phosphorylation or de-phosphorylation blocks hippocampal LTP and LTD, respectively. Stargazin is one of a family of *transmembrane AMPA receptor regulatory proteins (TARPs).* TARPs bind to all AMPA receptor subunits, are differentially expressed throughout the brain, and mediate the delivery of AMPA receptors to the surface and synapse of neurons.

Formation of Long-Term Memories Requires Gene Expression

The mechanisms described above are especially important for short-term forms of plasticity that underlie short-term memories. While the formation of short-term plasticity, and short-term memory, have been shown to rely on modifications of preexisting proteins at the synapse, the formation of long-term memories differs in that it depends upon new gene expression. This can be thought of in the context of the different effects of extracellular stimulation that were discussed in Chapter 15: stimuli can produce short-term changes by altering the activity of preexisting enzymes and proteins in the cell, or long-term functional changes by altering the expression of genes in the cell (see Figure 15-3). Studies in many systems and species, including in rodent hippocampus, have demonstrated that LTP and LTD can be divided into transient forms of plasticity that do not require gene expression and *l*ong-lasting

forms (L-LTP and L-LTD) that require both mRNA and protein synthesis.

The extreme morphological polarity and compartmentalization of neurons adds significant challenges to stimulus-induced changes in gene expression. First, to turn on transcription, signals must be relayed to the nucleus from the synapse, which in many cases is located at great distances from the cell body. Neurons are specialized for rapid signaling between compartments by electrical signaling. However, many studies have also shown that signaling molecules, including kinases, phosphatases, and transcriptional regulators, are actively transported from stimulated synapses to the nucleus to regulate transcription. In most cases, this long-distance retrograde transport has been shown to involve dynein motor protein–mediated transport along microtubules (as described in Chapter 18). How signaling is faithfully maintained during this long-distance transport in order to couple synaptic stimulation with gene expression is an area of active research.

The second great challenge in understanding the mechanism of stimulus-induced gene expression during synaptic plasticity derives from the fact that each neuron has a single nucleus and yet can form thousands of synapses. Long-lasting forms of synaptic plasticity are often synapse-specific, that is, they involve changes in synaptic strength at some but not all synapses formed by a single neuron. Since long-term plasticity requires transcription, synapse specificity begs the question of how gene expression can be spatially regulated

in such a highly compartmentalized cell. One important mechanism involves the localization of mRNAs and their local translation in response to synaptic stimulation, as was discussed in Chapter 9. Indeed, L-LTP of hippocampal synapses has been shown to require the translation of mRNAs that are localized in dendrites and at synapses. Electron micrographic studies have identified polyribosomes, actively translating ribosomes, at the base of spines in hippocampal neurons and have further shown that the number of spines controlling polyribosomes greatly increases after induction of L-LTP. Together, these studies have focused attention on the importance of post-transcriptional gene regulation in neurons during synaptic plasticity and on a host of questions about mRNA localization and regulated translation: What mRNAs are localized to synapses? How are they localized? How is their translation regulated by synaptic activity? What is the specific function of the locally translated protein? Why are some mRNAs translated into protein in the neuronal cell body and then transported to synapses, and others translated directly at synapses?

Indicative of the importance of post-transcriptional gene regulation in the proper functioning of the nervous system, mutations in an RNA-binding protein, the fragile X mental retardation protein, FMRP, cause a common form of mental retardation, fragile X syndrome (FXS). Mutations in the gene encoding FMRP, the *fmr1* gene, constitute the most common single-gene cause of autism. The most common mutations that lead to FXS

are expansions of CGG repeats in the *fmr1* gene that leads to gene methylation and silencing, as was described for Huntington's disease in Chapter 6. FMRP is a translational repressor that binds target mRNA and prevents its translation. A population of FMRP localizes to the base of dendritic spines, where it is thought to maintain mRNAs in a dormant state until synaptic stimulation triggers their translation. Genetically modified mice that lack FMRP serve as remarkably good models for the human disease. The mice show deficits in learning, reflective of the intellectual disability in FXS patients. Both mice and humans have abnormalities in the structure of their synaptic spines, which are elongated, like immature spines, rather than stubby, like mature spines. Studies in the mouse have revealed excessive basal translation of mRNAs at synapses and have further shown alterations in protein-synthesis-dependent forms of hippocampal LTD. Together, these findings indicate that synaptic translation of localized mRNAs is critical to the formation and the experience-dependent plasticity of neural circuits and that alterations in this process are a cause of neurodevelopmental and cognitive disorders. ■

KEY CONCEPTS OF SECTION 23.5

Forming and Storing Memories

- Experience changes the number and strength of connections between neurons in the brain through a process known as *synaptic plasticity*. Synaptic plasticity provides a biological basis for the formation and storage of memories.
- Studies of habituation and sensitization demonstrate that learning produces changes in synaptic connectivity. Experiments looking at the gill-withdrawal reflex in marine

mollusk *Aplysia californica* show that habituation involves decreases in the connectivity of the sensory and motor neurons that give rise to the gill-withdrawal reflex, while sensitization involves increases in sensory-motor connectivity (see Figure 23-42).

- The hippocampus is a region of the brain that is required for the formation of long-lasting memories (see Figure 23-43). Hippocampal synapses undergo activity-dependent forms of synaptic strengthening called *long-term potentiation (LTP)* and activity-dependent forms of synaptic weakening called *long-term depression (LTD)*.
- Changes in synaptic strength can be mediated by presynaptic, trans-synaptic, or postsynaptic mechanisms.
- Activity generates a constitutively active form of CaMKIIα in the postsynaptic compartment, which phosphorylates substrates in the postsynaptic density, including glutamate receptors, leading to changes in synaptic strength.
- Activity regulates the trafficking of AMPA glutamate receptors in the postsynaptic membrane. LTP is accompanied by an increase in insertion of AMPA receptors in the postsynaptic density, while LTD is accompanied by a decrease in the concentration of AMPA receptors in the postsynaptic density (see Figure 23-43c).
- Short-term forms of synaptic plasticity involve modifications of preexisting proteins at the synapses, but long-term forms require new mRNA and protein synthesis.
- Synapse-specific forms of plasticity involve the local translation of synaptically localized mRNAs.
- Fragile X syndrome is caused by null mutations in the gene encoding the RNA-binding protein FMRP. FMRP regulates local translation at synapses. Mice lacking FMRP have abnormal synapses and exhibit learning impairments and deficits in hippocampal LTD.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content of this chapter:

- Perspectives for the Future
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

acetylcholine
action potential
afferent neurons
Agrin
astrocytes
axon
dendrites
depolarization
efferent neurons
endocytosis
excitatory receptor
excitatory synapse

glial cells
glomeruli
hippocampus
hyperpolarization
inhibitory receptor
inhibitory synapse
interneurons
mechanosensors
membrane potential
motor neurons
MuSK
myelin sheath
nervous system
neuromuscular junction (NMJ)
neurons
neurotransmitter
nociceptors
node of Ranvier (node)
odorants
olfactory receptors
oligodendrocytes
optogenetics
refractory period
repolarization
saltatory conduction
Schwann cells
sensory neurons
SM proteins

SNAREs
synapse elimination
synapses
synaptic plasticity
synaptic vesicles
tastants
taste receptors
voltage-gated channels

Review the Concepts

1. What is the role of glial cells in the brain and other parts of the nervous system?
2. The resting potential of a neuron is approximately −70 mV inside compared with outside the cell. How is the resting potential maintained in animal cells?
3. Name the three phases of an action potential. Describe for each the underlying molecular basis and the ion involved. Why is the term "voltage-gated channel" applied to Na^+ channels involved in the generation of an action potential?
4. Explain how the crystal structures of potassium ion channels suggest the way in which the voltage-sensing domains interact with other parts of the proteins to open and close the ion channels. How does this structure-function relationship apply to other voltage-gated ion channels?

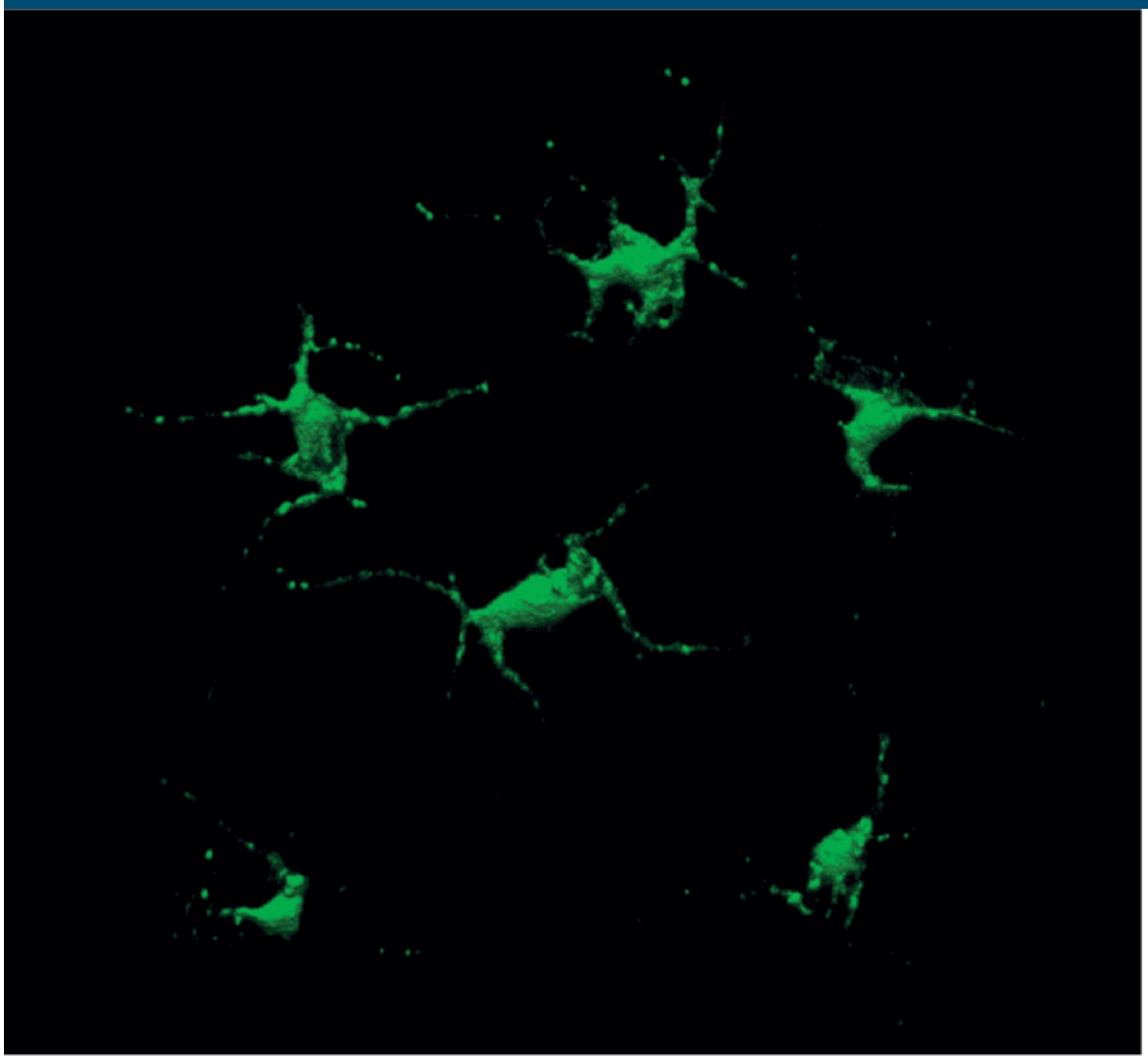
5. Explain why the strength of an action potential doesn't decrease as it travels down an axon.
6. Explain why the membrane potential does not continue to increase but rather plateaus and then decreases during the course of an action potential.
7. What does it mean to say that action potentials are "all or none"?
8. What prevents a nerve signal from traveling "backward" toward the cell body?
9. Why is the cell unable to initiate another action potential if stimulated during the refractory period?
10. Myelination increases the velocity of action potential propagation along an axon. What is myelination? Myelination causes clustering of voltage-gated Na^+ channels and Na^+/K^+ pumps at nodes of Ranvier along the axon. Predict the consequences to action potential propagation of increasing the spacing between nodes of Ranvier by a factor of 10.
11. Describe the mechanism of action for addictive drugs such as cocaine.
12. Acetylcholine is a common neurotransmitter released at the synapse. Predict the consequences for muscle activation of decreased acetylcholine esterase activity at nerve-muscle synapses.
13. Describe the ion dynamics of the muscle-contraction process.
14. Following the arrival of an action potential in stimulated cells, synaptic vesicles rapidly fuse with the presynaptic

membrane. This happens in less than 1 ms. What mechanisms allow this process to take place at such great speed?

15. Neurons, particularly those in the brain, receive multiple excitatory and inhibitory signals. What is the name of the extension of the neuron at which such signals are received? How does the neuron integrate these signals to determine whether or not to generate an action potential?
16. Explain the mechanism by which action potentials are prevented from being propagated to a postsynaptic cell if transmitted across an inhibitory synapse.
17. What is the role of dynamin in recycling synaptic vesicles? What evidence supports this?
18. Compare and contrast electrical and chemical synapses.
19. Compare the structures and functions of the receptor molecules for salty and sour taste; the taste-receptor molecules for sweetness, bitterness, and umami; and odor-receptor molecules.
20. Describe a synaptic mechanism underlying the formation of memory.

CHAPTER 24

Immunology



Courtesy of M. Boes and H. L. Ploegh.

Dendritic cells in the skin have class II MHC molecules on their surface. Those shown here were engineered to express a class II MHC–GFP fusion protein, which fluoresces green.

OUTLINE

24.1 Overview of Host Defenses
24.2 Immunoglobulins: Structure and Function
24.3 Generation of Antibody Diversity and B-Cell Development
24.4 The MHC and Antigen Presentation
24.5 T Cells, T-Cell Receptors, and T-Cell Development
24.6 Collaboration of Immune-System Cells in the Adaptive Response

---

Metazoan animals and the cells within them are surrounded by a sea of pathogens — viruses, bacteria, and fungi — and have evolved a number of defenses to counter them, collectively termed the *immune system*. **Immunity** is a state of protection against the harmful effects of exposure to pathogens. Even insects such as *Drosophila* have an immune system with two types of defenses: (1) physical barriers and chemical defenses, and (2) cells that comprise an innate immune system that can recognize and destroy certain pathogens (Figure 24-1). Mechanical and chemical defenses operate continuously. Innate immune responses, which involve cells and molecules that are present in the body at all times, are rapidly activated (in minutes to hours). However, because they recognize molecules that are common to many foreign organisms, their ability to distinguish among different pathogens is somewhat limited.

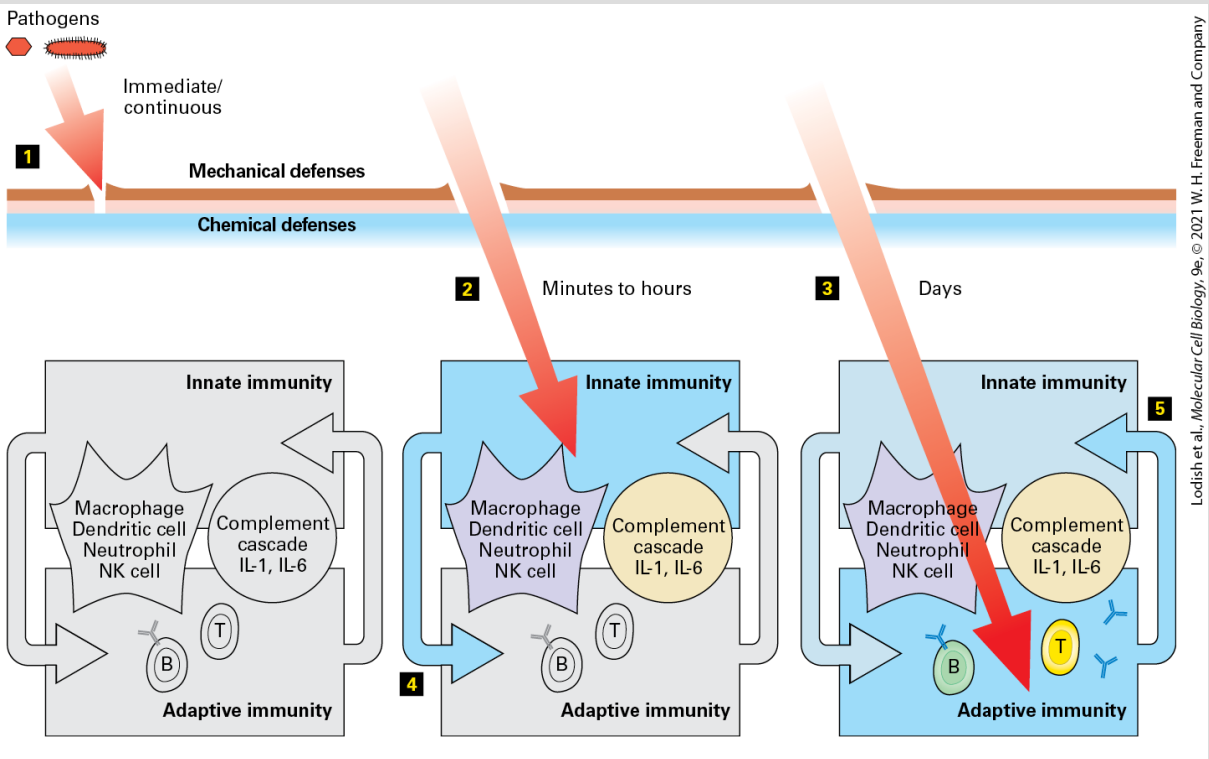


FIGURE 24-1 The three layers of vertebrate immune defenses. *Left:* Mechanical defenses consist of epithelia and skin. Chemical defenses include the low pH of the gastric environment and antibacterial enzymes in tears. These barriers provide continuous protection against invaders. Pathogens must physically breach these defenses (step 1) to infect the host. *Middle:* Pathogens that have breached the mechanical and chemical defenses (step 2) are handled by cells and molecules of the innate immune system (blue). These include phagocytic cells (neutrophils, dendritic cells, macrophages) that engulf the invader and destroy it in lysosomes, natural killer (NK) cells, complement proteins, and certain interleukins (IL-1, IL-6). Innate defenses are activated within minutes to hours of infection. *Right:* In vertebrates, pathogens that are not cleared by the innate immune system are dealt with by the adaptive immune system (step 3), in particular B lymphocytes and their antibody products (Y-shaped icons), and several types of T cells and the many proteins they make that can destroy pathogens. Full activation of adaptive immunity requires days. The products of an innate response may potentiate an ensuing adaptive response (step 4). Likewise, the products of an adaptive immune response, including antibodies, may enhance innate immunity (step 5). Several cell types and secreted products are used in both the innate and adaptive immune systems and serve to connect these two layers of host defense.

In addition to these defense mechanisms, vertebrates have evolved a sophisticated immune system, the *adaptive immune system,* that can recognize and respond to many types of molecules not normally present in the body, including those that are parts of specific pathogens. The adaptive immune system generates secreted proteins termed *antibodies,* each of which binds to a specific foreign target molecule (see Figure 3-22). The adaptive immune system also generates many types of cells that attack and destroy pathogenic invaders. Adaptive immune responses take several days to develop fully and are highly specific; that is, they can distinguish between closely related pathogens based on very small molecular differences in the structures of their proteins or other polymers. The adaptive immune system makes adjustments to these threats over time, changing in response to the types and abundances of pathogens to which the host is exposed.

All pathogens have found ways to disarm the immune system or manipulate it to their own advantage, and host-pathogen interactions are an evolutionary work in progress. Virtually all pathogens have relatively short generation times compared with the hosts they infect and thus can evolve sophisticated countermeasures against their hosts' immune system. This explains why, despite the evolution of remarkably sophisticated adaptive immune systems, pathogenic viruses, bacteria, and parasites continue to pose a threat to human populations. Seasonal outbreaks of influenza caused by new strains of influenza virus are just one example.

These sophisticated defenses come at a price: an immune system capable of dealing with a massively diverse collection of rapidly evolving pathogens can sometimes mistake the host's own tissues for pathogens and mount an attack against its own cells and tissues, a phenomenon called *autoimmunity*. Even so, we have learned to exploit the workings of the immune system to create vaccines that protect against a variety of infectious diseases. Vaccines are remarkably cost effective and have contributed to eliminating the scourge of many epidemics, such as outbreaks of smallpox.

In this chapter, we deal mainly with the vertebrate immune system, with particular emphasis on those molecules, cell types, and pathways that uniquely distinguish the immune system from other types of cells and tissues. Four remarkable features characterize the vertebrate immune system. They are specificity, diversity, memory, and tolerance. Specificity is the immune system's ability to distinguish between closely related substances. **Diversity** is the system's capacity to specifically recognize an astoundingly large number ($>10^6$) of different molecules. **Memory** is a host's ability to recall previously experienced exposure to a foreign substance and more rapidly and effectively defend itself from that substance the next time it is encountered. **Tolerance** is the ability to avoid mounting an immune-system attack against the host's own cells and tissues; that is, the ability to distinguish the body's own tissues ("self") from foreign materials including pathogens ("nonself").

Any material that can evoke an immune response is referred to as an **antigen**. As we shall see, the immune system achieves specificity and diversity by generating a large number of distinct proteins, such as antibodies and specific cell-surface receptors, each of which can bind very tightly to a target antigen such as a pathogenic molecule, but not to other, perhaps structurally very similar, molecules.

Memory and tolerance depend on complex cellular systems. They are accomplished through the generation of a massively diverse set of cell-surface receptors that bind specific antigens. Cells that display these receptors have been "trained" so that they are largely unresponsive to self components (self-tolerant).

From a practical perspective, the powers of the immune system can be exploited therapeutically. Today there is a multibillion-dollar market for monoclonal antibodies, which are used in the treatment of inflammatory conditions, autoimmune diseases, and cancer. The molecules that constitute the adaptive immune system — antibodies in particular — are also indispensable tools for the cell biologist, as we saw in Chapters 3 and 4. Antibodies allow visualization and isolation of the molecules they recognize with pinpoint precision. Their ability to do so has been invaluable in the accurate description of the components that make up the cell and its organelles and their localization, both in cells and in tissues. The technique of *immunofluorescence*, for example, is widely used by cell biologists to study cell morphology and behavior, while *immunoblotting (Western*

blotting) has become an indispensable tool in the study of signal transduction.

The ways in which these foreign materials are recognized and eliminated involve molecular and cell biological principles unique to the immune system. We begin this chapter with a brief sketch of the organization of the mammalian immune system, introducing the essential cell players in innate and adaptive immune responses and describing inflammation, a localized response to injury or infection that leads to the activation of immune-system cells and their recruitment to the affected site. In the next two sections, we discuss the structure and function of **antibody** (or **immunoglobulin**) molecules, which bind to specific molecular features on antigens, and how variability in antibody structure contributes to the recognition of specific antigens. The enormous diversity of antigens that can be recognized by the adaptive immune system finds its explanation in unique rearrangements of segments of DNA to generate antigen-specific receptors in B and T lymphocytes, commonly called **B cells** and **T cells**, which are the white blood cells that carry out antigen recognition. These gene rearrangements permit adaptation to a wide variety of pathogens by setting the specificity of antigen-binding receptors on lymphocytes; they also determine cell fate in the course of lymphocyte development.

Although the gene rearrangement mechanisms that give rise to antigen-specific receptors on B and T cells are very similar, the manner in which these receptors bind to (recognize) antigens is very different. The receptors on B cells — antibody molecules anchored in the plasma membrane — can interact with intact antigens directly, but the receptors on T cells cannot. Instead, as described in Section 24.4, the receptors on T cells recognize forms of antigen that are processed by target cells: small peptides and other small molecules that are displayed (or "presented") on the surfaces of these cells by specialized cell-surface glycoproteins. These glycoproteins are encoded by genes in a region of the genome called the **major histocompatibility complex (MHC)**. These MHC-encoded glycoproteins, also called *MHC products*, help determine the host's ability to mount both T-cell and B-cell responses to antigens.

Understanding these fundamental properties of the immune system has allowed us to answer a number of very practical questions: How can we best make antibodies that afford protection against an infectious agent? How can we raise antibodies to specific proteins we want to study in the laboratory? Knowledge of antigen processing and presentation thus informs both vaccine design to protect against infectious disease and the generation of tools essential for research. MHC-encoded glycoproteins also play a key role in an individual's development of tolerance for his or her own antigens. We conclude the chapter with an integrated view of the immune response to a pathogen, highlighting the collaboration between

different immune-system cells that is required for an effective immune response.

24.1 Overview of Host Defenses

Because the immune system evolves in the presence of microbes, including pathogenic ones, we begin our overview of host defenses by examining where typical pathogens are found in a host and where they replicate. Then we introduce the basic concepts of innate and adaptive immunity, including some of the key cellular and molecular players.

Pathogens Enter the Body Through Different Routes and Replicate at Different Sites

Human skin has a surface area of some 20 square feet. The epithelial surfaces that line our airways, gastrointestinal tract, and urogenital tract present an even more formidable surface area of about 4000 square feet. All these surfaces are exposed on a daily basis to viruses, bacteria, and fungi in the environment. Some of these bacteria, called *commensal bacteria*, do not usually cause disease and in fact can be beneficial, helping to provide key nutrients or to maintain healthy skin. The microbiota also helps tune reactivity and composition of both the innate and adaptive immune systems via engagement of their receptors and the release of small molecules. It is thought that at any point in time, an adult human may be carrying as much as 3 pounds of microbes, against

which most of us do not develop an overt inflammatory immune reaction. These commensal microbes are not pathogenic as long as they remain on these outer surfaces of the body. If the normal barrier function of the epithelia that compose these surfaces is compromised, however, and these microbes enter the body, they can be pathogenic. Food-borne pathogens and sexually transmitted agents target the epithelia to which they are exposed. The sneeze of a flu-infected individual releases millions of virus particles in aerosolized form, ready for inhalation by a new host. Rupture of the skin, even if only by minor abrasions, or of the epithelial barriers that protect the underlying tissues provides an easy route of entry for pathogens, which then gain access to a rich source of nutrients (for bacteria) and to the cells required for replication (for viruses).

Replication of viruses is confined strictly to the cytoplasm or nuclei of host cells, where viral protein synthesis and replication of the viral genetic material occur. Viruses can then spread to other cells either as free virus particles (virions) released from the initially infected cell or by direct transfer to an adjacent cell (cell-to-cell spreading). Many bacteria can replicate in the extracellular spaces of the body, but some are specialized to invade host cells and survive and reproduce within those cells. Such intracellular bacteria, such as *Mycobacterium tuberculosis,* the causative agent of tuberculosis, reside either in the membrane-delimited vesicles through which they enter cells by endocytosis or phagocytosis (see Figure 17-19) or in the cytoplasm if they escape from these vesicles. An effective host

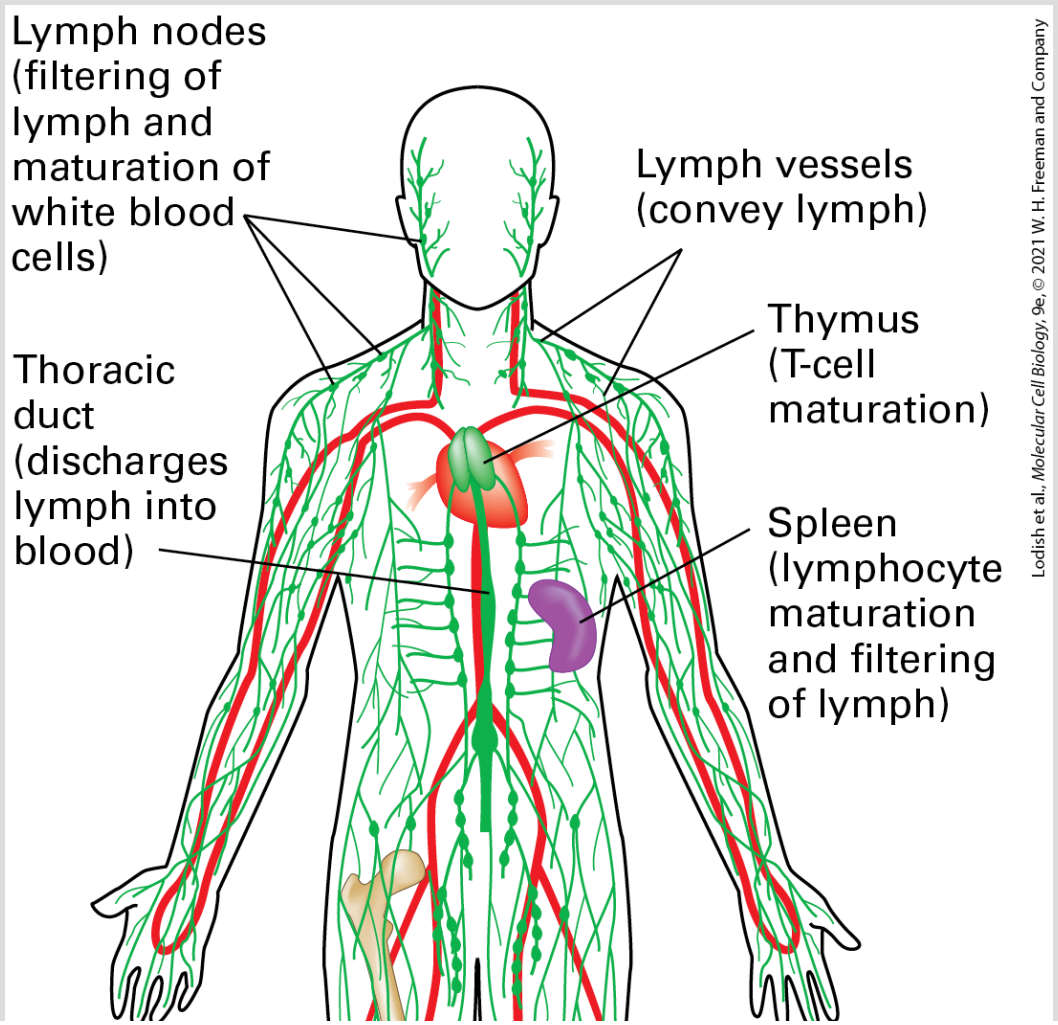
defense system, therefore, needs to be capable of eliminating not only extracellular viruses and bacteria, but also host cells that harbor pathogens.

Parasitic eukaryotes can also cause disease. Some of these parasites, such as the protozoans that cause sleeping sickness (trypanosomes) or malaria (*Plasmodium* species; see Figure 1-25), have very complex life cycles and have evolved complex countermeasures to avoid destruction by the host's immune system.

Cells of the Innate and Adaptive Immune Systems Circulate Throughout the Body and Take Up Residence in Tissues and Lymph Nodes

The circulatory system (Figure 24-2) is responsible for moving blood throughout the body. Blood comprises cells (red and white blood cells, platelets) and liquid (plasma, which contains dissolved substances including proteins, ions, and small molecules). In addition to the hemoglobin-containing, oxygen-carrying erythrocytes (red blood cells) that compose the overwhelming majority of blood cells, the blood also contains *leukocytes* (white blood cells) and platelets (involved in blood clotting). Leukocytes encompass a variety of cell types, including lymphocytes (B and T cells), monocytes (precursors to the scavenger cells called

macrophages), dendritic cells, neutrophils, and natural killer (NK) cells, all of which have distinct functions in the immune system. In contrast to erythrocytes, which never leave the circulation until they get old and die, leukocytes leave the circulation and enter target tissues to help protect the body from invaders. The circulatory system moves leukocytes from the sites where they are generated (bone marrow, thymus, fetal liver) to the sites where they can be activated (lymph nodes, spleen), and then to the site of infection. Once leukocytes arrive at a given location, they may leave and re-enter the circulation in the course of their tasks.



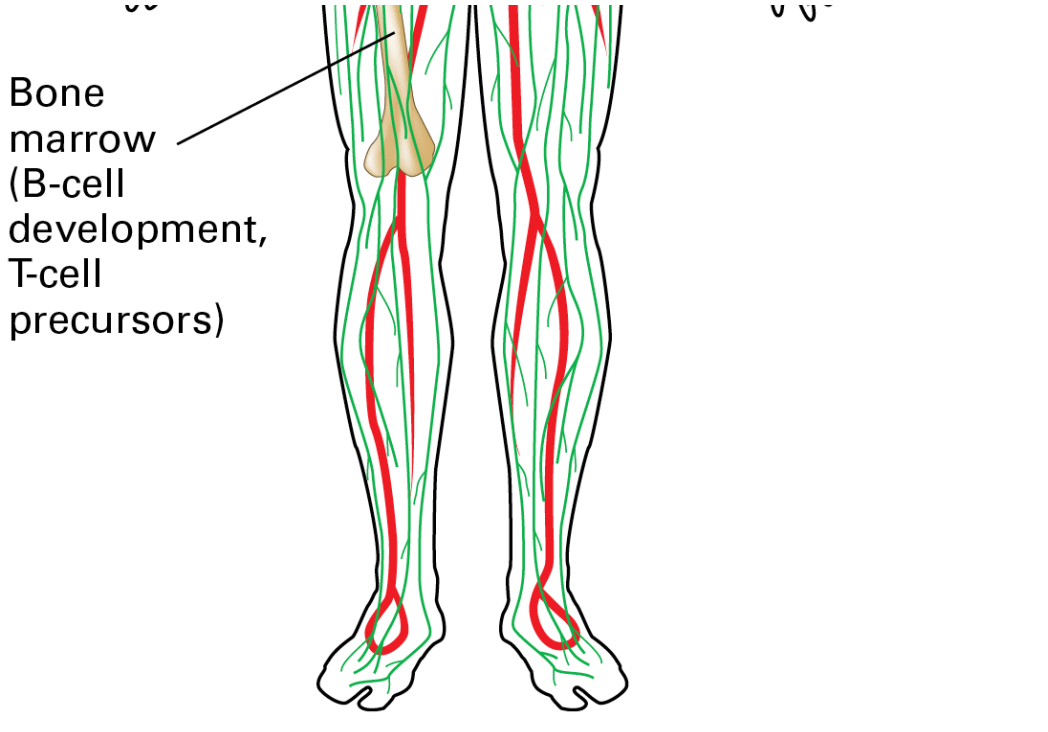


FIGURE 24-2 The circulatory and lymphatic systems. Positive arterial pressure exerted by the pumping heart is responsible for the movement of liquid from the circulatory system (red) into the interstitial spaces of the tissues, so that all cells of the body have access to nutrients and can dispose of waste. This interstitial fluid, whose volume is roughly three times that of all blood in the circulation, is returned to the circulation in the form of lymph, which passes through specialized anatomic structures called *lymph nodes*. The primary lymphoid organs, where lymphocytes are generated, are the bone marrow (B cells, T-cell precursors) and the thymus (T cells). The initiation of an immune response involves the secondary lymphoid organs (lymph nodes, spleen).

The immune system, an interconnected system of vessels, organs, and cells, can be divided into primary and secondary lymphoid organs (see Figure 24-2). *Primary lymphoid organs* are the sites at which **lymphocytes** — the subset of leukocytes that includes B and T cells — are generated and acquire their functional properties. These organs include the thymus, where T cells are generated from

progenitors, and the bone marrow, where B cells are generated. Adaptive immune responses, which require functionally competent lymphocytes, are initiated in *secondary lymphoid organs*, which include lymph nodes and the spleen. All of the cells within lymphoid organs are ultimately derived from hematopoietic stem cells (see Figure 22-18), generated initially in the fetal liver and subsequently in the bone marrow. The total number of lymphocytes in a young man is estimated to be 500×10^9. Roughly 15 percent of these cells are found in the spleen, 40 percent in the other secondary lymphoid organs (tonsils, lymph nodes), 10 percent in the thymus, and 10 percent in the bone marrow; the remainder circulate in the bloodstream or reside in other tissues.

In normal circumstances, the pressure exerted by the pumping heart not only drives transport of the blood within blood vessels, but also forces cell-free liquid across blood vessel walls into the underlying tissue. This liquid delivers both nutrients and proteins, some of which carry out defensive functions. Its volume is up to three times the total blood volume. To maintain homeostasis, the fluid that leaves the circulation must ultimately return, and it does so in the form of *lymph*, via lymphatic vessels. At their most distal ends, lymphatic vessels are open to collect the interstitial fluid that bathes the cells in tissues. The lymphatic vessels merge into larger collecting vessels, which deliver lymph to *lymph nodes* (Figure 24-3). A lymph node consists of a capsule organized into areas that are defined by the cell types that inhabit them. Blood vessels entering a

lymph node deliver B and T cells to it. The lymph that arrives in a lymph node carries soluble antigens and specialized cells that have encountered ("sampled") antigens from the tissue drained by that particular afferent lymphatic vessel. In the lymph node, the cells and molecules required for the adaptive immune response interact, respond to the newly acquired antigenic information, and then execute the necessary steps to rid the body of the pathogen (see Figure 24-3).

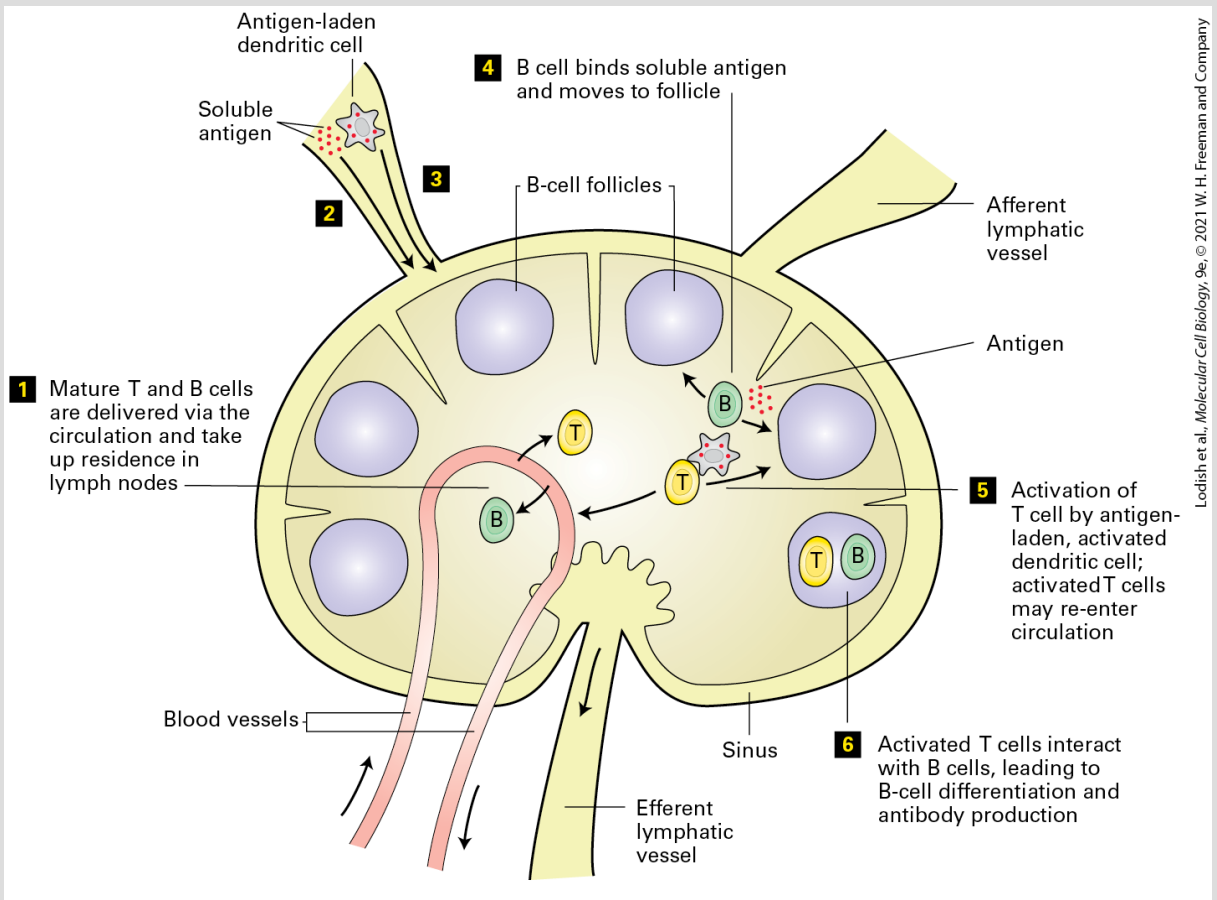


FIGURE 24-3 Initiation of the adaptive immune response in lymph nodes. Recognition of antigen by B and T cells (lymphocytes) located in lymph nodes initiates an adaptive immune response. T and B lymphocytes leave the circulation and take up residence in lymph nodes (step 1). Lymph carries antigen in two forms, soluble antigen and antigen-laden dendritic

cells; both are delivered to lymph nodes via afferent lymphatic vessels (steps 2 and 3). Dendritic cells, thus named because of their spine-like projections, are specialized cells of a type termed “antigen-presenting cells”; they degrade antigens into peptides or other small fragments and display them by means of MHC proteins on their surface, thus “presenting” antigens to T cells. Soluble antigen is recognized by B cells (step 4), and antigen-laden dendritic cells present antigen to T cells (step 5). Productive interactions between T and B cells (step 6) allow B cells to move into follicles and differentiate into plasma cells, which produce large amounts of secreted immunoglobulins (antibodies). Efferent lymphatic vessels return lymph from the lymph node to the circulation.

Lymph nodes can be thought of as filters in which antigenic information gathered from distal sites throughout the body is collected and displayed to the immune system in a form suitable to evoke a response. All the relevant steps that lead to activation of a resting lymphocyte take place in lymphoid organs. Cells that have received proper instructions to become functionally active leave the lymph node via efferent lymphatic vessels that ultimately return lymph to the bloodstream. Such activated cells recirculate through the bloodstream and, now ready for action, may reach a location where they again leave the circulation in response to chemotactic cues, move into tissues, and seek out pathogenic invaders, destroy virus-infected cells, or produce the antibodies that recognize and tag the invaders for destruction.

The exit of lymphocytes and other leukocytes from the circulation, the recruitment of these cells to sites of infection, the processing of antigenic information, and the return of immune-system cells to the circulation are all carefully regulated processes that involve specific

cell-adhesion events, chemotactic cues, and the crossing of endothelial barriers, as we will see later in this chapter.

Having introduced the principal types of cells that form the immune system, we return to the first two lines of defense against foreign pathogens: physical barriers and chemical defenses, and an innate immune system.

Mechanical and Chemical Boundaries Form a First Layer of Defense Against Pathogens

As noted already, mechanical and chemical defenses form the first line of host defense against pathogens (see Figure 24-1). Mechanical defenses, which operate continuously, include skin, epithelia, and arthropod exoskeletons, all barriers that can be breached only by mechanical damage or through specific enzymatic attack. Chemical defenses include the low pH found in gastric secretions as well as enzymes such as *lysozyme*, found in tears and in intestinal secretions, that can attack microbes directly.

The essential nature of mechanical defenses is immediately obvious in the case of burn victims. When the integrity of the skin (epidermis and dermis; Figure 1-4) is compromised, the rich source of nutrients in the underlying tissues is exposed, and airborne

bacteria or otherwise harmless commensal bacteria found on the skin can multiply unchecked, ultimately overwhelming the host. Viruses and bacteria have evolved strategies to breach the integrity of these physical barriers. For example, certain pathogenic bacteria (e.g., "flesh-eating bacteria," which are highly pathogenic strains of *Streptococcus*) secrete collagenases that compromise the integrity of connective tissue and so facilitate access of the bacteria to underlying tissue.

Innate Immunity Provides a Second Line of Defense

The innate immune system is activated once the mechanical and chemical defenses have failed and the presence of an invader is sensed (see Figure 24-1). The innate immune system comprises cells and molecules that are always available for immediate response to pathogens. **Phagocytes**, cells that ingest and destroy pathogens (see Figure 17-19), have receptors on their surface that bind to molecules frequently present on pathogens but absent from normal body cells. Many yeasts and other fungi, for instance, have cell walls that contain polymers of the sugar mannose. Such polymers are not found on normal body proteins, and macrophages utilize a *cell surface mannose receptor* to bind many fungal pathogens. Macrophages are widespread throughout tissues and epithelia and can be recruited to sites of infection. Animals that lack an adaptive

immune system, such as insects, rely exclusively on innate defenses to combat infections. Likewise, plants rely exclusively on innate defenses and lack adaptive immunity altogether. Cells of the innate immune system, otherwise resembling lymphocytes, are now categorized as innate lymphoid cells (ILCs) and will be discussed later in this chapter.

Phagocytes and Antigen-Presenting Cells

The innate immune system includes macrophages, neutrophils, and dendritic cells. All of these cells are phagocytic and come equipped with other types of pathogen recognition receptors such as **Toll-like receptors** (**TLRs**; see Figure 24-37 for their molecular structure) and *scavenger receptors* on their cell surface. Like the mannose receptor, these receptors detect broad patterns of pathogen-specific markers, such as bacterial cell-wall constituents or nucleic acids that contain unmethylated CpG or double-stranded RNA, and are thus key sensors for detecting the presence of bacteria or viruses. When these markers bind to TLRs, the cells produce effector molecules, including antimicrobial peptides. Dendritic cells and macrophages whose TLRs have detected pathogens also function as **antigen-presenting cells (APCs)** by processing and displaying foreign materials to antigen-specific T cells, thus bridging the innate and adaptive immune systems. The structure and function of TLRs and their role in activating dendritic cells are described in detail in Section 24.6.

Inflammasomes and Non-TLR Nucleic Acid Sensors

Mammalian cells possess a family of proteins that are capable of recognizing all manner of nonself components and of perceiving danger signals. The molecules recognized by these proteins span a range from components of the bacterial cell wall to uric acid crystals, to heme degradation products, and even to asbestos and silica (Figure 24-4). Once recognized, these danger signals activate the assembly of a multiprotein complex called the **inflammasome**, which activates the effector proteins involved in inflammation. Proteins that make up the inflammasome contain modules that mediate interactions with adapter proteins that ultimately allow a physical connection with and activation of caspase-1, an enzyme that is critical in the production of cytokines that cause inflammation (a process described below). As we will see in Section 24.6, the inflammasome plays an important role in bridging the innate and adaptive immune response.

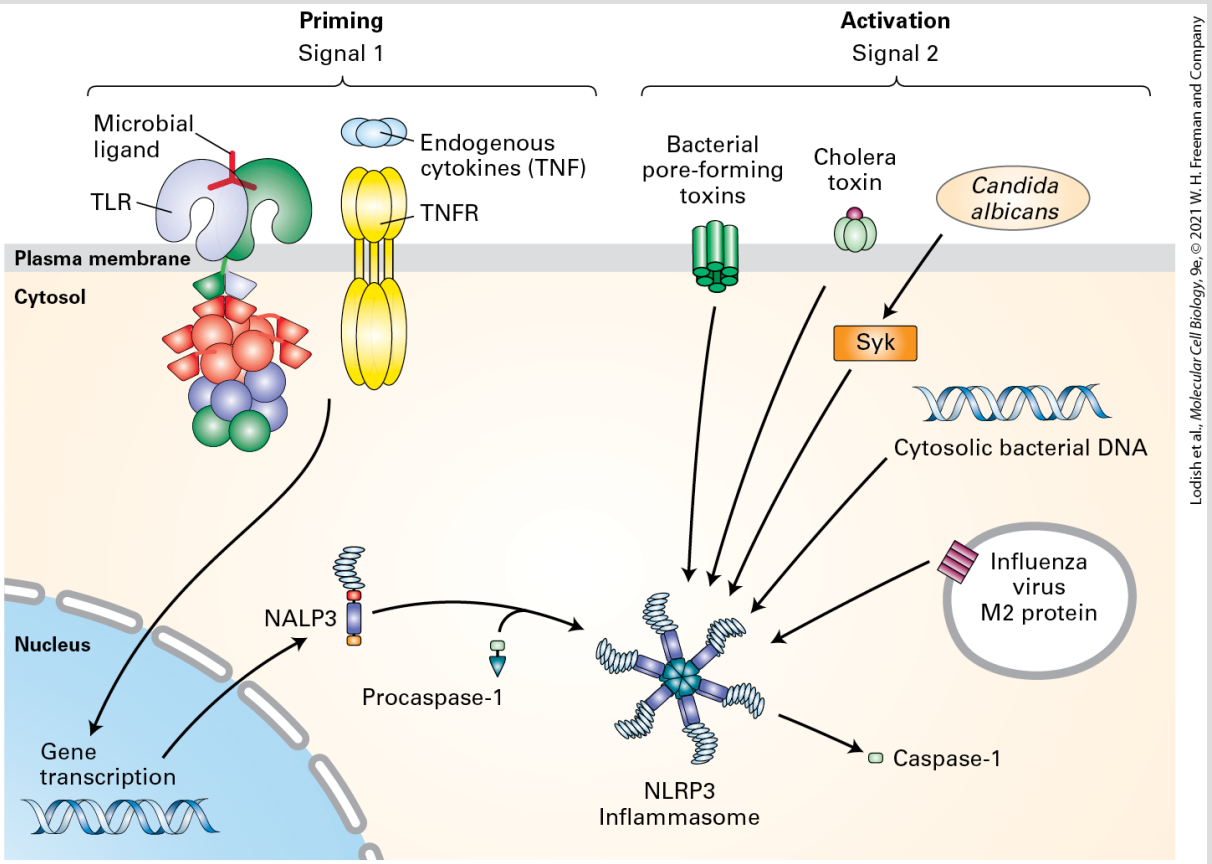
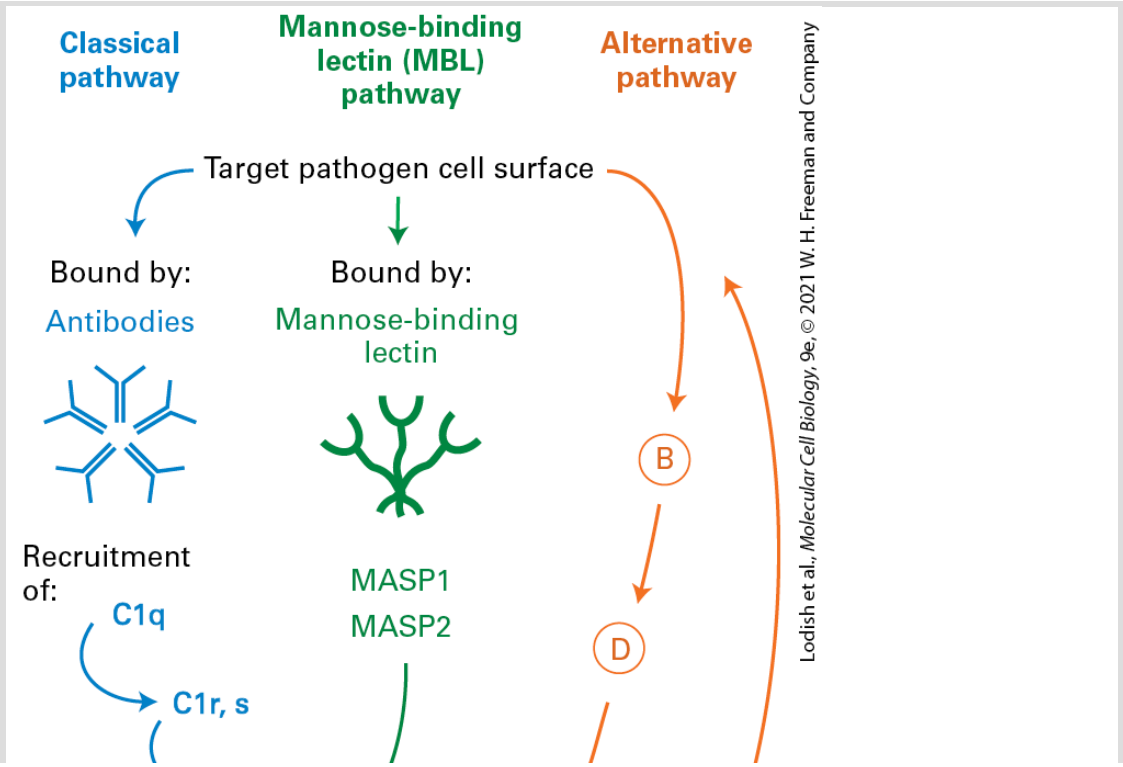


FIGURE 24-4 The NLRP3 inflammasome. The NLRP3 inflammasome activates caspase-1 only after receiving two signals. Signal 1 is provided by microbial antigens recognized via Toll-like receptors (TLRs) or by binding of endogenous cytokines such as TNF to the TNF receptor (TNFR). Signal 1 causes the up-regulation of NLRP3 (also called *NALP3*) and pro-IL-1β. Signal 2, which activates the NLRP3 inflammasome, can be provided by bacterial pore-forming toxins, by influenza virus M2 protein, by fungal particles via the kinase Syk (as shown for *Candida albicans*), or by cholera toxin. Cytosolic bacterial DNA can also activate the NLRP3 inflammasome, although the molecular details of this mechanism are not yet understood.

The Complement System

Another important component of the innate immune system is the complement system, a collection of serum proteins that can bind

directly to microbial or fungal surfaces. This binding initiates a cascade of protease activation that culminates in, among other things, the formation of a *membrane attack complex*, a multiprotein complex that forms pores in the pathogen's protective membrane (Figure 24-5). The cascade of complement activation is conceptually similar to the blood-clotting cascade, with amplification of the reaction at each successive stage of activation. At least three distinct pathways can activate the complement system. The *classical pathway* requires the presence of antibodies produced in the course of an adaptive immune response and bound to their antigens on the surface of the target microbe. How such antibodies are produced will be described in Section 24.3. This complement pathway represents an example of components of the innate immune system acting in concert with the adaptive immune system.



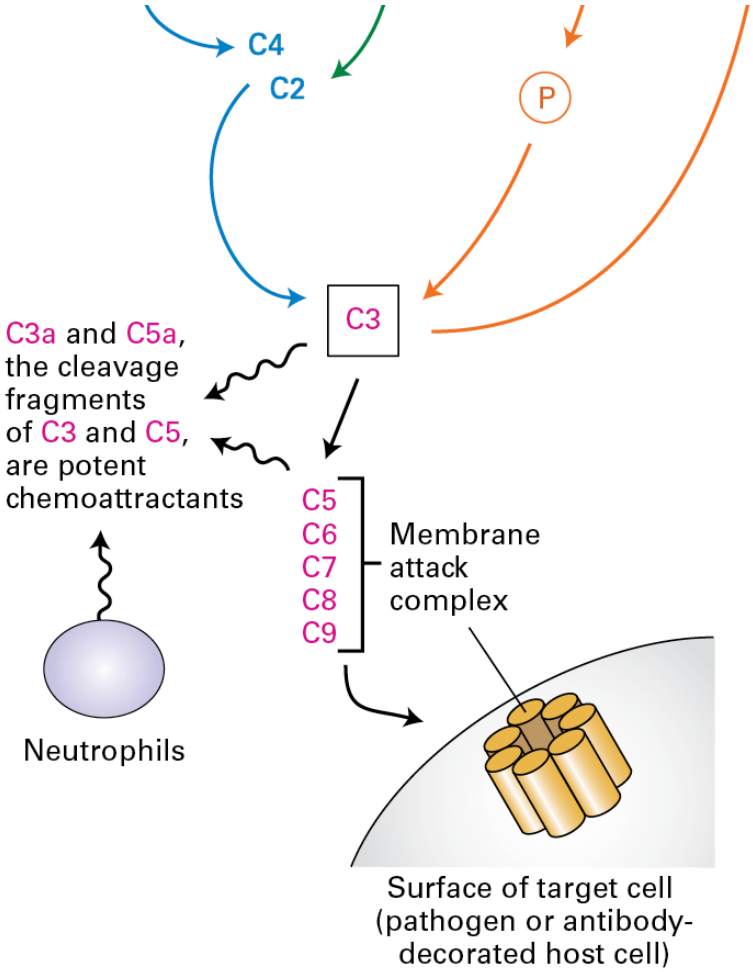


FIGURE 24-5 Three pathways of complement activation. The classical pathway involves the formation of antibody-antigen complexes. In the mannose-binding lectin pathway, mannose-rich structures found on the surfaces of many pathogens are recognized by mannose-binding lectin. The alternative pathway requires deposition of a special form of the serum protein C3, a major complement component, onto a microbial surface, upstream of which are factors B, D, and P. Each of the activation pathways is a cascade of proteases in which the downstream component is itself a protease. Amplification of activity occurs with each successive step. All three pathways converge on C3, which cleaves C5 and thus triggers formation of the membrane attack complex, leading to destruction of target cells. The small fragments of C3 and C5 generated in the course of complement activation are chemoattractants. They initiate inflammation by attracting neutrophils, phagocytic cells that can kill bacteria at short range or upon ingestion.

In addition to the classical pathway of complement activation, pathogens that contain mannose-rich cell walls activate the complement cascade through the *mannose-binding lectin pathway*. (Note that this is distinct from the macrophage mannose receptor mentioned previously.) Mannose-binding lectin binds to distinctive groups of mannose sugars on the surface of the pathogen, which triggers activation of two mannose-binding lectin–associated proteases, MASP-1 and MASP-2. Protease activation leads to activation of the downstream components of the complement cascade as shown in Figure 24-5. Finally, many microbial surfaces have physical and chemical properties, albeit incompletely understood, that result in activation of complement via the *alternative pathway*, an activation cascade that includes factors B, D, and P, all proteins found in plasma.

The three pathways converge on the activation of complement protein C3. This protein is synthesized as a precursor that contains an internal, strained thioester linkage between a cysteine and a glutamate residue, requiring proteolytic conversion to become fully reactive. C3 is deposited only on antigen-antibody complexes that are nearby. Surfaces that are properly decorated with mannose-binding lectin or that receive C3 deposits via the alternative pathway are similarly targeted. This proximity restriction limits the effects of complement to nearby surfaces, avoiding an inappropriate attack on cells that do not display the targeted antigens.

Regardless of the activation pathway, activated C3 unleashes the terminal components of the complement cascade, complement proteins C5 through C9, culminating in formation of the membrane attack complex, which inserts itself into almost any adjacent biological membrane and renders it permeable by forming a pore. The resulting loss of electrolytes and small solutes leads to lysis and death of the target cell. Whenever complement is activated, the membrane attack complex is formed and results in death of the cell onto which it is deposited. The direct microbe-killing (microbicidal) effect of a fully activated complement cascade is an important mechanism of host defense.

All three complement activation pathways also generate C3a and C5a cleavage fragments, which bind to G protein–coupled receptors and function to attract neutrophils and other cells involved in inflammation. In addition, phagocytic cells, such as **macrophages**, which recognize cells whose surfaces are covalently labeled with fragments from C3, ingest and destroy those cells.

The complement cascade thus fulfills multiple roles in host defense: it can destroy the membranes that envelope a pathogen (bacteria, viruses); it covalently decorates the targeted pathogen so that it may be more readily ingested by phagocytic cells capable of killing the pathogen and presenting its contents to cells that will initiate an adaptive immune response. Finally, the act of complement activation yields signals to attract cells of the innate (neutrophils, macrophages, dendritic cells) and adaptive (lymphocytes) immune

systems to the site of infection. These cues are called *chemotactic signals*.

Natural Killer Cells

In addition to bacterial and eukaryotic parasitic invaders, the innate immune system also defends against viruses. When the presence of a virus-infected cell is detected, still other cell types of the innate immune system become active, seeking out virus-infected target cells and killing them. For instance, when many types of cells (not just immune-system cells) are infected, they synthesize and secrete a class of proteins called *type I interferons* that act as intercellular signals, warning the immune system that an infection is present. The interferons are classified as **cytokines**, small, secreted proteins that help regulate immune responses in a variety of ways. We will encounter other cytokines and discuss some of their receptors as the chapter progresses.

Interferons activate **natural killer (NK) cells**, which belong to the group of *i*nnate *l*ymphoid *c*ells-1 (ILC1). These innate lymphoid cells will be discussed in Section 24.5. Activated NK cells help protect the body in several ways. First, they can kill host cells infected by a virus (hence the name "natural killer"), preventing those infected cells from making additional virus particles that would spread the infection. Second, NK cells secrete type II interferon γ which is essential for orchestrating many other aspects of antiviral defenses

(Figure 24-6). Third, NK cells can kill target cells that have been decorated by antibodies. NK cells recognize their targets by means of several classes of surface receptors capable of yielding stimulatory (promoting cell killing) or inhibitory signals.

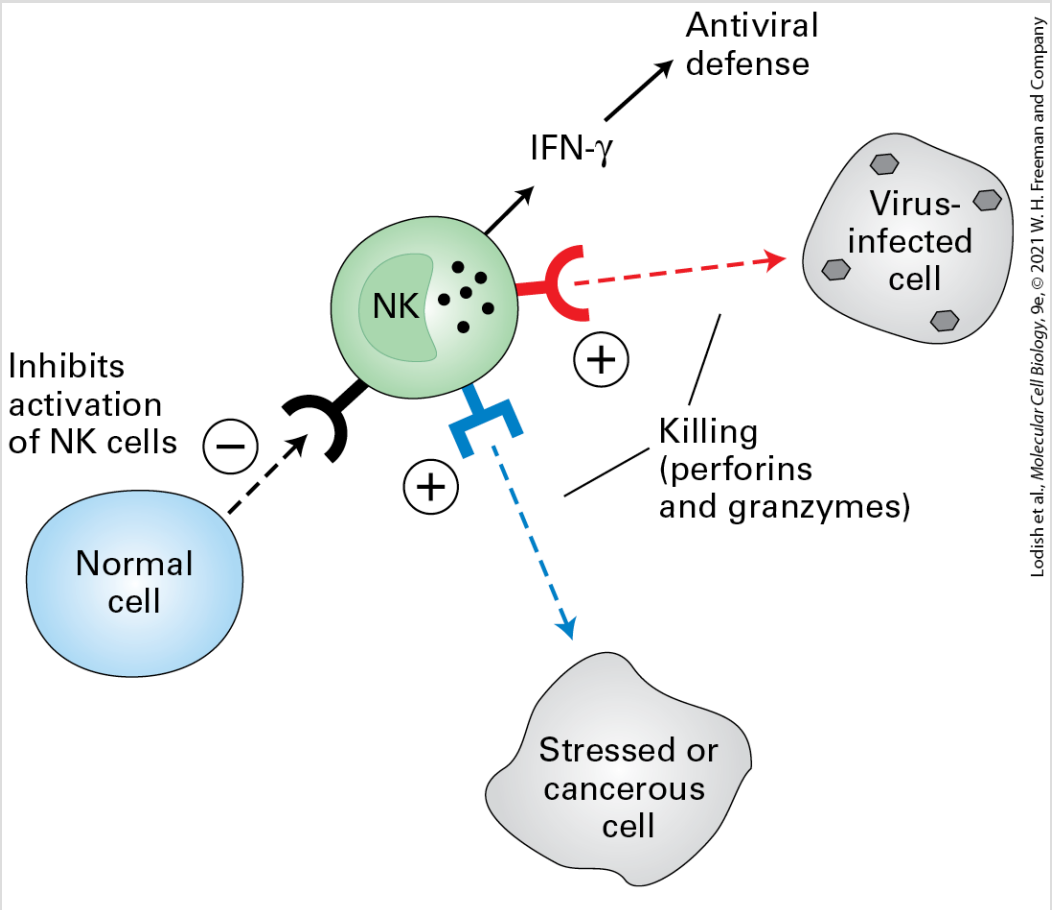


FIGURE 24-6 Natural killer cells. Natural killer (NK) cells or ILC1s are an important source of the cytokine interferon γ (IFN-γ), which is involved in antiviral defenses, and can kill virus-infected and cancerous cells directly by means of perforins. These pore-forming proteins allow access to the cytoplasm of the target cell by serine proteases called *granzymes*. Granzymes can also initiate apoptosis through activation of caspases (see Chapter 22). Receptors on NK cells identify infected or stressed cells and stimulate the NK cell to kill them. Other receptors identify normal cells and inhibit NK cell activation.

Inflammation Is a Complex Response to Injury That Encompasses Both Innate and Adaptive Immunity and Helps Destroy Pathogens

When a vascularized tissue (one that is supplied with blood vessels) is injured, the stereotypical response that follows is **inflammation**. The injury may be a consequence of physical or chemical processes, such as torn muscles, a simple paper cut, or infection with a pathogen. Inflammation, also called the *inflammatory response*, is characterized by four classic signs: *redness, swelling, heat*, and *pain*. These signs are caused by increased leakiness of blood vessels (vasodilation), attraction of immune-system cells to the site of damage, and the production of soluble mediators of inflammation, which are responsible for the sensation of heat and pain. Inflammation provides immediate protection through the activation of the cell types and soluble products that together mount the innate immune response and create a local environment conducive to the initiation of the adaptive immune response. If it is not properly controlled, however, inflammation can also be a major cause of tissue damage.

Figure 24-7 depicts the key players in the inflammatory response to bacterial pathogens and the subsequent initiation of an adaptive immune response. Tissue-resident **dendritic cells** sense the

presence of pathogens via their TLRs and respond by releasing small soluble proteins such as cytokines and **chemokines**; the latter act as chemoattractants for immune-system cells. **Neutrophils** leave the circulation and migrate to the site of injury or infection in response to the cytokines and chemokines produced there (see Figure 20-40). Neutrophils, which constitute almost half of all circulating leukocytes, are phagocytic (see Figure 17-19), directly ingesting and destroying pathogenic bacteria and fungi. Neutrophils can interact with a wide variety of pathogen-derived macromolecules via their TLRs. Engagement of these receptors, described in detail below, activates the neutrophils, which produce more cytokines and chemokines. The latter can attract more leukocytes — neutrophils, macrophages, and ultimately lymphocytes (T and B cells) — to the area to fight the infection. Activated neutrophils can also release bacteria-destroying enzymes (e.g., lysozyme and proteases) as well as small peptides with microbicidal activity, collectively called *defensins*. Activated neutrophils also turn on enzymes that generate the superoxide anion radical and other reactive oxygen species (see Section 12.4), which can kill microbes at short range. Neutrophils can also undergo a reaction referred to as *NETosis*. By committing suicide and releasing their nuclear DNA into the extracellular space, neutrophils throw out a fibrous net that can capture microbes and platelets (Figure 24-8). This reaction helps confine invaders to the site where neutrophil activation occurs. Another cell type that contributes to the inflammatory response is the tissue-resident *mast*

cell. When activated by a variety of physical or chemical stimuli, mast cells release histamine, a small molecule that binds to G protein–coupled histamine receptors. This binding leads to increased vascular permeability and thereby facilitates access to the tissue by plasma proteins (e.g., complement) that can act against the invading pathogen.

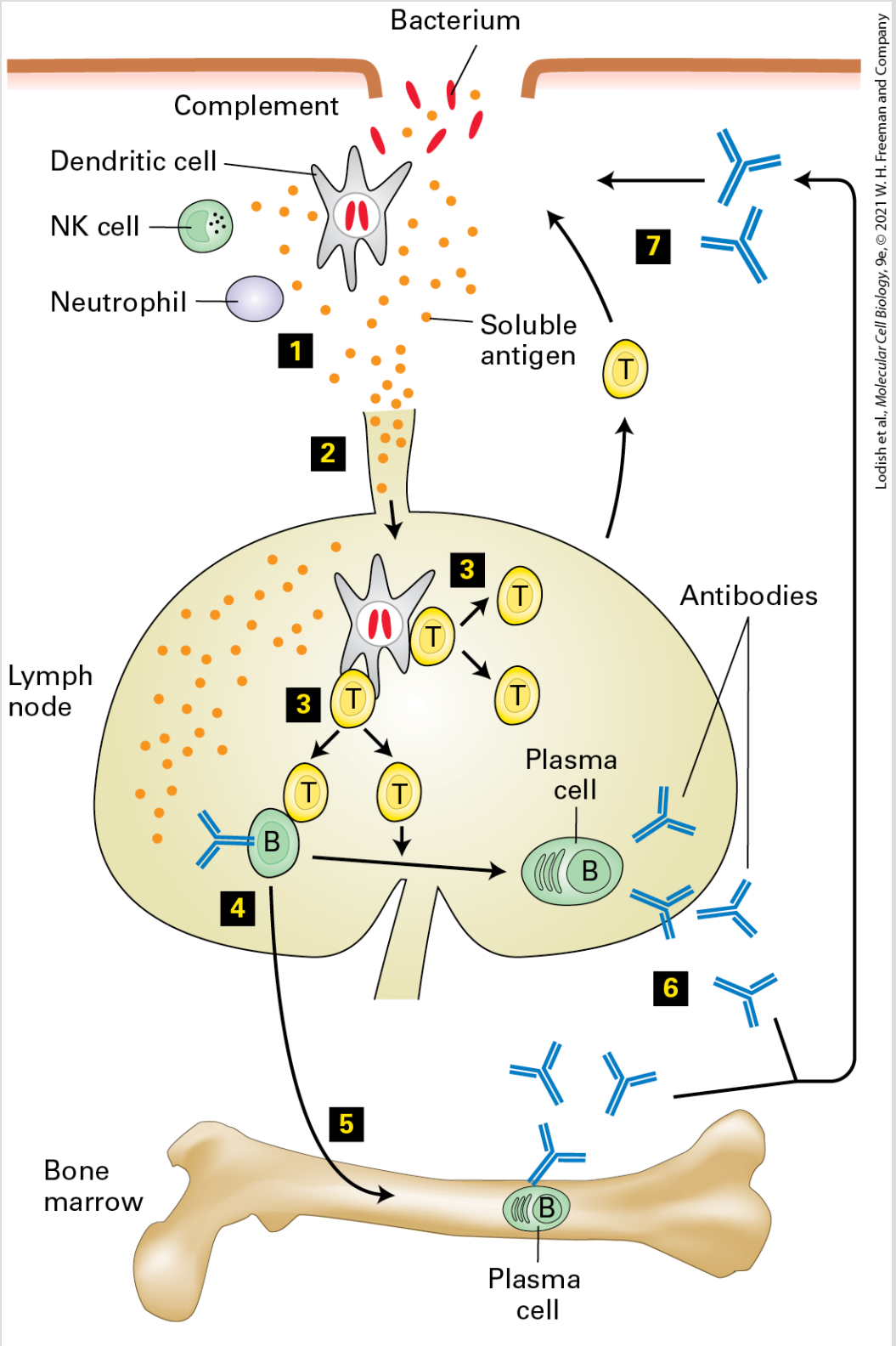


FIGURE 24-7 Interplay of innate and adaptive immune responses to a bacterial pathogen. Once a bacterium breaches the host's mechanical and

chemical defenses, the bacterium is exposed to components of the complement cascade, as well as to innate immune-system cells that confer immediate protection (step 1). Various inflammatory proteins induced by tissue damage contribute to a localized inflammatory response. Local destruction of the bacterium results in the release of bacterial antigens, which are delivered to the lymph nodes (step 2) via the afferent lymphatic vessels that drain the tissue. Dendritic cells acquire antigen at the site of infection, become migratory, and move to the lymph nodes, where they activate T cells (step 3). In the lymph nodes, antigen-stimulated T cells proliferate and acquire effector functions, including the ability to help B cells (step 4), some of which may move to the bone marrow and complete their differentiation into antibody-secreting plasma cells (step 5). In later stages of the immune response, activated T cells provide additional assistance to antigen-experienced B cells to yield plasma cells that secrete antigen-specific antibodies at a high rate (step 6). Antibodies produced as a consequence of the initial exposure to the bacterium act in synergy with complement to eliminate the infection (step 7), should it persist, or afford rapid protection in the case of re-exposure to the same pathogen.

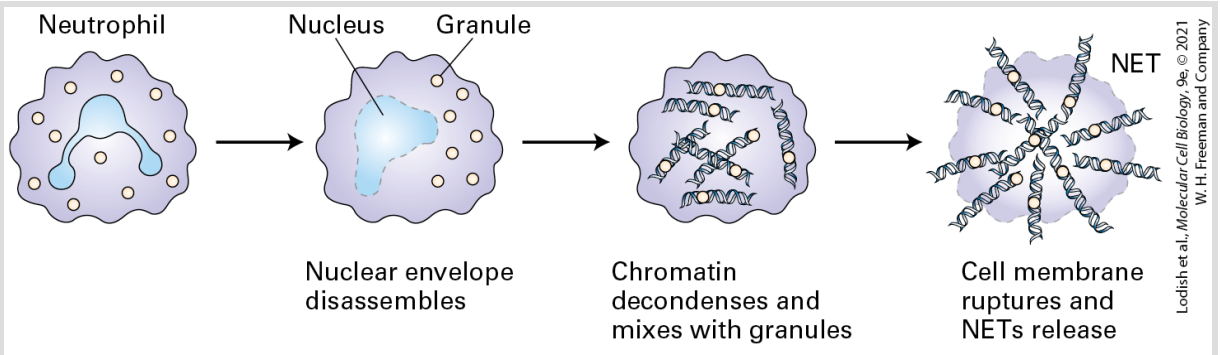


FIGURE 24-8 In NETosis neutrophils sacrifice themselves to facilitate capture of bacterial pathogens. Upon encounter with bacterial products such as lipopolysaccharides, neutrophils can undergo a form of cell death that results in the release of a network of nuclear DNA and associated molecules (a neutrophil extracellular trap or “NET”). NETs contain antimicrobial components and can kill bacteria extracellularly without reliance on phagocytosis. NETs can physically limit further spread of the pathogens.

A very important early response to infection or injury is the activation of a variety of plasma proteases, including the proteins of the complement cascade discussed above. As we have seen, the cleavage fragments produced during activation of these proteases attract neutrophils to the site of tissue damage (see Figure 24-5). They further induce production of cytokines such as interleukins 1 and 6 (IL-1 and IL-6), which cause inflammation. The recruitment of neutrophils also depends on an increase in vascular permeability, which is controlled in part by lipid-signaling molecules (e.g., prostaglandins and leukotrienes) that are derived from phospholipids and fatty acids. All of these events occur rapidly, starting within minutes of injury. A failure to resolve the cause of this immediate response may result in chronic inflammation with ensuing tissue damage, in which cells of the adaptive immune system play an important role.

When the pathogen burden at the site of tissue damage is high, it may exceed the capacity of innate defense mechanisms to deal with the infection. Moreover, some pathogens have acquired, in the course of evolution, tools to disable or bypass innate immune defenses. In such situations, the adaptive immune response is required to help control the infection. This response depends on specialized cells that are parts of both the adaptive and innate immune systems, including macrophages and dendritic cells, which are capable of ingesting and killing pathogens as well as presenting antigens to the adaptive immune system. Dendritic cells, in particular, can initiate an adaptive immune response by delivering

newly acquired pathogen-derived antigens to secondary lymphoid organs (see Figure 24-7).

Adaptive Immunity, the Third Line of Defense, Exhibits Specificity

Adaptive immunity is the term reserved for the highly specific recognition of foreign substances by antigen-specific receptors, the full elaboration of which requires days or weeks after occurrence of the initial exposure. Lymphocytes bearing antigen-specific receptors are the key cells responsible for adaptive immunity. The B lymphocytes will differentiate into plasma cells that produce a secreted version of their cell surface antigen receptors, proteins called *immunoglobulins* or *antibodies*. Immunoglobulins can neutralize (render inactive) not only bacterial toxins but also harmful agents such as viruses by binding directly to them in a manner that prevents the virus from attaching itself to host cells. The generation of neutralizing antibodies is the rationale underlying virtually all vaccination strategies. Vaccination is a form of active immunization that consists of deliberately exposing an individual to a foreign antigen to elicit protective immunity by generating an adaptive immune response (described below) and antibodies that bind the foreign antigen.

In the same vein, antibodies raised against snake venoms can be administered to the victims of snake bites to protect them from intoxication, provided the administration occurs relatively soon after the bite. The antibodies bind to the toxic proteins in the venom, keeping them from binding to their targets in the host, and in so doing neutralize them. This procedure, called *passive immunization*, can save lives by instant neutralization of a noxious substance such as a toxin. Passive immunization is also used prophylactically to protect those who travel to areas where a disease such as viral hepatitis is endemic; administration of serum from immune individuals provides temporary protection against infection. Antibodies can thus have immediate protective effects. Given that today's medical advances allow the survival of individuals whose immune systems are severely compromised (e.g., cancer patients receiving chemotherapy or radiation, transplant patients with a pharmacologically suppressed immune system, patients who suffer from AIDS, individuals with inborn deficiencies of the immune system), passive immunization can be of immediate practical importance. The deliberate exposure of an animal such as a mouse or rabbit to a foreign substance (*immunization*) allows the production of *antisera* that specifically recognize that substance (the *antigen*). These antisera have become standard components of the cell biologist's toolbox.

KEY CONCEPTS OF SECTION 24.1

Overview of Host Defenses

- Mechanical and chemical defenses provide protection against most pathogens. This protection is immediate and continuous, yet possesses little specificity. Innate and adaptive immunity provide defenses against pathogens that breach the body's mechanical or chemical boundaries (see Figure 24-1).
- The circulatory and lymphatic systems distribute the molecular and cellular players in innate and adaptive immunity throughout the body (see Figure 24-2).
- Innate immunity is mediated by the complement system (see Figure 24-5) and several types of leukocytes, including natural killer cells, neutrophils, and other phagocytic cells, such as macrophages and dendritic cells. The cells and molecules of the innate immune system are deployed rapidly (minutes to hours). Molecular patterns diagnostic of the presence of pathogens can be recognized by Toll-like and other receptors, but the specificity of recognition is modest, as these receptors are capable of recognizing rather broad sets of related molecules.
- Adaptive immunity is mediated by T and B lymphocytes. These cells require days for full activation and deployment, but they can distinguish between closely related antigens. This specificity of antigen recognition is the key distinguishing feature of adaptive immunity.
- Innate and adaptive immunity act in a mutually synergistic fashion. Inflammation, an early response to tissue injury or infection, involves a series of events that combines elements of innate and adaptive immunity (see Figure 24-7).

24.2 Immunoglobulins: Structure and Function

Immunoglobulins (also called *antibodies*), produced by B cells, are the best understood of the molecules that confer adaptive immunity. An individual human has the capacity to make a limitless number of different antibodies, but any given specific antibody is typically made only when the individual has been exposed to the antigen (immunized) to which the antibody will bind specifically; hence antibody production is an adaptive immune response. In this section, we describe the structural organization of immunoglobulins, their diversity, and how they bind to antigens. The mechanisms that generate diverse antibodies are described in Section 24.3.

Immunoglobulins Have a Conserved Structure Consisting of Heavy and Light Chains

Immunoglobulins are abundant serum proteins that fall into several classes with distinct structural and functional properties. Immunoglobulins were identified as the class of serum proteins responsible for antibody activity when they were biochemically purified from serum isolated from immunized animals (called

antiserum). They were purified based on their abilities to mediate the killing of microbes and to bind directly to their corresponding, or *cognate*, antigens. Immunoglobulins of the most common class are composed of two identical *heavy (H) chains*, covalently attached to two identical *light (L) chains* (Figure 24-9; the various classes will be described in the next section). The typical immunoglobulin (sometimes abbreviated *Ig*) therefore has a twofold, symmetric structure, described as H_2L_2. One H_2L_2 antibody molecule can usually bind to two antigen molecules (bivalent binding; see below). An exception to this basic H_2L_2 architecture occurs in the immunoglobulins made by camelids (camels, llamas, vicuñas) and sharks. These animals can make immunoglobulins that are heavy-chain dimers (H_2) and lack light chains.

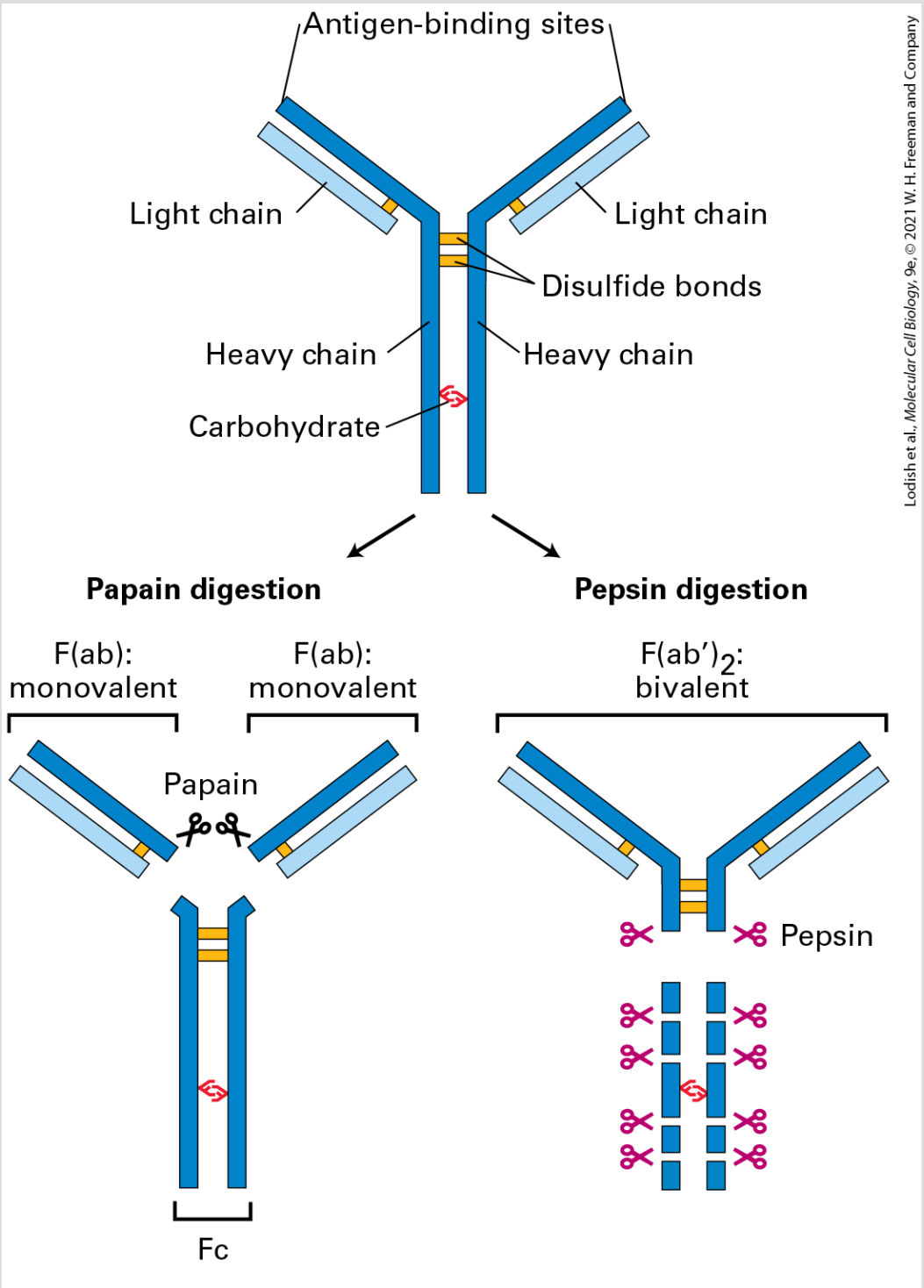


FIGURE 24-9 The basic structure of an immunoglobulin molecule. Antibodies are serum proteins also known as *immunoglobulins*. They are twofold, symmetric structures composed of two identical heavy chains and two identical light chains. Fragmentation of antibodies with proteases yields

fragments that retain antigen-binding capacity. The protease papain yields monovalent F(ab) fragments, and the protease pepsin yields bivalent $\mathrm{F(ab')_2}$ fragments. The Fc fragment is unable to bind antigen, but this portion of the intact molecule has other functional properties. F(ab) fragments are now commonly made using recombinant DNA technology instead of relying on proteolytic digestion.

A biochemical approach was used to answer the question of how antibodies manage to distinguish among related molecules, that is, how one antibody can bind to its specific antigen but not to another, structurally very similar, molecule. Proteolytic enzymes were used to fragment immunoglobulins, which are rather large proteins (~150 kDa), to identify the regions in the protein that are directly involved in antigen binding (Figure 24-9). The protease papain yields fragments, called *F(ab)* for *a*ntigen *b*inding *f*ragment, that can bind a single antigen molecule (monovalent fragments), whereas the protease pepsin yields bivalent fragments, referred to as $F(ab')_2$ (F = fragment; ab = antibody) that exhibit two binding sites, termed *bivalent binding*. These enzymes are used to convert intact immunoglobulin molecules into monovalent or bivalent reagents. Although F(ab) fragments are incapable of binding two antigens, a property often termed *cross-linking* when applied to antigens displayed at a cell surface, $\mathrm{F(ab')_2}$ fragments can do so. Researchers frequently take advantage of this property to cross-link and thus activate surface receptors. Many receptors, such as the EGF receptor, dimerize upon engagement of ligand (*ligand-induced dimerization*), a prerequisite for full activation of downstream

signaling cascades (see Chapter 16). Many receptors on immune-system cells behave in a similar fashion. The portion released upon papain digestion and incapable of antigen binding is called *Fc* because of its ease of crystallization (F = fragment; c = crystallizable).

Multiple Immunoglobulin Isotypes Exist, Each with Different Functions

Immunoglobulins can be divided into different classes, termed **isotypes**, based on their distinct biochemical properties. There are two light-chain isotypes, κ and λ. The heavy chains show more variation; in mammals, the major heavy-chain isotypes are μ, δ, γ, α and ε. These heavy chains can associate with either κ or λ light chains. Depending on the vertebrate species, further subdivisions occur within the α and γ isotypes, and fish possess an isotype that is not found in mammals. The fully assembled immunoglobulin (Ig) derives its name from the heavy chain: antibodies with μ heavy chains yield IgM; α chains, IgA; γ chains, IgG; δ chains, IgD; and ε chains, IgE. The general structures of the major Ig isotypes are depicted in Figure 24-10. By means of the unique structural features of the Fc portions of their heavy chains, each of the different Ig isotypes carries out specialized functions.

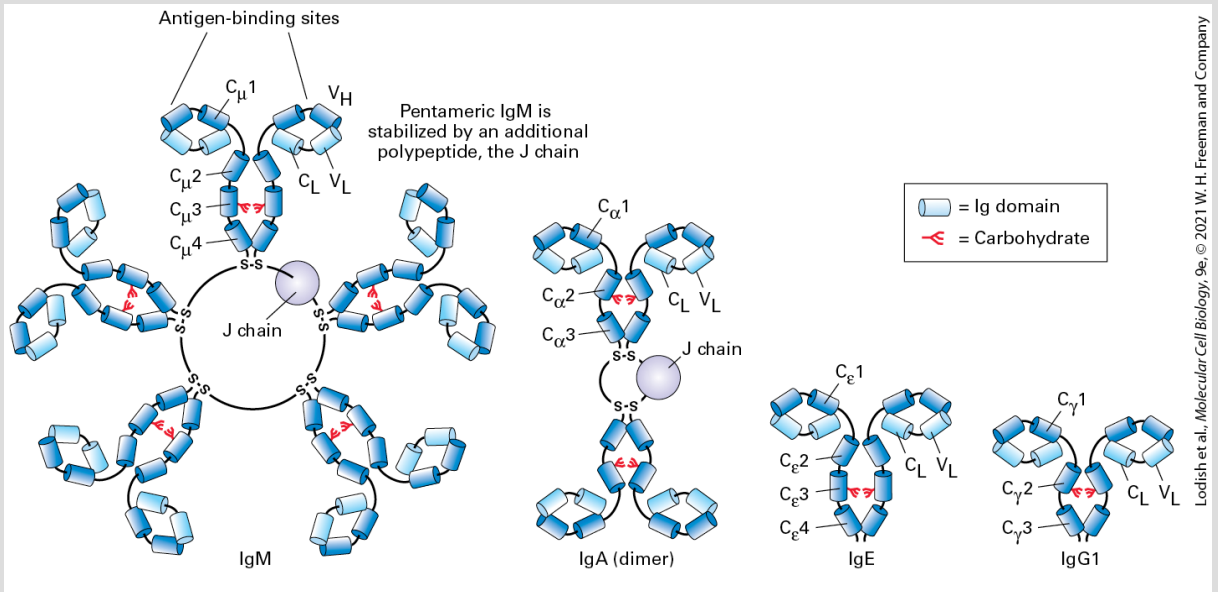
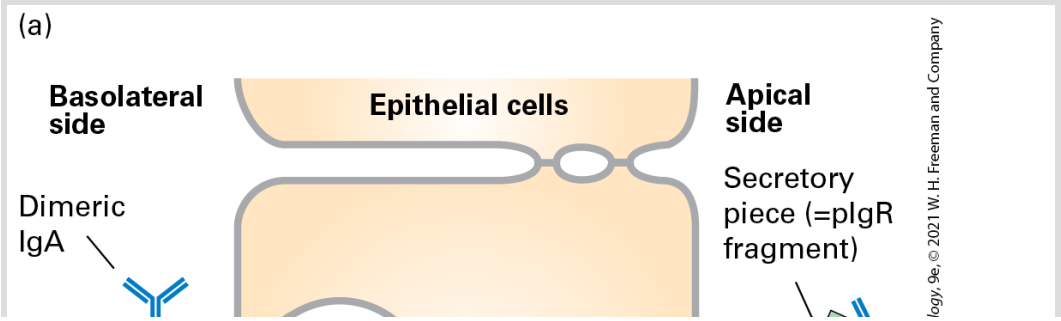


FIGURE 24-10 Immunoglobulin isotypes. The different classes of immunoglobulins, called *isotypes*, may be distinguished biochemically and by immunological techniques. In mice and humans, there are two light-chain isotypes (κ and λ) and five heavy-chain isotypes (μ, δ, γ, ε, α). Each isotype defines a class of immunoglobulin based on the identity of the heavy chain. IgG, IgE, and IgD (not shown) are monomers with generally similar overall structures. IgM and IgA can occur in serum as pentamers and dimers, respectively, accompanied by an accessory subunit, the J chain, in covalent disulfide linkage. This volume-rendered depiction of the immunoglobulins highlights their modular design, with each barrel representing an individual Ig domain. Different isotypes have different functions. See Figure 24-13 for definitions of abbreviations.

The IgM molecule is secreted as a pentamer of H_2L_2 chains, stabilized by disulfide bonds between the ends of the heavy chains and an additional chain, the J chain. In its pentameric form, IgM possesses 10 identical antigen-binding sites (2 for each H_2L_2), which allow high-avidity interactions with surfaces that display the cognate antigen. *Avidity* is defined as the product of the *strength* of interactions (affinity) of the available individual binding sites and the *number* of such binding sites. Combining many low-affinity

interactions can lead to a high-avidity interaction (as in Velcro). Upon its deposition on a surface that carries the cognate antigen, the pentameric IgM molecule assumes a conformation that is highly conducive to activation of the complement cascade; it thus is an effective means of damaging the membrane onto which it is adsorbed and onto which complement proteins are deposited as a consequence.

The IgA molecule also interacts with the J chain, forming a dimer of H_2L_2 molecules. Dimeric IgA can bind to an IgA receptor found on the basolateral plasma membrane of epithelial cells, where its binding triggers receptor-mediated endocytosis. Subsequently, the IgA receptor is cleaved, and the dimeric IgA, with a fragment of the receptor (the so-called "secretory piece") still attached, is released from the apical side of the epithelial cell. This process, called **transcytosis**, is an effective means of delivering immunoglobulins from the basolateral side of an epithelium to the apical side (Figure 24-11a). Tears and other secretions, especially in the gastrointestinal tract, are rich in IgA — grams of immunoglobulin are secreted from the blood into the lumen of the GI tract each day! — and so provide protection against environmental pathogens.



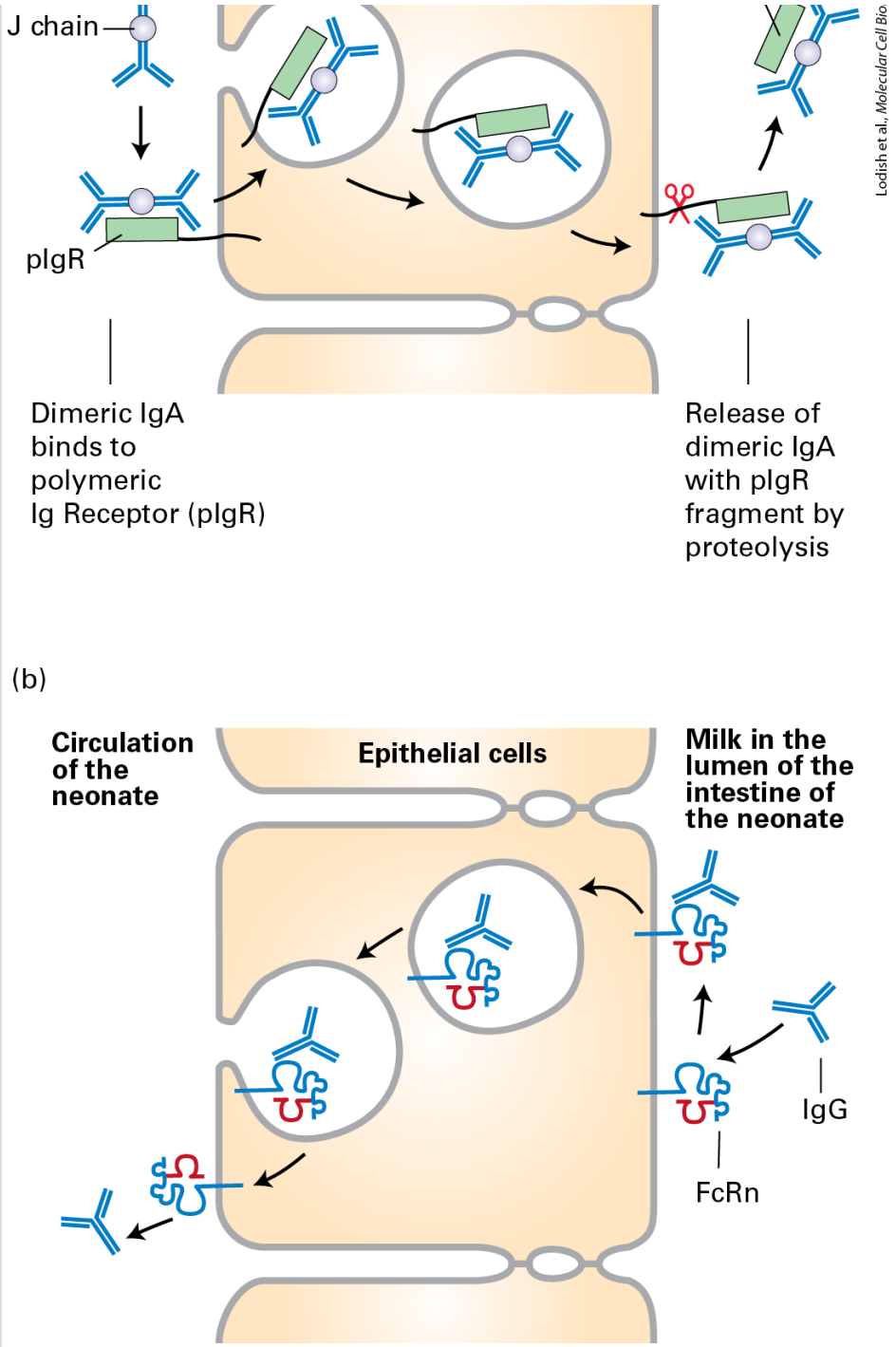


FIGURE 24-11 Transcytosis of IgA and IgG. (a) IgA, found in tears and in the secretions of various mucous membranes, must be transported across the epithelium. IgA binds to the polymeric IgA receptor on the basolateral surface of an epithelial cell and is endocytosed. As the resulting complex is transported across the epithelial monolayer, a portion of the receptor is cleaved, and the

IgA, still bound to a portion of the receptor, the secretory piece, is released at the apical side. (b) Suckling rodents acquire Ig from their mother's milk. At the apical surface of its intestinal epithelial cells, the newborn possesses the neonatal Fc receptor (FcRn), whose structure resembles that of class I MHC molecules (see Figure 24-24). After this receptor binds to the Fc portion of IgG, transcytosis moves the acquired IgG to the basolateral side of the epithelium. In humans, the syncytial trophoblast in the placenta expresses FcRn and so mediates acquisition of IgG from the maternal circulation and its delivery to the fetus (transplacental transport).

The IgG isotype is important for neutralization of virus particles. This isotype also helps prepare particulate antigens, such as viruses or larger fragments of bacteria, for acquisition by cells equipped with receptors specific for the Fc portion of the IgG molecule (see below).

The immune system of the newborn mammal is immature, but protective antibodies are transferred from the mother to the newborn via the mother's milk. The receptor responsible for capturing maternal IgG is the neonatal Fc receptor (FcRn), which in rodents is present on intestinal epithelial cells. Maternal IgG captured by an FcRn on the luminal side of the newborn's intestinal tract is transcytosed across the gut epithelium and released into the infant's circulation, generating passive immune protection of the infant rodent (Figure 24-11b). In humans, FcRn is found on fetal cells that contact the maternal circulation in the placenta. Transcytosis of IgG antibodies from the maternal circulation across the placenta delivers maternal antibodies to the fetus. These

maternal antibodies will protect the newborn until its immune system is sufficiently mature to produce antibodies on its own. In adults, FcRn is also expressed on endothelial cells and helps control the turnover of IgG in the circulation as well as the delivery of IgG across the endothelial barrier and into underlying tissue.

As we will see in Section 24.3, the IgM and IgD isotypes are expressed as membrane-bound receptors on newly generated B cells. Here the μ chains have an important role in B-cell development and activation.

Each Naive B Cell Produces a Unique Immunoglobulin

The *clonal selection theory* stipulates that each *naive* B lymphocyte (not yet having seen its specific antigen) expresses on its plasma membrane an antigen-binding receptor of unique specificity. The receptor is an antibody extended at its C-terminus by a hydrophobic sequence that anchors the protein in the plasma membrane, as we discuss later (see Figure 24-19). When a lymphocyte encounters the antigen for which it is specific, clonal expansion occurs. This allows an amplification of the response, culminating in production and secretion by plasma cells of large amounts of this specific antibody (the same one made by the original precursor cell) (Figure 24-12). (Note that clonal expansion is rapid cell division of this single cell to

form a group of cells — a clone — that all originate from this precursor cell.) The antigen-specific antibody is responsible for binding to the antigen and subsequently mediating the clearing of the antigen out of the body. In a typical immune response, the antigen that elicits the response is of complex composition: even the simplest virus contains several distinct proteins, and each protein may present to the immune system several molecularly distinct features that can be recognized independently of one another. Thus many individual lymphocytes respond to different parts of a given antigen and expand into independent clones in response to it, each producing its own antigen-specific receptor and antibody of unique structure and therefore with unique binding characteristics (affinity). Because each lymphocyte is endowed with a unique receptor and clonally expands in response to antigen, this response of multiple, independent B cell precursors is characterized as *polyclonal*.

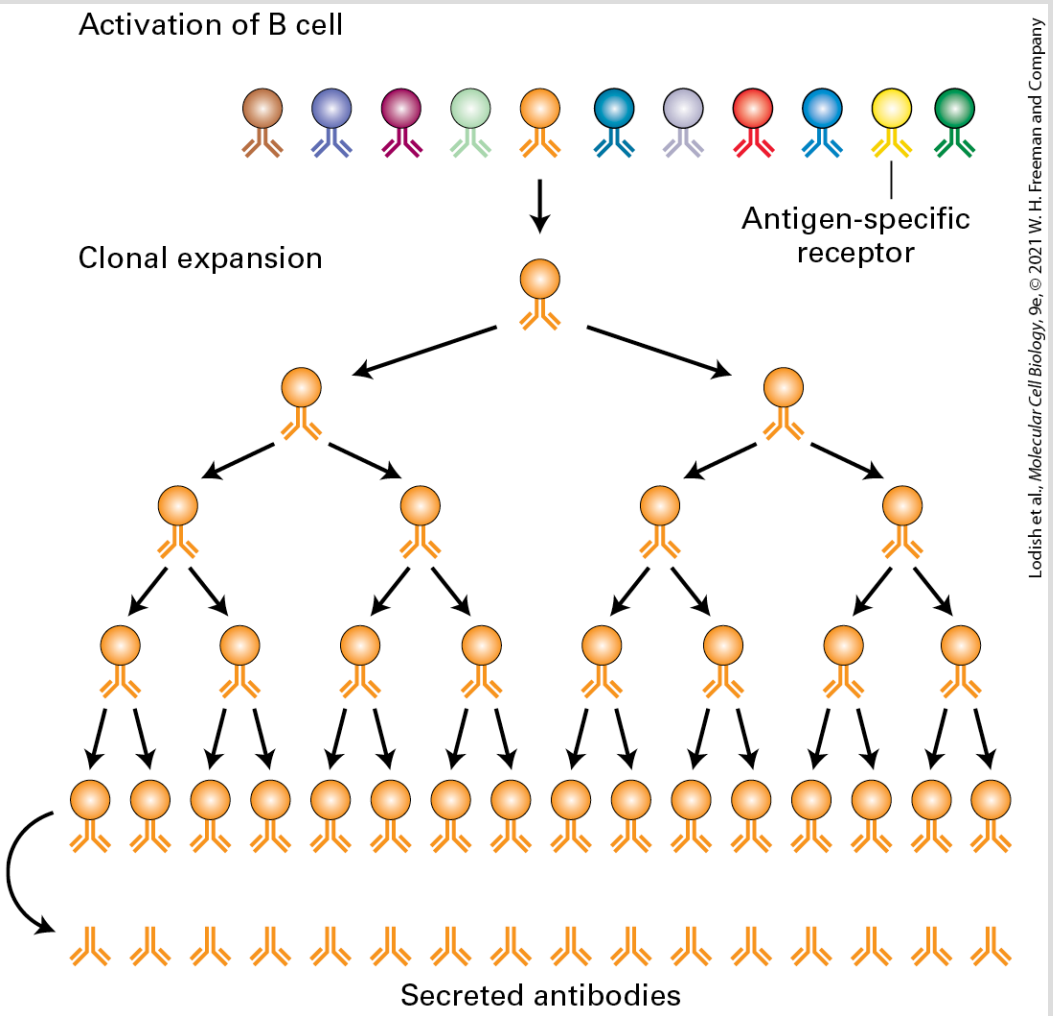


FIGURE 24-12 Clonal selection. The clonal selection theory proposes the existence of a large set of B lymphocytes, each expressing on its surface a unique antigen-specific receptor (indicated by different colors). The receptor is an antibody extended at its C-terminus by a hydrophobic sequence that anchors it in the plasma membrane (see Figure 24-19). The antigen that fits with the receptor carried by a particular lymphocyte binds to it and stimulates that lymphocyte to expand clonally. From a modest number of antigen-specific cells, a large number of cells of the desired specificity may be generated. Many of these cells will differentiate into plasma cells, where a form of the antigen-specific receptor lacking the membrane anchor is secreted as an antibody. Note that, like secreted antibodies, the cell surface antigen-specific receptors each have two H and two L chains.

B-cell tumors, which represent malignant clonal expansions of individual B-lymphocytes, enabled the first molecular analysis of the processes that underlie the generation of antibody diversity. A key observation was that tumors derived from lymphocytes may produce large quantities of a single type of secreted immunoglobulin. Some of the light chains of these immunoglobulins are secreted in the urine of patients with such tumors. These light chains, called *Bence-Jones proteins* after their discoverers, can be readily purified and afforded the first target for a protein chemical analysis of immunoglobulins.

Two key observations emerged from this work. First, no two independent B-cell tumors produced light chains with identical biochemical properties, suggesting that they were all unique in sequence. Second, the differences in amino acid sequence that distinguished one light chain from another were not randomly distributed but were clustered in a domain referred to as the *variable region of the light chain*, or V_L. This domain comprises the N-terminal ~110 amino acids of the light chain. The remainder of the sequence is identical for the different light chains (provided they derive from the same isotype, either κ or λ) and is therefore referred to as the *constant region*, or C_L. Immunoglobulins unique to each individual patient were subsequently purified from the patients' serum. Sequencing of the heavy chains from these preparations showed that the variable residues that distinguished one heavy chain from

another were again concentrated in a well-demarcated domain, referred to as the *variable region of the heavy chain*, or V_H.

An alignment of variable-region sequences obtained from different light chains showed a nonrandom pattern of regions of variability, revealing three *hypervariable regions* — HV1, HV2, and HV3 — that are sandwiched between what are called *framework regions* (Figure 24-13a). (Similar alignments for the immunoglobulin heavy-chain sequences also yielded hypervariable regions.) In the properly folded three-dimensional structure of immunoglobulins, these hypervariable regions are in close proximity (Figures 24-13b and 24-14) and make contact with antigen. Thus that portion of an Ig molecule containing the hypervariable regions constitutes the antigen-binding site. For this reason, hypervariable regions are also referred to as *complementarity-determining regions (CDRs)*.

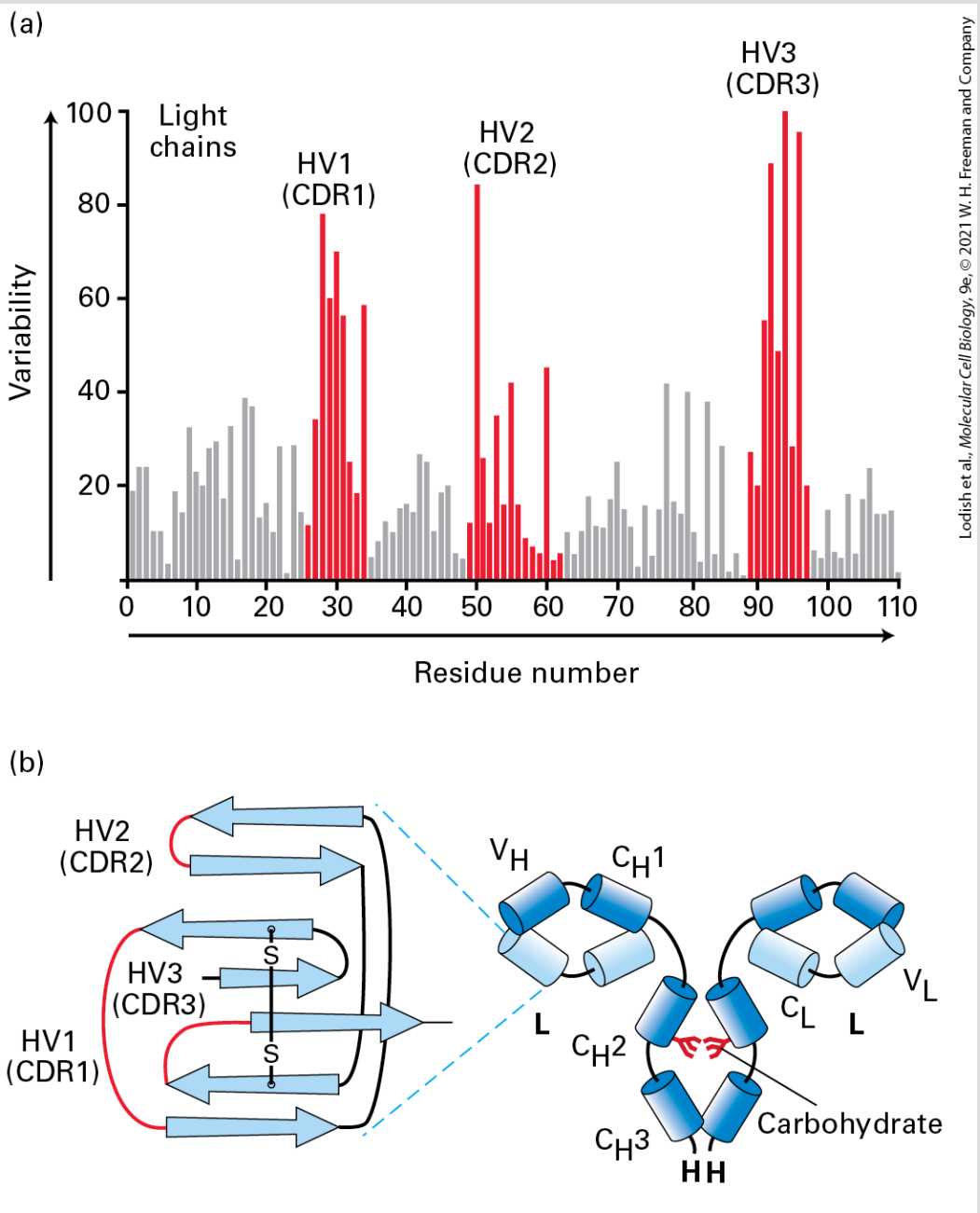


FIGURE 24-13 Hypervariable regions and the immunoglobulin fold. (a) Amino acid variability varies with residue position in Ig light chains. Here the percentage of variable-region sequences with variant amino acids is plotted for each position in the sequence. Positions for which many different amino acid side chains are present are assigned a high variability index; those that are invariant among the sequences compared are assigned a value of 0. This analysis reveals three regions of increased variability: hypervariability (HV) regions 1, 2, and 3; these regions are also called *complementarity-determining*

regions (CDRs) because the surface topology these regions create are complementary to the surface of the antigen to which the immunoglobulin binds. (b) Volume-rendered depiction of $F(ab')_2$ fragment (*right*) and ribbon diagram of a typical Ig light-chain variable region (V_L) with the positions of the hypervariable regions indicated in red (*left*). The hypervariable regions are found in the loops that connect the β strands and make contact with antigen. The β strands (rendered as arrows) make up two β sheets and constitute the framework region, which has a similar fold in all antibody domains. Each variable and constant domain has this characteristic three-dimensional structure, called the *immunoglobulin fold*. L = light chain; H = heavy chain; $V_H = \text{heavy-chain}$ variable region; $V_L = \text{light-chain}$ variable region; $C_H1, C_H2, C_H3 = \text{heavy-chain}$ constant domains; $C_L = \text{light-chain}$ constant region.

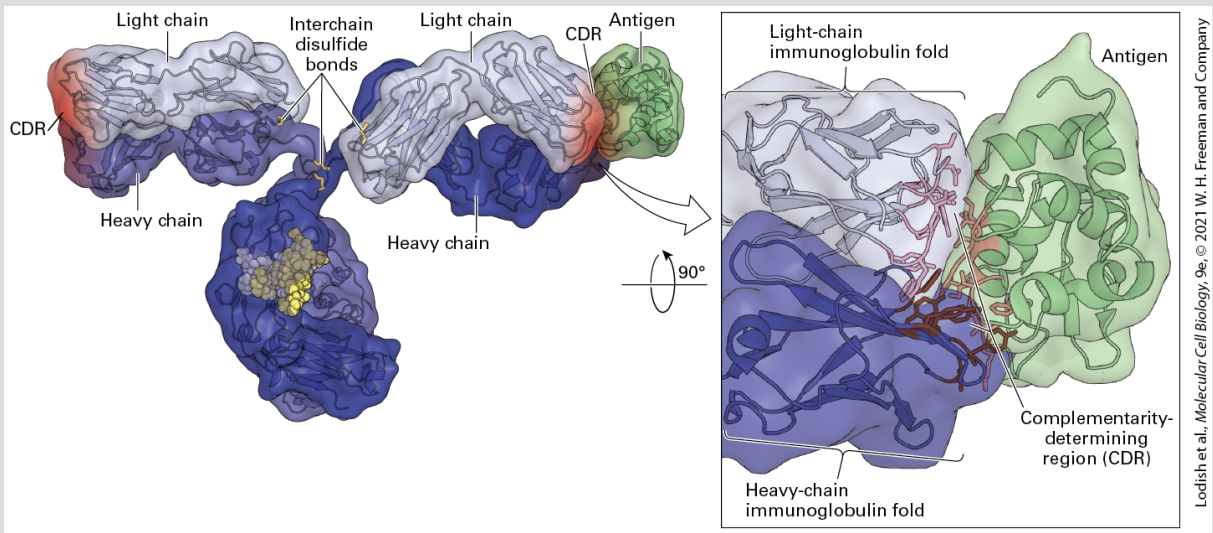


FIGURE 24-14 Immunoglobulin structure. This model shows the three-dimensional structure of an immunoglobulin molecule complexed with hen egg-white lysozyme (a protein antigen) as determined by x-ray crystallography.

[Data from E. A. Padlan et al., 1989, *Proc. Nat'l Acad. Sci. USA* **86**:5938, PDB ID 1igt, 3hfm.]

The difficulty of encoding directly in the inherited genome (germ line) all the information necessary to generate the enormously diverse antibody repertoire (more than a million different antibody molecules in what we now know is a genome encoding about 20,000 independent genes) led to suggestions of unique genetic mechanisms to account for this diversity. Given the size of a typical antibody heavy chain and light chain (each heavy chain–light chain combination, if encoded as such, would require 2.5–3.5 kb of DNA, depending on the isotype), it is immediately obvious that encoding a set of antibody molecules of sufficient diversity to provide adequate protection against the wide array of pathogens and other foreign substances to which an organism is exposed would rapidly exhaust its DNA-coding capacity. We shall see that, indeed, unique mechanisms are at work to create an adequately diverse set of antibodies.

Immunoglobulin Domains Have a Characteristic Fold Composed of Two β Sheets Stabilized by a Disulfide Bond

Both the variable and constant domains of immunoglobulins fold into a compact, three-dimensional structure, an immunoglobulin (Ig) domain, composed exclusively of β sheets (see Figure 24-13b). A typical Ig domain contains two β sheets (one with three strands and one with four strands) held together like a sandwich by a disulfide

bond. The residues that point inward are mostly hydrophobic and help stabilize this sandwich structure. The residues exposed to the aqueous environment show a greater frequency of polar and charged side chains. The spacing of the cysteine residues that make up the disulfide bond and a small number of strongly conserved residues characterize this evolutionarily ancient structural motif, termed the **immunoglobulin fold**. The basic immunoglobulin fold is also found in numerous eukaryotic proteins that are not directly involved in antigen-specific recognition, including the Ig superfamily of cell-adhesion molecules, or Ig CAMs (see Chapter 20).

The region on an antigen that makes contact with the corresponding antibody is called an **epitope**; the portion of the immunoglobulin that makes contact with antigen is the **paratope**. A protein antigen usually contains multiple epitopes, which are often exposed loops or surfaces on the protein and are thus able to bind to different antibody molecules. Each homogeneous antibody preparation derived from a clonal population of B cells recognizes a single molecularly defined epitope on the corresponding antigen.

To determine the structure of an antibody complexed to its cognate epitope on an antigen, it is important to have a source of homogeneous immunoglobulin and of antigen in pure form (see Chapter 3). As we have seen, homogeneous immunoglobulins can be obtained from B-cell tumors, but in that case, the antigen for which the antibody is specific is usually not known. The breakthrough for

generating homogeneous antibody preparations suitable for structural analysis was the development of techniques to obtain antibodies from hybridomas by use of a special selection medium. The creation of immortalized cell lines that produce antibodies of defined specificity, called **monoclonal antibodies**, has yielded essential tools for the cell biologist. Monoclonal antibodies are widely used not only for the specific detection of macromolecules, but also for detection and quantitation of drugs, drug metabolites, and even signaling molecules such as cAMP. Monoclonal antibodies can detect proteins and their modifications (e.g., phosphorylation, nitrosylation, methylation, acetylation, etc.) as well as complex carbohydrates, (glyco)lipids, and nucleic acids and their modifications. They have therefore found widespread use in the laboratory as well as for diagnostic and therapeutic purposes.

We now have detailed insights into the structure of a large number of monoclonal antibodies, each in a complex with its specific antigen. There are no hard-and-fast rules that describe these interactions, other than the usual rules of molecular complementarity of proteins with other molecules and macromolecules (see Chapter 3). The CDRs make the most important contributions to the antigen-antibody interface. The CDR3 of the V_H region of the Ig heavy chain plays a particularly prominent role, as does the CDR3 of the V_L region of the Ig light chain.

An Immunoglobulin's Constant Region Determines Its Functional Properties

As we have seen, antibodies recognize antigen via their variable regions. Their constant regions determine which effector molecules they recruit to neutralize the pathogen.

Antibodies attached to a viral or microbial surface can be recognized directly by cells that express receptors specific for the Fc portion of immunoglobulins. These **Fc receptors (FcRs)**, which are specific for individual classes and subclasses of immunoglobulins, display considerable structural and functional heterogeneity. By means of FcR-dependent events, specialized phagocytic cells such as dendritic cells and macrophages can bind to antibody-decorated particles or cells and then ingest and destroy them. The decoration of an antigenic target with bound antibodies, or its covalent modification with complement components, is called **opsonization**. FcR-dependent events also allow some immune-system cells (e.g., monocytes and natural killer cells) to directly engage target cells that display viral or other antigens to which antibodies are attached. This engagement may induce the immune-system cells to release toxic small molecules (e.g., reactive oxygen species) or the contents of cytotoxic granules, including perforins and granzymes. *Perforins* are proteins that can attach themselves to the surface of the engaged target cell and form pores in its membrane. These newly formed pores allow access by *granzymes*, proteases that initiate a sequence

of events that will ultimately kill the target cell (see Figure 24-6). This process, called *antibody-dependent cell-mediated cytotoxicity (ADCC)*, illustrates how cells of the innate immune system interact with, and benefit from, the products of the adaptive immune response.

Antigen-antibody (immune) complexes of some immunoglobulin isotypes can initiate the classical pathway of complement activation (see Figure 24-5). IgM and IgG3 are particularly good at complement activation, but all IgG classes can in principle activate complement, whereas IgA and IgE are unable to do so. The large amounts of IgA found in the gut contribute to its barrier function by neutralizing gut-resident microbes.

KEY CONCEPTS OF SECTION 24.2

Immunoglobulins: Structure and Function

- Most immunoglobulins (antibodies) are composed of two identical heavy (H) chains and two identical light (L) chains (H_2L_2). Each chain contains a variable (V) region and a constant (C) region. Proteolytic fragmentation yields monovalent F(ab) and bivalent $F(ab')_2$ fragments, which contain variable-region domains that retain antigen-binding capability (see Figure 24-9). The Fc fragment contains constant-region domains and determines their ability to activate complement components or bind to receptors specific for Fc regions expressed on leukocytes.
- Immunoglobulins are divided into classes based on the constant regions of their heavy chains (see Figure 24-10). In mammals, there are five major classes: IgM, IgD, IgG, IgA, and IgE; the corresponding heavy chains are referred to as *μ, δ, γ, α* and *ε*. There are two major classes of light chains, κ and λ, again characterized by the attributes of their constant regions.

- IgM and IgA can form higher order structures: IgM forms pentamers (five identical H_2L_2 copies); and IgA forms dimers (two identical H_2L_2 copies).
- Each individual B lymphocyte expresses an immunoglobulin of unique sequence and is therefore uniquely specific for a particular antigen. Upon recognition of antigen, only a B lymphocyte that bears a receptor specific for it will be activated and expand clonally (clonal selection) (see Figure 24-12).
- The antigen specificity of antibodies is conferred by their variable regions, which contain regions of high sequence variability, called *hypervariable* or *complementarity-determining regions* (see Figure 24-13a). These hypervariable regions are positioned at the tip of the variable region, where they can make specific contacts with the antigen for which a particular antibody is specific.
- The repeating immunoglobulin domains that make up immunoglobulin molecules have a characteristic three-dimensional structure, the immunoglobulin fold, which consists of two β sheets held together in a sandwich by a disulfide bond (see Figure 24-13b).
- The constant regions of the heavy chains endow antibodies with unique effector functions, such as the capacity to bind complement, the ability to be transported across epithelia, or the ability to interact with receptors specific for the Fc portion of immunoglobulins.

24.3 Generation of Antibody Diversity and B-Cell Development

Pathogens have short replication times, are quite diverse in their genetic makeup, and evolve quickly by mutation, generating enormous antigenic variation. An adequate defense must therefore be capable of mounting an equally diverse response. Antibodies provide the diversity required for successful host defense. The timing of the antibody response and its necessary adjustment to changes in the antigenic makeup of the pathogen pose unique demands on the organization and regulation of the adaptive immune system. A unique mechanism has evolved that allows not only virtually limitless variation in the set of antibodies that can be produced (called the *repertoire*), but also rapid improvement in the binding affinity of those antibodies to meet the demands posed by an ongoing viral or bacterial infection. Because optimal antibody production by B cells requires assistance from T cells, we will see below that the molecular mechanisms underlying lymphocyte diversity are fundamentally similar for B and T cells.

B cells, which are responsible for antibody production, make use of a unique mechanism by which the genetic information required for synthesis of immunoglobulin heavy and light chains is stitched together from separate DNA sequences, or Ig gene segments, to create a functional transcription unit. The recombination mechanism that combines Ig gene segments itself dramatically

expands the variability in sequence precisely at those sites where these genetic elements are joined together. This mechanism for generating a diverse array of antibodies is fundamentally different from meiotic recombination, which occurs only in germ cells, and from alternative splicing of exons (see Chapter 7). Because this recombination mechanism occurs in somatic cells but not in germ cells, it is known as *somatic gene rearrangement* or *somatic recombination*. This unusual recombination mechanism, unique to antigen-specific receptors on B and T lymphocytes, makes it possible to specify an enormously diverse set of receptors with minimal expenditure of DNA coding space. The discovery of somatic recombination is detailed in Classic Experiment 24.1.

The ability to combine discrete genetic elements at will (combinatorial diversity), in addition to the generation of yet more sequence diversity in the encoded receptors by the underlying recombination mechanisms themselves, allows adaptive immune responses against a virtually limitless array of antigens, including molecules encoded by the host and chemical structures not naturally found in nature. Thus there are mechanisms at work not only to create this enormous diversity, but also to impose tolerance to curtail unwanted reactivity against self components. Neither mechanism is perfect: the adaptive immune system cannot generate receptors for all foreign substances. Furthermore, the unavoidable price we pay for how we generate B- and T-cell receptors is the

likelihood of self-reactive receptors, which are implicated in **autoimmunity**.

A Functional Light-Chain Gene Requires Assembly of V and J Gene Segments

Genes encoding intact immunoglobulins do not exist already assembled in the genome, ready for expression. Instead, the required gene segments are brought together and assembled in the course of B-cell development. The organization of the region of the genome containing the immunoglobulin genes is shown in Figure 24-15. In B cells, the DNA in this region is rearranged to generate assembled and fully functional immunoglobulin-encoding genes in each B cell and its descendants. Although in the course of B-cell development the rearrangement of heavy-chain genes occurs before the rearrangement of light-chain genes, we discuss light-chain genes first, because of their less complex organization.

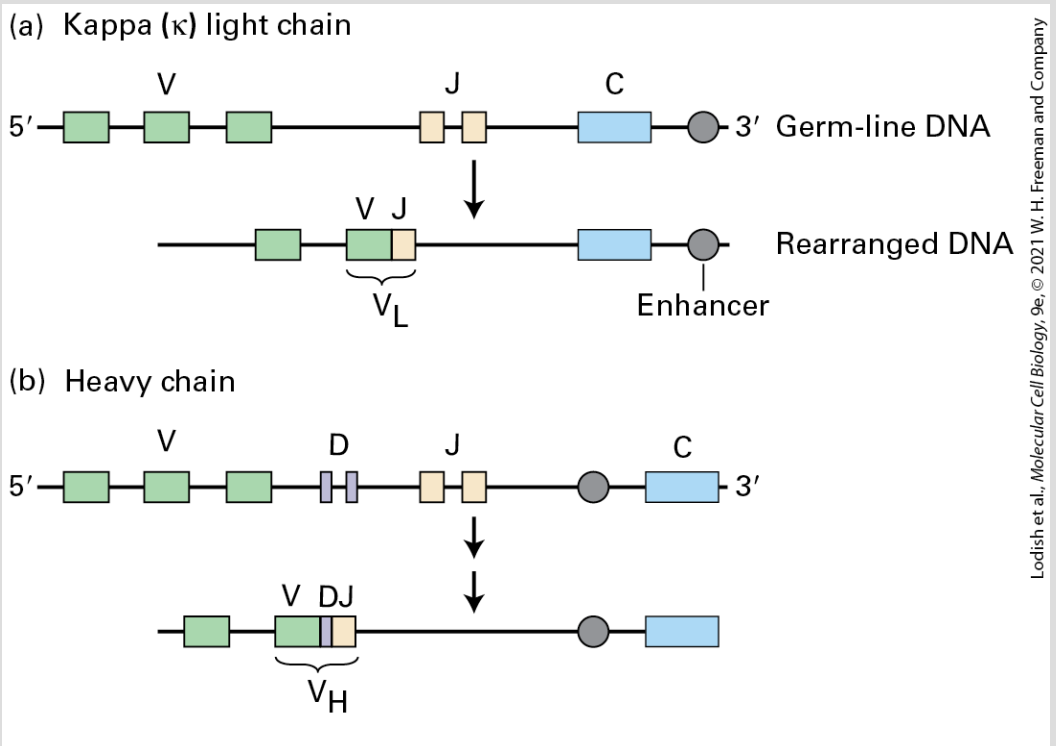


FIGURE 24-15 Overview of somatic gene rearrangement in immunoglobulin DNA. The stem cells that give rise to B cells contain multiple gene segments encoding portions of immunoglobulin heavy and light chains. During development of a B cell, somatic recombination of these gene segments yields functional light-chain genes (a) and heavy-chain genes (b). Each V gene segment carries its own promoter. Rearrangement brings an enhancer linked to or contained within the immunoglobulin locus close enough to the combined sequence to activate transcription. The light-chain variable region (V_L) is encoded by two joined gene segments, and the heavy-chain variable region (V_H) is encoded by three joined segments. Note that the chromosomal regions encoding immunoglobulins contain many more V, D, and J segments than those shown. In addition, the κ light-chain locus contains a single constant (C) segment, as shown, but the heavy-chain locus contains several distinct C segments (not shown) that encode the different immunoglobulin isotypes, as we detail later.

The immunoglobulin light-chain genes consist of clusters of V gene segments, followed downstream by a single C segment. Each V gene

segment carries its own promoter sequence and encodes the bulk of the light-chain variable region, although a small piece of the nucleotide sequence encoding the light-chain variable region is missing from the V gene segment. This missing portion is provided by one of the multiple J segments located between the V segments and the single C segment in the unrearranged (or germ-line) κ light-chain locus (Figure 24-15a). (This J segment is a genetic element, not to be confused with the J chain, a polypeptide subunit of the pentameric IgM molecule and found also in association with IgA; see Figure 24-10.) In the course of B-cell development, commitment of a B-cell precursor to use a particular V gene segment — a random process — results in its physical juxtaposition with one of the J segments, again a random choice, to form an exon (V_L) encoding the entire light-chain variable region (Figure 24-15a). This DNA rearrangement not only generates an intact and functional light-chain gene, but also places the promoter sequence of the rearranged gene within controlling distance of enhancer elements, located near the light-chain constant-region exon, that are required for its transcription. Only a fully rearranged light-chain gene is transcribed and subsequently translated into protein.

Recombination Signal Sequences

Detailed DNA sequence analysis of the light-chain and heavy-chain regions revealed a conserved sequence element at the 3′ end of each V gene segment. This conserved element, called a *recombination*

signal sequence (RSS), is composed of heptamer and nonamer sequences separated by a 23-bp spacer. At the 5′ end of each J segment, there is a similarly conserved RSS that contains a 12-bp spacer (Figure 24-16). The 12- and 23-bp spacers separate the conserved heptamer and nonamer sequences by one and two turns of the DNA helix, respectively.

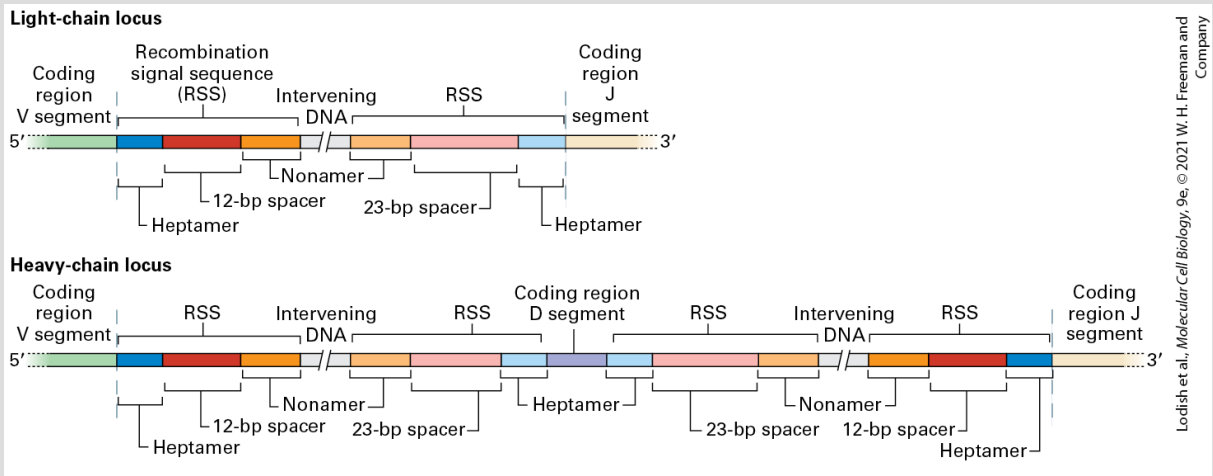
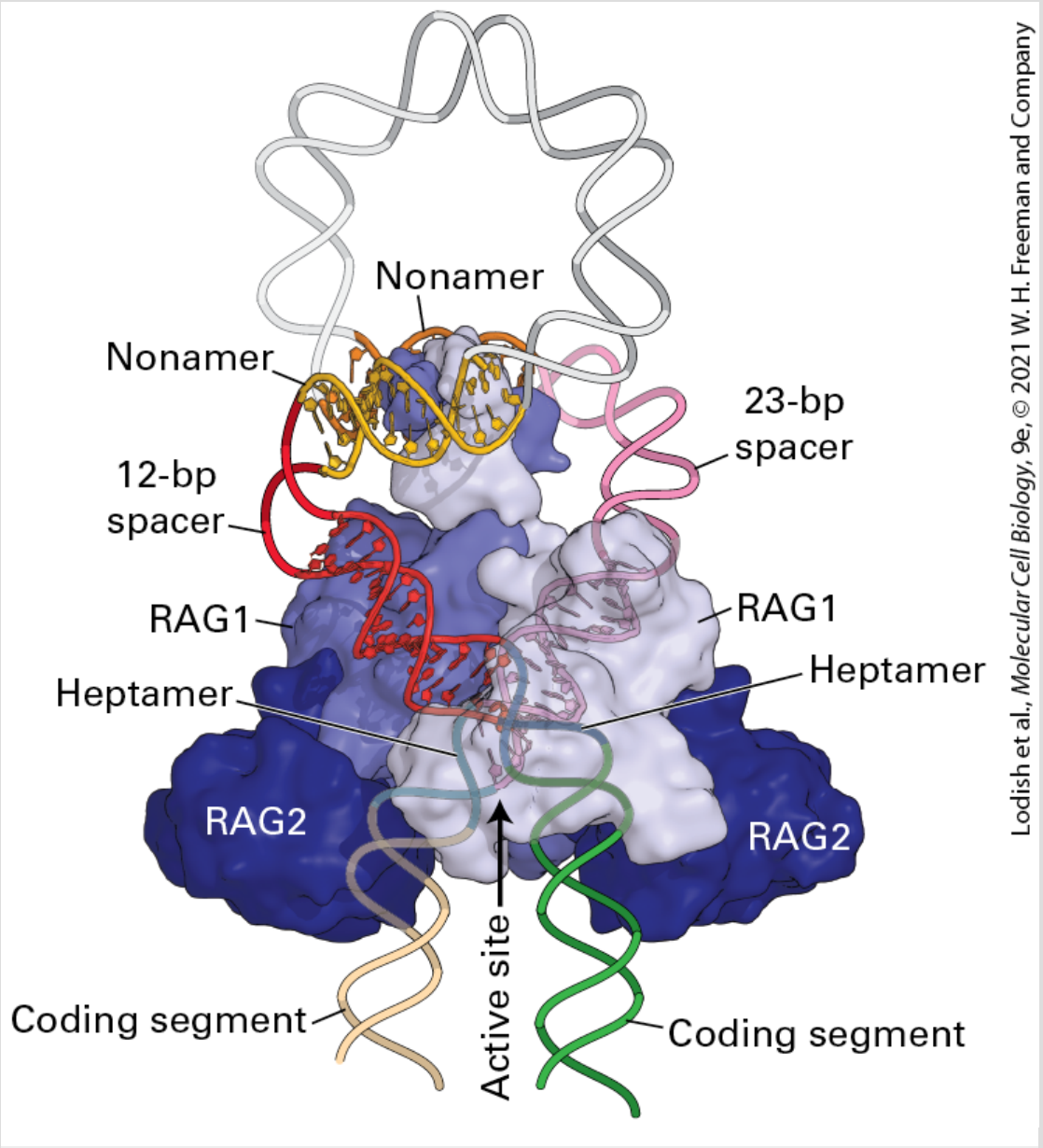


FIGURE 24-16 DNA elements involved in V(D)J recombination in the immunoglobulin loci. Schematic representation of the immunoglobulin light-chain (*top*) and heavy-chain locus (*bottom*). At its 3′ end, each immunoglobulin V region coding sequence abuts a recombination signal, composed of a conserved heptamer (7 bp) and a conserved nonamer (9 bp) sequence, separated by a 12-bp spacer of variable sequence. The combination of heptamer, nonamer, and spacer is called a *recombination signal sequence (RSS)*. Recombination requires juxtaposition of the gene segment that recombines with a second gene segment with an RSS, which carries a heptamer and nonamer sequence that is complementary and antiparallel to the first RSS. For the second RSS, the spacer is 23-bp long, and of variable sequence. For the light-chain locus, the RSSs that flank the V and J segments are separated by intervening DNA. Recombination can occur only between elements that carry RSSs of different spacer length (the “12-23” rule). The heavy-chain locus comprises three sequence elements that can rearrange. The V region of the heavy chain carries an RSS at its 3′ end, the D segments are flanked on their 5′ and 3′ ends by an RSS, and the J segments carry an RSS at their 5′ end. The organization of the immunoglobulin heavy-

chain locus thus enables rearrangements of D to J, and of V to DJ, but direct rearrangements of V to J are not allowed, as this would violate the 12-23 rule.

Somatic recombination is catalyzed by two enzymes, the RAG1 and RAG2 recombinases, which are expressed only in lymphocytes. Thus these rearrangements do not occur in any other cells of the body. Juxtaposition of the two gene segments to be joined is stabilized by the RAG1/RAG2 complex (Figure 24-17). The structure of the RAG1/RAG2 complex juxtaposes the RSSs so that proper covalent joining of the immunoglobulin gene segments can occur, with concomitant release of the intervening DNA, which is released as a covalently closed, circular DNA molecule.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 24-17 RAG1/RAG2 structure. The RAG1 and RAG2 proteins (light and dark purple) are responsible for bringing together the immunoglobulin gene segments for rearrangement. RAG1/RAG2 complexes with the recombination signal sequences, positioning the 12- and 23-bp spacer sequences to enable cleavage at the boundary of the coding sequence and the heptamer of the RSS.

[Data from M.-S. Kim et al., 2015, *Nature* **518**:507–511, PDB ID 4wwx; and F. F. Yin et al., 2009, *Nat. Struct. Mol. Biol.* **16**:499–508, PDB ID 3gna.]

Defects in the synthesis of RAG proteins obliterate the possibility of somatic gene rearrangements. As described below, the rearrangement process is essential for B-cell development; consequently, RAG deficiency leads to the complete absence of B cells. People with defects in RAG gene function suffer from severe immunodeficiency. Targeted deletion of RAG genes in mice likewise leads to a complete defect in immunoglobulin (and T-cell receptor) gene rearrangement, resulting in a developmental block in the generation of B and T lymphocytes. ■

Junctional Imprecision

In addition to the random selection of V and J gene segments, processing of the intermediates created in the course of somatic recombination provides an additional means for expanding the variability of immunoglobulin sequences. This additional variability is created at the junction of the segments to be joined by random insertion or deletion of nucleotides to one of the cut ends before the V and J segments are joined, a process termed *junctional imprecision*. Whenever a V and a J segment recombine, the sequence and reading frame of the VJ product cannot be predicted; only one in three recombination reactions results in a reading frame that is compatible with light-chain synthesis.

Light-chain diversity therefore arises not only from the combinatorial use of V and J gene segments, but also from

junctional imprecision. Inspection of the three-dimensional structure of the light chain shows that the highly diverse joint generated as a consequence of junctional imprecision forms part of a loop — hypervariable region 3 (HV3) — that projects into the antigen-binding site and makes contact with antigen (see Figure 24-13b).

Rearrangement of the Heavy-Chain Locus Involves V, D, and J Gene Segments

The organization of the heavy-chain locus is more complex than that of the κ light-chain locus. The heavy-chain locus contains not only a large tandem array of V segments (each equipped with its own promoter) and multiple J segments, but also multiple D (diversity) segments (see Figure 24-15b). Somatic recombination of a V, a D, and a J segment generates a rearranged sequence encoding the heavy-chain variable region (V_H).

At the 3′ end of each V segment in heavy-chain DNA, there are conserved heptamer and nonamer sequences separated by spacer DNA, similar to the RSSs in light-chain DNA. These RSSs are also found in complementary and antiparallel configuration at the 5′ end and the 3′ end of each D segment (Figure 24-16). The J segments are similarly equipped at their 5′ end with the requisite RSS. The spacer

lengths in these RSSs are such that D segments can join to J segments, and V segments to already rearranged DJ segments. However, neither direct V-to-J nor D-to-D joining is allowed, in compliance with the 12-23-bp spacer rule. Heavy-chain rearrangements proceed via the same mechanisms described above for light-chain rearrangements.

In the course of B-cell development, the heavy-chain locus is always rearranged first, starting with D-J rearrangement. D-J rearrangement is followed by V-D-J rearrangement. In the course of the D-J and V-D-J rearrangements, an enzyme called *terminal deoxynucleotidyl transferase (TdT)* may add nucleotides to free 3′ OH ends of DNA in a template-independent fashion. Up to a dozen or so nucleotides, called the *N-region* or *N-nucleotides*, may be added, generating additional sequence diversity at the junctions whenever D-J and V-D-J rearrangements occur (Figure 24-18, step **7**). As with rearrangements forming the light-chain gene, only one in three rearrangements yields the proper reading frame for the rearranged VDJ sequence. If the rearrangement yields a sequence encoding a functional protein, it is called *productive*. Although the heavy-chain locus is present on each of two homologous chromosomes, only one productive rearrangement is permitted, as we will see below.

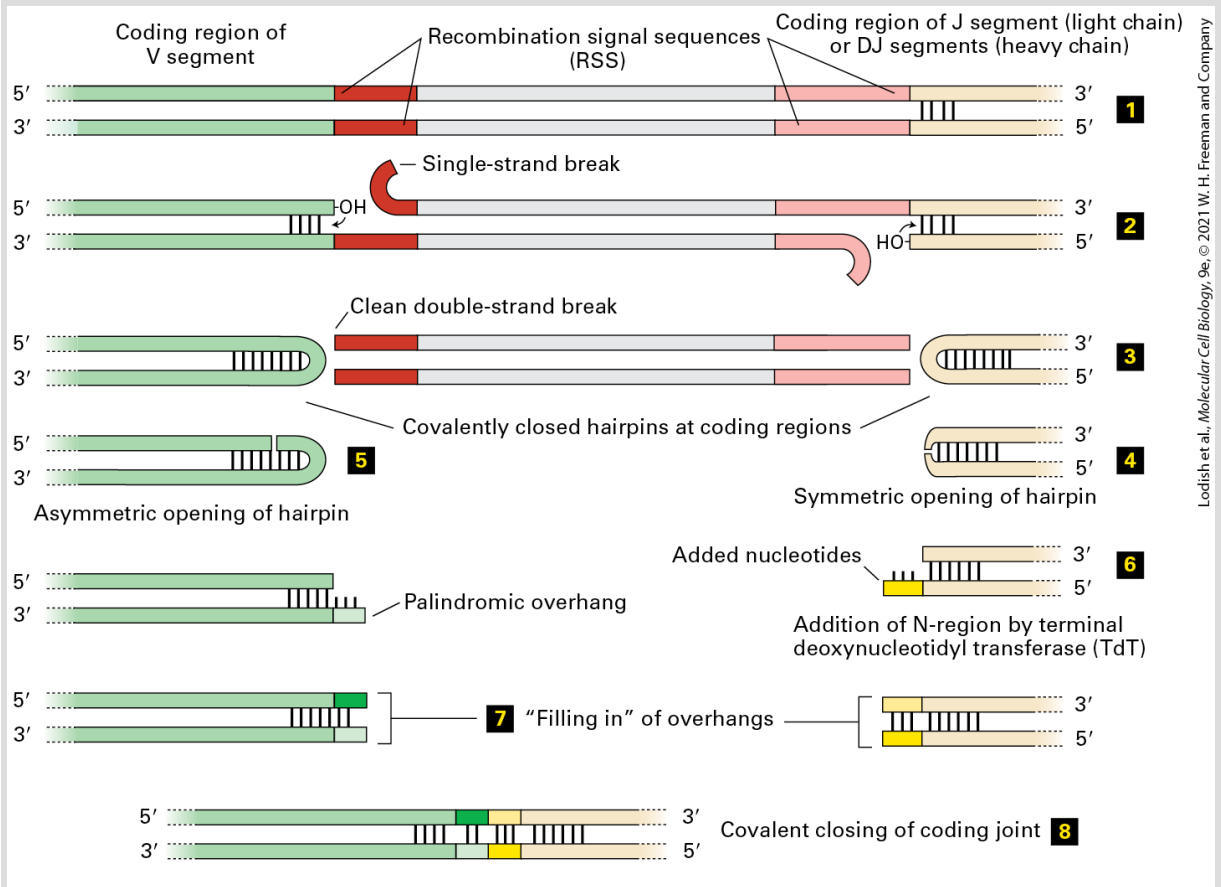


FIGURE 24-18 The sequential steps in V(D)J recombination. Step 1: Linear depiction of the immunoglobulin gene segments to be recombined. Each immunoglobulin gene segment carries a recombination signal sequence. For the V regions, at their $3'$ end, for the J segments (light chain) at their $5'$ end and for a rearranged DJ segment (heavy chain) also at the $5'$ end. Step 2: The RAG1/RAG2 recombinase makes a single strand cut to generate a free $3'$ OH hydroxyl group. Step 3: The $3'$ OH group attacks the bottom strand to create a covalently closed hairpin. The intervening DNA, with an RSS at either end, is excised precisely at the boundaries of the RSS and yields a covalently closed "excision circle" (circularization step not shown in the figure). The gene segments that participate in rearrangement are held in close proximity by the RAG1/RAG2 recombinase and associated proteins. Step 4: The covalently closed hairpins at the end of the coding sequences can be opened either symmetrically (*left*) or (step 5) asymmetrically (*right*). If the hairpin is opened asymmetrically (*left*), the unpaired nucleotides may be removed by nucleases to produce a "flush" end (not shown) or result in a short palindromic sequence that can be rendered flush by "filling in" the missing nucleotides (step 7); this adds nucleotides to where the actual junction will occur, thus creating additional variability in the final product of rearrangement. These nucleotides are called *P nucleotides* because of their palindromic

nature. Even if the hairpin opens symmetrically (step 4, *right*), the enzyme terminal deoxynucleotidyl transferase (TdT) can add random nucleotides to the free 3′ OH in a template-independent manner (step 6), thus increasing sequence diversity at the joint. The target of TdT activity is confined to V, D, and J elements of the immunoglobulin heavy-chain locus. Once the ends of the coding sequences have been rendered flush (step 7), the coding ends are ligated together by nonhomologous end joining.

An enhancer located downstream of the cluster of J segments and upstream of the constant-region segment activates transcription from the promoter at the 5′ end of the rearranged VDJ sequence (see Figure 24-15). Splicing of the primary transcript produced from the rearranged heavy-chain gene generates a functional mRNA encoding the μ heavy chain. For both heavy-chain and light-chain genes, somatic recombination places the promoters upstream of the V segments within functional reach of the enhancers necessary to allow transcription, so that only rearranged VJ and VDJ sequences, and not the V segments that remain in the germ-line configuration, are transcribed.

Somatic Hypermutation Allows the Generation and Selection of Antibodies with Improved Affinities

In addition to somatic recombination and junctional imprecision, antigen-activated B cells can undergo an additional diversity-generating process called *somatic hypermutation*. Upon exposure to

antigen and receipt of the proper signals, most of which are provided by T cells, expression of activation-induced deaminase (AID) is turned on. This enzyme deaminates cytosine residues, converting them to uracil. When a B cell that carries this lesion replicates its DNA, it may place an adenine on the complementary strand, thus generating a G-to-A transition (see Figure 5-14). Alternatively, the uracil may be excised by DNA glycosylase to yield an abasic site. Such abasic sites, when copied, give rise to an A, G, or C in the complementary DNA strand. That is, a C:G base pair can be replaced by somatic hypermutation with a T:A, A:T, or G:C base pair. Mutations thus accumulate with every successive round of B-cell division, yielding numerous mutations in the rearranged VJ and VDJ segments. Error-prone filling by DNA polymerase of gaps created by nucleotide excision repair also contributes to somatic hypermutation.

The process of somatic hypermutation occurs when lymphocytes reside in specialized microanatomic structures known as *germinal centers*. These structures, which arise within the follicles of secondary lymphoid organs upon immunization, consist of foci of thousands of rapidly proliferating and hypermutating B cells. In addition to B cells, germinal centers contain follicular dendritic cells, a cell type that serves as a depot for antigen that can be retrieved by B cells, and a small number of helper T cells specialized in providing selective signals that control B cells. Many of the somatic mutations induced by AID are deleterious, in that they

reduce the affinity of the encoded antibody for an antigen, but some improve the encoded antibody's affinity for an antigen. In a process analogous to Darwinian evolution, B cells carrying surface immunoglobulins with affinity-increasing mutations have a selective advantage in picking up antigen from follicular dendritic cells, which allows them to successfully compete for signals from the limiting number of helper T cells residing in the germinal center (described in some detail in Section 24.6). These signals trigger the proliferation, and thus the clonal selection, of B cells that have on their surface antibodies with higher affinity for the target antigen. This triggers further proliferation and additional mutations and differentiation into antibody-secreting plasma cells or memory B cells. The net result is the generation of a B-cell population whose antibodies, as a rule, show a higher affinity for the antigen.

In the course of an immune response, or upon repeated immunization, the adaptive immune response therefore exhibits *affinity maturation* — an increase in the average affinity of antibodies for an antigen as a function of time after antigen exposure — as the result of somatic hypermutation and selection. Antibodies produced following this phase of the adaptive immune response display affinities for antigen in the nanomolar (or better) range. For reasons that are not fully understood, but likely related to the chromatin structure of the rearranged immunoglobulin loci, the activity of AID is focused mainly on rearranged VJ and VDJ segments, and this targeting may therefore require active transcription. The entire

process of somatic hypermutation is strictly antigen-dependent and shows an absolute requirement for interactions between the B cells and certain T-cell types.

B-Cell Development Requires Input from a Pre-B-Cell Receptor

B cells destined to make immunoglobulins must rearrange the necessary gene segments to assemble functional heavy-chain and light-chain genes. These rearrangements occur in a carefully ordered sequence during the development of a B cell, starting with heavy-chain rearrangements. Moreover, as noted above, the rearranged heavy-chain contains, near its C-terminus, a membrane-spanning segment that traverses the plasma membrane (Figure 24-19; also see Figure 24-20). This membrane-bound receptor first executes a cell-fate decision necessary to drive further B-cell development (and antibody synthesis) by permitting subsequent rearrangement of the light-chain genes. Only a productive rearrangement that yields an in-frame VDJ combination can generate a complete μ heavy chain. The production of that μ chain serves as a signal to the B cell that it has successfully accomplished rearrangement and that no further rearrangements of the heavy-chain locus on the remaining gene copy are required. Recall that each lymphocyte precursor starts out with two immunoglobulin locus–bearing, homologous chromosomes in the germ-line

(unrearranged) configuration. In accordance with clonal selection theory, which stipulates that each lymphocyte ought to come equipped with a single antigen-specific receptor, continued rearrangement would entail the risk of producing B cells with two different heavy chains, each with different specificity — an undesirable outcome.

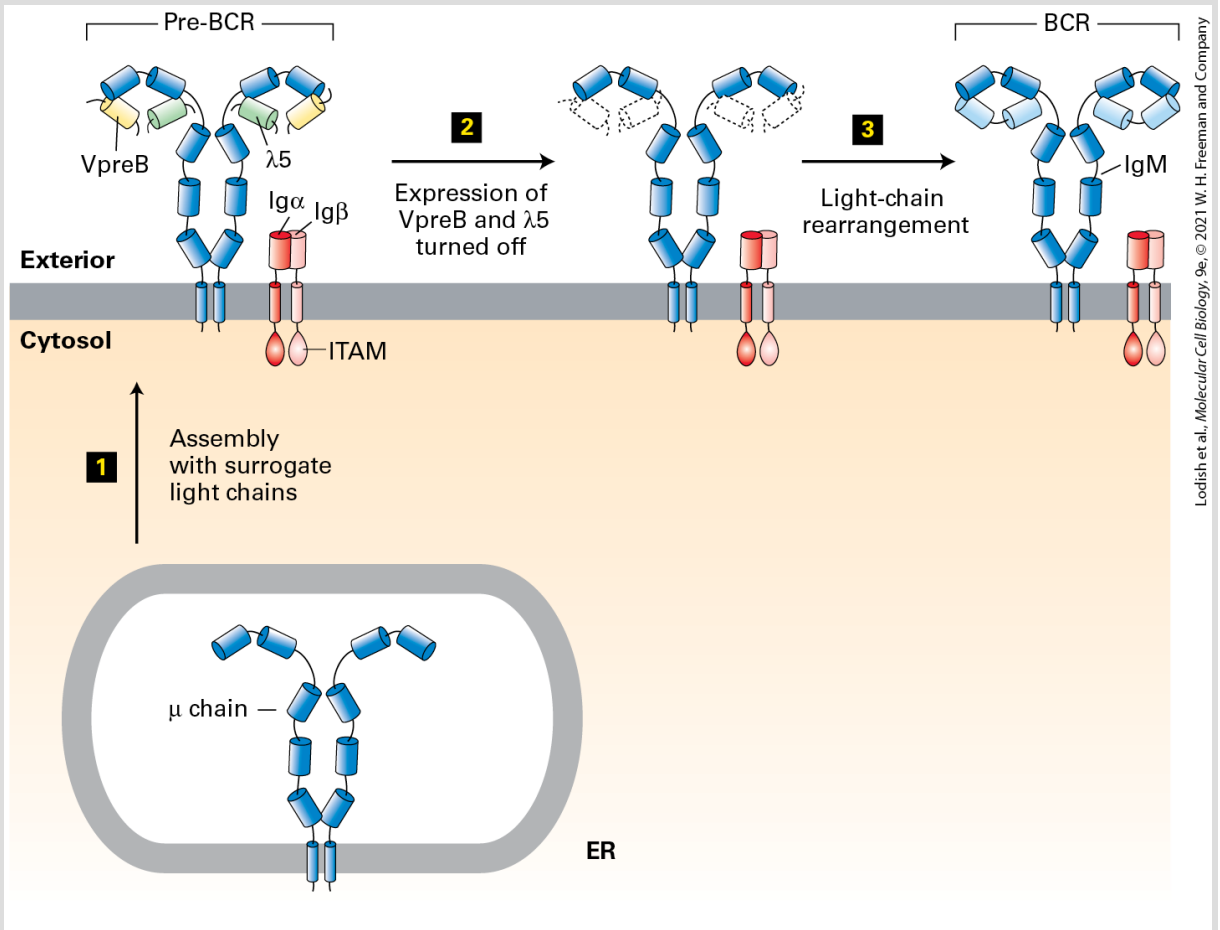


FIGURE 24-19 Structure of the pre-B-cell receptor and its role in B-cell development. Successful rearrangement of V, D, and J heavy-chain gene segments allows synthesis of membrane-bound μ heavy chains in the endoplasmic reticulum (ER) of a pre-B cell. At this stage, no light-chain gene rearrangement has occurred. Newly made μ chains assemble with surrogate light chains, composed of λ5 and VpreB, and Igα/Igβ to yield the pre-B-cell receptor, pre-BCR (step 1). This receptor drives proliferation of those B cells that carry it. It

also suppresses rearrangement of the heavy-chain locus on the other chromosome and so mediates allelic exclusion. In the course of proliferation, the synthesis of λ5 and VpreB is shut off (step 2), resulting in "dilution" of the available surrogate light chains and reduced expression of the pre-BCR. As a result, rearrangement of the light-chain loci can proceed (step 3). If this rearrangement is productive, the B cell can synthesize light chains and complete assembly of the B-cell receptor (BCR), which consists of a membrane-bound IgM and associated Igα and Igβ. The B cell is now responsive to antigen-specific stimulation.

Successful rearrangement of V, D, and J segments in the heavy-chain locus thus allows the synthesis of a complete μ chain. B cells at this stage of development are called *pre-B cells*, as they have not yet completed assembly of a functional light-chain gene and therefore cannot engage in antigen recognition. As with other membrane-spanning proteins (see Chapter 13), the μ chain is synthesized in the endoplasmic reticulum. It becomes part of a membrane-bound signaling receptor whose expression is essential for B-cell development to proceed in an orderly fashion.

In pre-B cells, newly made μ chains form a complex with two so-called surrogate light chains, λ5 and VpreB (see Figure 24-19). The μ chain itself possesses a very short (three amino acids) cytosol-facing C-terminal tail and is therefore incapable of recruiting cytoplasmic components for the purpose of signal transduction. Instead, pre-B cells express two auxiliary transmembrane proteins, called *Igα* and *Igβ*, each of which carries in its cytoplasmic tail an *i*mmunoreceptor *t*yrosine-based *a*ctivation *m*otif, or ITAM. The entire complex, including μ chain, λ5, VpreB, Igα, and Igβ, constitutes the *pre-B-cell*

receptor (pre-BCR). Engagement of this receptor by (unknown) suitable signals results in recruitment and activation of a Src-family tyrosine kinase, which phosphorylates tyrosine residues in the ITAMs. In their phosphorylated form, ITAMs recruit other molecules essential for signal transduction (see below). Because no functional light chains are yet part of this receptor, it is presumed to be incapable of antigen recognition, as the antigen-binding site has contributions from both the heavy and light chains (see Figure 24-14) (except in camelids, which produce only H-chain antibodies).

The pre-B-cell receptor has several important functions. First, it shuts off expression of the RAG recombinases, so that rearrangement of the other (allelic) heavy-chain locus cannot proceed. This phenomenon, called *allelic exclusion*, ensures that only one of the two available copies of the heavy-chain locus will be rearranged and thus expressed as a complete μ chain. Second, because of the association of the pre-B-cell receptor with Igα and Igβ, the receptor becomes a functional signal-transduction unit. Signals that emanate from the pre-BCR initiate proliferation of the pre-B cell to expand the numbers of those B cells that have undergone productive D-J and V-D-J recombination.

In the course of this expansion, expression of the surrogate light chains, VpreB and λ5, subsides. The progressive dilution of VpreB and λ5 with every successive cell division results in insufficient fully assembled pre-BCR in the endoplasmic reticulum. As a

consequence, the heavy chains are degraded (see Chapters 13 and 14) and the amount of pre-BCR signaling decreases. This reduction in signaling allows re-initiation of expression of the RAG recombinases, which now target the κ or λ light-chain locus. A productive light-chain V-J rearrangement also shuts off rearrangement of the allelic locus (allelic exclusion). Upon completion of a successful V-J light-chain rearrangement, the B cell can make both μ heavy chains and κ or λ light chains and assemble them into a functional **B-cell receptor (BCR)**, which can recognize antigen (see Figure 24-19).

Once a B cell expresses a complete BCR on its cell surface, it can recognize antigen, and all subsequent steps in B-cell activation and differentiation require engagement with the antigen for which that BCR is specific. The BCR not only plays a role in driving B-cell proliferation upon a successful encounter with antigen, but also functions as a device for receptor-mediated endocytosis, an essential step that allows the B cell to process the acquired antigen and convert it into a signal that sends out a call for assistance by T lymphocytes; this antigen-presentation function of B cells is described in later sections.

During an Adaptive Response, B Cells Switch from Making Membrane-Bound Ig to Making Secreted Ig

As just described, the B-cell receptor, a membrane-bound IgM, provides a B cell with the ability to recognize a particular antigen, an event that triggers clonal selection and proliferation of that B cell, thus increasing the number of B cells specific for that antigen (see Figure 24-12). However, key functions of immunoglobulins, such as neutralization of antigens or killing of bacteria, require that those products be released by the B cell so that they can accumulate in the extracellular environment and act at a distance from the site where they were produced.

Whether to synthesize membrane-bound or secreted immunoglobulin is a choice made by the B cell during processing of the heavy-chain primary transcript. As shown in Figure 24-20, the μ locus contains two exons (TM1 and TM2) that together encode a C-terminal domain that spans the plasma membrane and anchors the IgM in it. One polyadenylation site is found upstream of these exons; a second polyadenylation site is present downstream. If the downstream poly(A) site is chosen, then further processing yields an mRNA that encodes the membrane-bound form of μ. (As described above, this choice is necessary for formation of the B-cell receptor, which includes membrane-bound IgM.) If the upstream poly(A) site is chosen, processing yields the secreted version of the μ chain. Similar arrangements are found for the other Ig constant-region gene segments (γ, α, ε), each of which can specify either a membrane-bound or a secreted heavy chain. The ability to switch between the membrane-anchored and the secreted form of immunoglobulin heavy chains by alternative use of polyadenylation

sites (*not* by alternative splicing) is so far unique to this family of gene products.

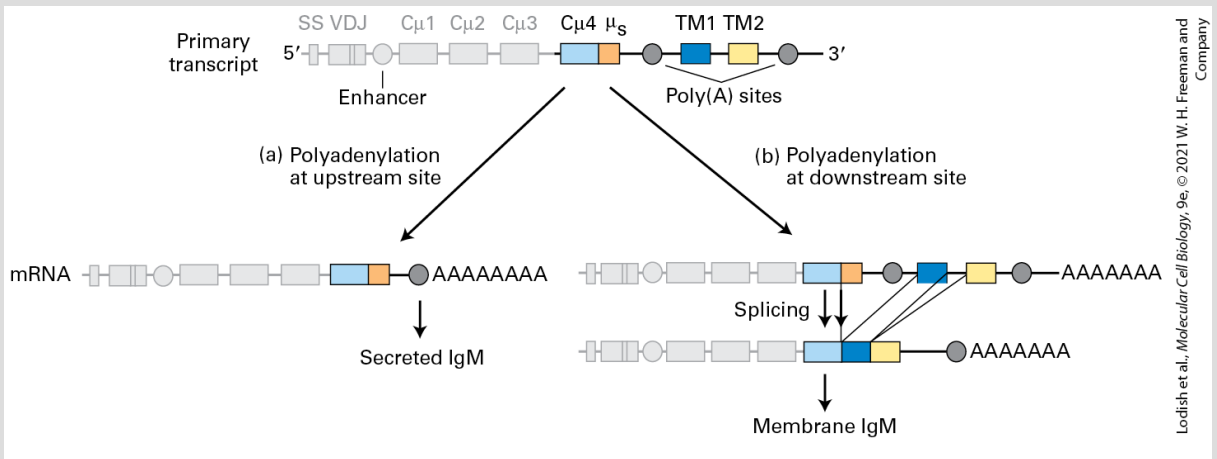


FIGURE 24-20 Synthesis of secreted and membrane IgM. The organization of the μ heavy-chain primary transcript is shown at the top: Cμ4 is the exon encoding the fourth μ constant-region domain; μ_s is a coding sequence unique for secreted IgM; TM1 and TM2 are exons that specify the transmembrane domain of the μ chain. Whether secreted or membrane-bound IgM is made depends on which poly(A) site is selected during processing of the primary transcript. (a) If the upstream poly(A) site is used, the resulting mRNA includes the entire Cμ4 exon and specifies the secreted form of the μ chain. (b) If the downstream poly(A) site is used, a splice donor site in the Cμ4 exon allows splicing to the transmembrane exons, yielding an mRNA that encodes the membrane-bound form of the μ chain. Similar mechanisms generate secreted and membrane-bound forms of other Ig isotypes. SS = signal sequence.

The capacity to switch from the synthesis of exclusively membrane-bound immunoglobulin to the synthesis of secreted immunoglobulin is acquired by B cells in the course of their differentiation. Terminally differentiated B cells, called *plasma cells*, are devoted almost exclusively to the synthesis of secreted antibodies (see Figure 24-7). Each plasma cell synthesizes and

secretes several thousand antibody molecules per second. It is this ramped-up production of secreted antibodies that underlies the effectiveness of the adaptive immune response in eliminating a pathogen and protecting against subsequent infection with the same pathogen. The protective value of antibodies is proportional to the concentration at which they are present in the circulation. Indeed, circulating antibody levels are often used as the key parameter to determine whether vaccination against a particular pathogen has been successful. The ability of plasma cells to secrete large amounts of immunoglobulins requires a massive expansion of the endoplasmic reticulum (ER), a hallmark of plasma cells. The unfolded-protein response (see Chapter 13) is initiated in B cells as an essential physiological mechanism to expand the ER and prepare the differentiating B cell for its future task as a highly active secretory cell. Interference with the unfolded-protein response abolishes the ability of B cells to turn into plasma cells.

B Cells Can Switch the Isotype of Immunoglobulin They Make

In the immunoglobulin heavy-chain locus, the exons that encode the constant segments of the μ chain lie immediately downstream of the rearranged VDJ exon (Figure 24-21, *top*). They are followed by exons that specify the constant segments of the δ chain. Transcription of a newly rearranged immunoglobulin heavy-chain

locus yields a single primary transcript that includes both the μ and δ constant regions. The splicing of this large transcript determines whether a μ chain or a δ chain will be produced. Downstream of the μ and δ exons are the exons that encode the constant segments of all the other heavy-chain isotypes. Upstream of each cluster of exons encoding one of the different isotypes (with the exception of the δ locus) is a repetitive sequence (switch region) that is recombination-prone, presumably because of its repetitive nature. Because each B cell necessarily starts out with surface IgM, recombination involving these sites, if it occurs, results in *class switching* from IgM to one of the other isotypes encoded downstream in the array of constant-region genes (Figure 24-21). The intervening DNA is deleted.

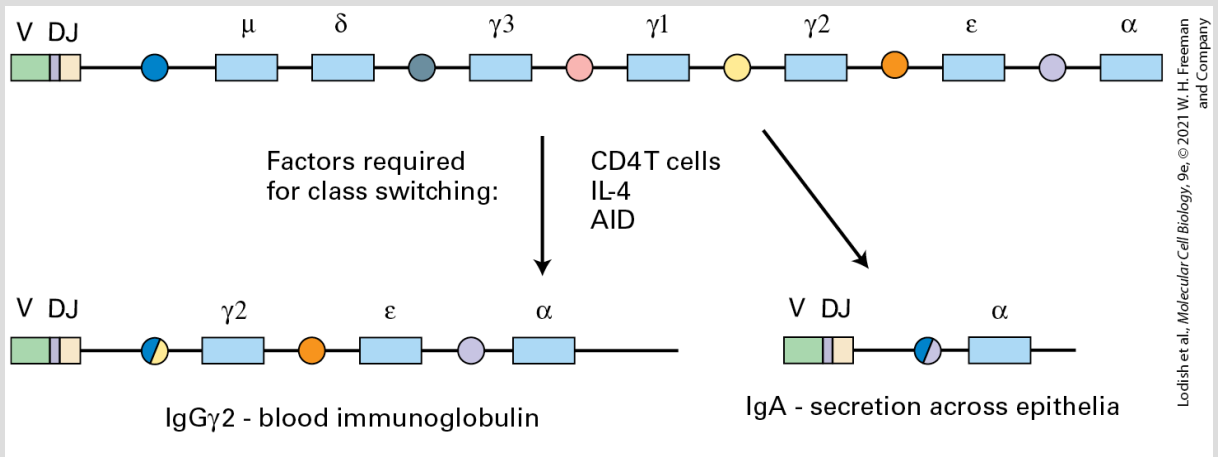


FIGURE 24-21 Class-switch recombination in the immunoglobulin heavy-chain locus. Class-switch recombination involves switch sites, which are repetitive sequences (colored circles) upstream of each of the heavy-chain constant-region genes. Recombination requires activation-induced deaminase (AID) as well as cytokines (e.g., IL-4) produced by certain helper T cells. Recombination eliminates the segment of DNA between the switch site upstream of μ exons and the constant region to which switching occurs. Class switching generates antibody molecules with the same specificity for antigen as that of the IgM-

bearing B cell that mounted the original response, but with different heavy-chain constant regions and therefore different effector functions.

In the course of its differentiation, a B cell can switch Ig classes sequentially. Importantly, the light chain is not affected by this process, nor is the rearranged VDJ segment with which the B cell started out on this pathway. Class-switch recombination thus generates antibodies with different constant regions, but identical antigenic specificity because the variable region has not changed. Each immunoglobulin isotype is characterized by its own unique constant region. As discussed previously, these constant regions determine the functional properties of the various isotypes. Class-switch recombination is dependent on the activity of AID and on the presence of antigen as well as on helper T cells. Somatic hypermutation and class-switch recombination occur concurrently, and their combined effect allows fine-tuning of the adaptive immune response with respect to the affinity of the antibodies produced and the effector functions employed.

KEY CONCEPTS OF SECTION 24.3

Generation of Antibody Diversity and B-Cell Development

- Functional antibody-encoding genes are generated by somatic rearrangement of multiple DNA segments at the heavy-chain and light-chain loci. These rearrangements involve V and J segments for immunoglobulin light chains and V, D, and J segments for immunoglobulin heavy chains (see Figure 24-15).
- Rearrangement of immunoglobulin gene segments is controlled by conserved recombination signal sequences (RSSs) composed of heptamers and nonamers

separated by 12- or 23-bp spacers (see Figure 24-16). Only those segments that have spacers of different lengths can rearrange successfully: two segments to be joined must possess a 12- and a 23-bp spacer, not two of identical length.

- The molecular machinery that carries out the rearrangement process includes proteins made only in lymphocytes (recombinases RAG1 and RAG2).
- Antibody diversity is created by the random selection of Ig gene segments to be recombined and by the ability of the heavy and light chains produced from rearranged Ig genes to associate with many different light chains and heavy chains, respectively.
- Junctional imprecision generates additional antibody diversity at the joints of the gene segments brought together by somatic gene rearrangements.
- Further antibody diversity arises after B cells encounter antigen as a consequence of somatic hypermutation, which can lead to the selection and proliferation of B cells producing the highest affinity antibodies, a process termed *affinity maturation*.
- During B-cell development, heavy-chain genes are rearranged first, leading to expression of the pre-B-cell receptor. Subsequent rearrangement of light-chain genes results in assembly of an IgM membrane-bound B-cell receptor (see Figure 24-19).
- Only one of the allelic copies of the heavy-chain locus and of the light-chain locus is rearranged (allelic exclusion), ensuring that a B cell expresses Ig with a single antigenic specificity.
- Polyadenylation at different poly(A) sites in an Ig primary transcript determines whether the membrane-bound or secreted form of an antibody is produced (see Figure 24-20).
- During an immune response, class switching allows B cells to adjust the class of antibody made and thus the effector functions of the immunoglobulins produced, while retaining the antibody's specificity for antigen (see Figure 24-21).

24.4 The MHC and Antigen Presentation

Antibodies can bind their target antigen without the involvement of other molecules; the presence of antigen and antibody is sufficient for their interaction. In the course of their differentiation, B cells receive essential assistance in the form of signals sent from T cells by a process that will be described in some detail below. This process, literally called *T-cell help*, is antigen-specific, and the T cells responsible for providing it are *helper T cells*. Although antibodies contribute to the elimination of bacterial and viral pathogens, it may also be necessary to destroy infected host cells that might serve as a source of new virus particles. This task is carried out by *cytotoxic T cells*.

Both helper T cells and cytotoxic T cells make use of an antigen-specific T-cell receptor on its surface that, as we see in Section 24.5, is encoded by genes that are generated by mechanisms analogous to those used by B cells to generate immunoglobulin genes — including gene rearrangements. However, T cells recognize their cognate antigens in a manner very different from that used by B cells. The antigen-specific receptors on T cells recognize short segments of protein antigens but can do so only when these peptides are noncovalently bound to a glycoprotein complex present on the external surface of an *antigen-presenting cell*. Importantly, each T cell

in an individual expresses many copies of a unique T-cell receptor on its surface, and different T-cell receptors will bind to different sets of peptides bound to the glycoprotein complex on an antigen-presenting cell.

The membrane glycoprotein complex that presents the antigen snippets is encoded in a region of genomic DNA called the *major histocompatibility complex (MHC)*. As we will see in detail, various antigen-presenting cells, in the course of their normal activities, digest pathogen-derived proteins (as well as their own proteins) and then present complexes consisting of an MHC protein bound to a protein snippet (usually a small peptide) on their cell surface. T cells can scrutinize these complexes, and if they detect a pathogen-derived peptide bound to the MHC molecule, the T cells take appropriate action, which may include killing the cell that carries the MHC-peptide complex. In this section, we describe the MHC and the proteins it encodes, then examine how MHC molecules are involved in antigen presentation and antigen recognition by T cells.

The MHC Determines the Ability of Two Unrelated Individuals of the Same Species to Accept or Reject Grafts

The MHC was discovered, as its name implies, as the genetic locus that controls acceptance or rejection of tissue grafts. At a time when

tissue culture had not yet been developed to the stage at which tumor-derived cell lines could be propagated in the laboratory, investigators relied on serial passage of tumor tissue in vivo (i.e., transplanting a tumor from one mouse to another). It was quickly observed that a tumor that arose spontaneously in one inbred strain of mice could be propagated successfully in the strain in which it arose but generally not in a genetically distinct strain of mice. Genetic analysis showed that a single major genetic locus in the recipient mouse was responsible for this behavior. Similarly, transplantation of healthy skin was feasible within the same strain of mice, but not when the recipient was of a genetically distinct background. Genetic analysis of transplant rejection likewise identified a single major locus — the same one responsible for tumor rejection — that controlled acceptance or rejection, which is an immune reaction. This locus has been expanded to the designation MHC to reflect the presence of multiple genes in this region. As we now know, all vertebrates that possess an adaptive immune system have a genetic region that corresponds to the major histocompatibility complex as originally defined in the mouse.

In mice, the genetic region responsible for graft rejection is called the *H-2 complex* (Figure 24-22a). In humans, the genetic region encoding the MHC was uncovered during the study of patients who underwent multiple blood transfusions that provoked an immune response. The human MHC region is called the *HLA complex* (Figure 24-22b). The typical mammalian MHC contains dozens of genes,

many encoding proteins of immunological relevance, not all of which are shown in the figure. All vertebrate MHCs encode a highly homologous set of proteins, although the details of organization and gene content show considerable variation between species, as seen for domestic chickens, mice, and humans. As we will see, several MHC proteins present antigenic peptides for recognition by the immune system. Since most cells in vertebrates express MHC proteins, they therefore have the potential to present antigens — especially antigen derived from pathogenic bacteria and viruses — to the adaptive immune system.

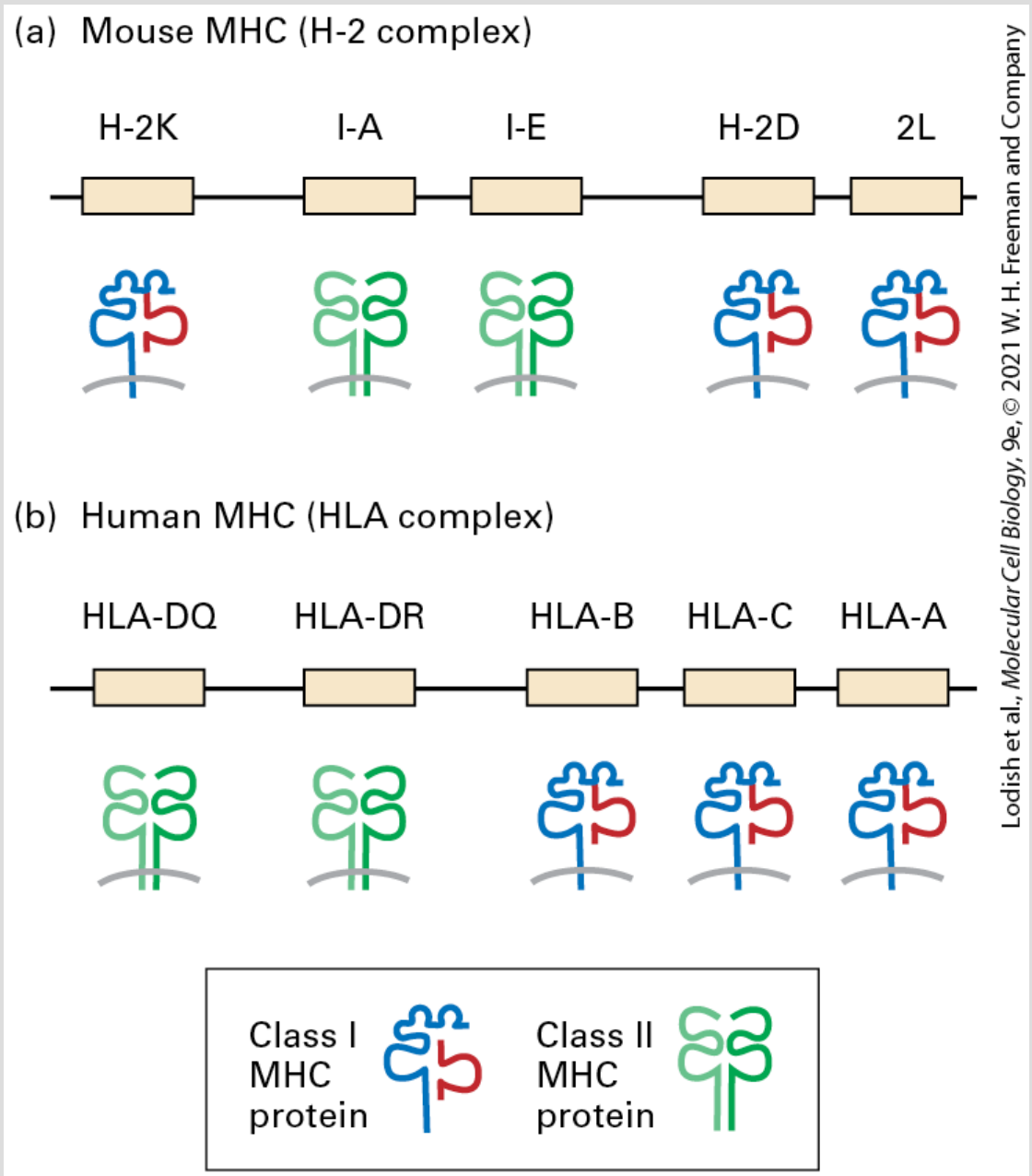


FIGURE 24-22 Organization of the major histocompatibility complex in mice and in humans. The major loci are depicted with schematic diagrams of their encoded proteins below. Class I MHC proteins are composed of an MHC-encoded glycoprotein with a single membrane-spanning sequence, in noncovalent association with a small subunit, called *β2-microglobulin*, which is not encoded in the MHC and is not membrane bound. Class II MHC proteins consist of two nonidentical single-pass transmembrane glycoproteins, both of which are encoded by the MHC.

Interestingly, the human fetus may be considered a tissue graft in the mother: the fetus shares only half of its genetic material with the mother, the other half being contributed by the father. Antigens encoded by the paternal alleles may differ sufficiently from their maternal counterparts to elicit an immune response in the mother. Such a response can occur because in the course of pregnancy, fetal cells that slough off into the maternal circulation can stimulate the maternal immune system to mount an antibody response against the paternal antigens. We now know that these antibodies recognize proteins encoded by the human MHC. The fetus itself is spared rejection because of the specialized organization of the placenta, which prevents initiation of an immune response by the mother against fetal tissue.

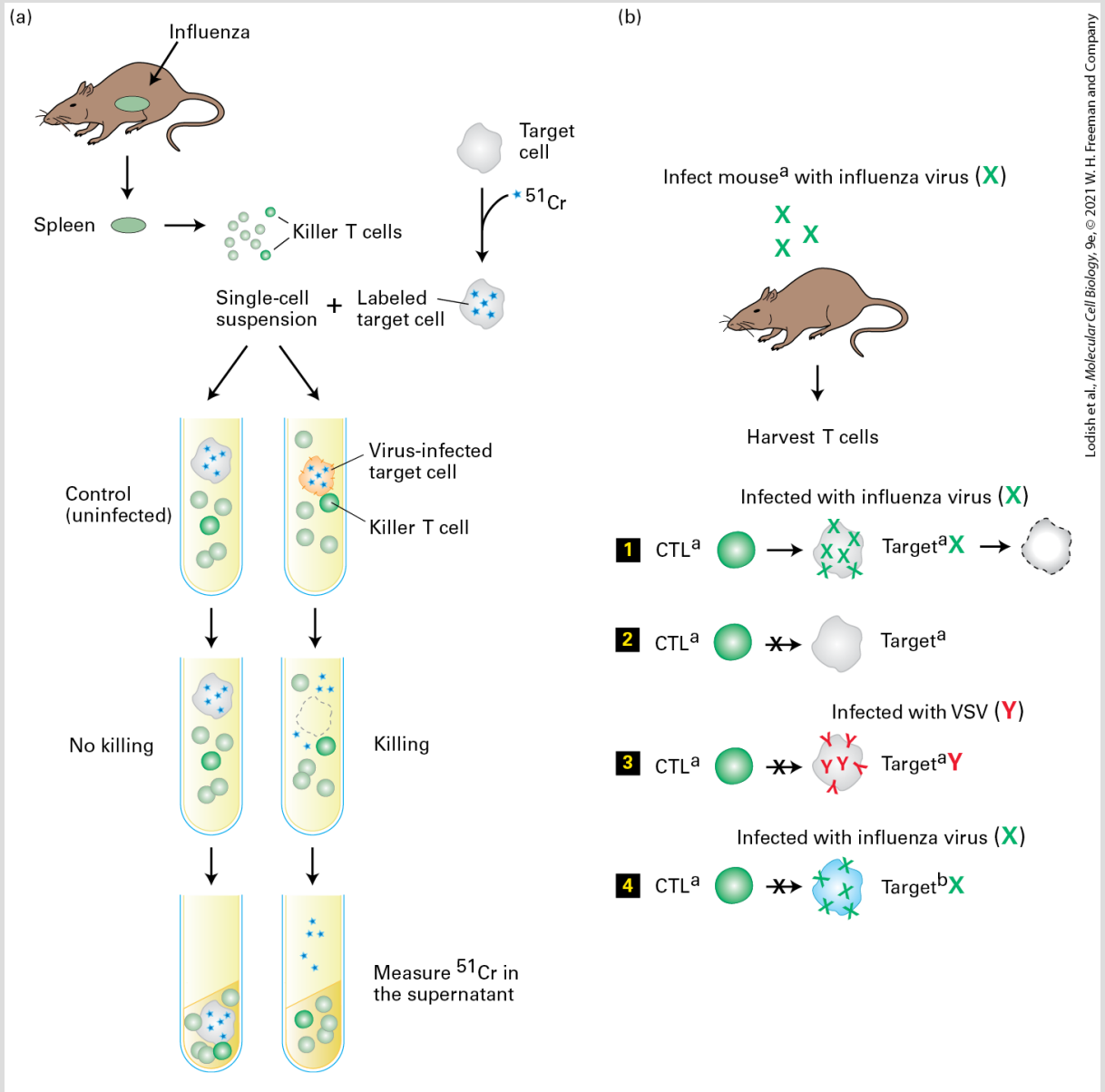
The Killing Activity of Cytotoxic T Cells Is Antigen Specific and MHC Restricted

Clearly MHC molecules did not evolve to prevent the exchange of surgical grafts. MHC molecules play an essential role in the recognition of virus-infected cells by a type of T cell termed a *cytotoxic T cell*, which is also called a *cytolytic T lymphocyte (CTL)*. In virus-infected cells, MHC molecules bind to peptides derived from the viral proteins and display these fragments on the cell surface, where CTLs, charged with eliminating the infection, can recognize them. How such fragments of antigen are generated and displayed

will be described below. CTLs that have receptors capable of recognizing a particular peptide-MHC complex unleash a payload of lethal molecules onto the infected target cells, destroying the target-cell membranes. The destruction of these target cells can be readily measured by the release of their cytoplasmic contents when they physically disintegrate. Thus CTL killing of infected host cells requires: (1) MHC presentation of antigenic peptides from the pathogen on the host-cell surface, (2) CTLs expressing antigen-MHC–specific T-cell receptors on their surface that can recognize the MHC-antigen complex, and (3) the activation of the CTL killing machinery once the T-cell receptors have bound to the MHC-antigen complex.

Mice that have recovered from a particular viral infection are a ready source of CTLs that can recognize and kill target cells infected with the same virus. If CTLs are obtained from a mouse that has successfully cleared an infection with influenza virus, cytotoxic activity is observed against influenza-infected target cells, but not against uninfected controls (Figures 24-23a and 24-23b, experiment **2**). Moreover, the influenza-specific CTLs will not kill target cells infected with a different virus, such as vesicular stomatitis virus (Figure 24-23b experiment **3**). CTLs can even discriminate between closely related strains of influenza virus and can do so with pinpoint precision: differences of a single amino acid in the viral antigen may suffice to prevent recognition and killing by CTLs. These experiments show that CTLs are truly antigen specific and do not

simply recognize some attribute that is shared by all virus-infected cells, regardless of the identity of the virus.



EXPERIMENTAL FIGURE 24-23 Chromium (^{51}Cr) release assay allows direct demonstration of the cytotoxicity and specificity of cytotoxic T cells in a heterogeneous population of cells. (a) A suspension of spleen cells containing cytotoxic (killer) T cells is prepared from mice that have been exposed to a particular virus (e.g., influenza virus) and have cleared the infection. Target cells obtained from mice of the same strain are infected with the identical virus or left uninfected. After infection, cellular proteins are labeled

nonspecifically by incubation of the target-cell suspension with ^{51}Cr. When the radiolabeled target cells are incubated with the suspension of cytotoxic T cells, the killing of infected target cells results in release of the ^{51}Cr-labeled proteins. Uninfected target cells are not killed and retain their radioactive contents. Lysis of cells by cytotoxic T cells can therefore be readily detected and quantitated by measuring the radioactivity released into the supernatant. (b) Cytotoxic T cells (CTLs) harvested from mice that have been infected with influenza virus (X) can be tested against various target cells to determine the specificity of CTL-mediated killing. CTLs capable of lysing influenza virus–infected target cells (**1**) cannot kill uninfected cells (**2**) or cells infected with vesicular stomatitis virus (Y), for example (**3**). When these CTLs are tested on influenza virus–infected targets from a strain of mice that carries an altogether different MHC type (strain b, blue target cells), again no killing is observed (**4**). Cytotoxic T-cell activity is thus virus specific and restricted by the MHC.

In the example in Figure 24-23a, we assumed that the CTLs harvested from an influenza-immune mouse are assayed on influenza-infected target cells derived from that same strain of mouse (strain a). However, if target cells from a completely unrelated strain of mouse (strain b, blue-colored target cell in Figure 24-23b experiment **4**) are infected with influenza and used as targets, the CTLs from the strain a mouse are unable to kill the infected strain b target cells (Figure 24-23b, **1** vs. **4**). It is therefore not sufficient that the antigen (an influenza-derived protein) is present; recognition of the antigen by CTLs is *restricted* by mouse strain-specific elements. Genetic mapping has shown that these restricting elements are encoded by genes in the MHC. Thus CTLs from a mouse strain that is immune to influenza will kill influenza-infected target cells from another mouse strain only if the two strains match at the MHC loci for the relevant MHC molecules.

This phenomenon is therefore known as *MHC restriction*, and the MHC molecules involved are called *restriction elements*.

T Cells with Different Functional Properties Are Guided by Two Distinct Classes of MHC Molecules

The MHC encodes two types of glycoproteins essential for immune recognition, commonly called *class I* and *class II MHC molecules*. The genetic maps of mouse and human MHCs show the presence of several class I MHC genes and several class II MHC genes, although their arrangements differ between the two species (see Figure 24-22). In addition to class I and class II MHC molecules, the MHC encodes key components of the antigen-processing (e.g., proteolysis) and presentation machinery. Finally, the typical vertebrate MHC also encodes components of the complement cascade.

Both class I and class II MHC proteins are involved in presenting antigen to T cells, but they serve two broadly distinct functions. Class I MHC products present antigens to cytotoxic T cells, licensing them to destroy infected cells. Cytotoxic T cells are characterized by the expression of CD8, a surface glycoprotein that determines the ability of the T cells that carry it to interact with class I MHC products. Most, if not all, nucleated cells constitutively express class

I MHC molecules, and many can support replication of viruses. Cytotoxic T cells thus recognize and kill infected cells of all body types via cell-surface class I MHC molecules that have bound to it (display) a virus-derived antigen (peptide).

Class II MHC products are found exclusively on specialized antigen-presenting cells (APCs), also called *professional* APCs. These APCs present antigens bound to class II MHC molecules to a class of T lymphocytes called *helper T cells*. This presentation is the start of an adaptive immune response that also enables cytotoxic T cells to kill their targets and assists B cells in producing antigen-specific antibodies. B cells cannot undergo final differentiation into antibody-secreting plasma cells without assistance from helper T cells. Helper T cells express a surface glycoprotein called *CD4* and interact with class II but not class I MHC molecules on target cells. The constitutive expression of class II MHC molecules is confined to professional APCs, which include B cells, dendritic cells, and macrophages. (Several other cell types, such as some epithelia, can be induced to express class II MHC molecules under specific circumstances, but we will not discuss them.) Again, the underlying cell biology that describes the expression, assembly, and mode of antigen presentation by class II MHC molecules fits this functional specialization rather neatly, as we shall see below.

The two major groups of functionally distinct T lymphocytes — cytotoxic T cells and helper T cells — can thus be distinguished by

the unique profile of membrane proteins displayed at their cell surface and by the MHC molecules they use as targets; these targets are often referred to as *restriction elements*.

- Cytotoxic T cells: CD8 marker; class I MHC restricted
- Helper T cells: CD4 marker; class II MHC restricted

Both CD4 and CD8, along with many other proteins of the immune system, including B-cell and T-cell receptors and the polymeric IgA receptor, belong to the immunoglobulin (Ig) superfamily of proteins, all of which have one or more Ig domains. The molecular basis for the strict correlation between expression of CD8 and use of class I MHC molecules as the restriction element, or between expression of CD4 and use of class II MHC molecules as the restriction element, will become evident once the structure and mode of action of MHC molecules has been described.

MHC Molecules Are Highly Polymorphic, Bind Peptide Antigens, and Interact with the T-Cell Receptor

Polymorphisms in the MHC Locus Are the Basis of Transplant Rejection

Both class I and class II MHC molecules are highly *polymorphic*; that is, thousands of allelic variants, encoding similar proteins but with slightly different amino acid sequences, exist among individuals of the same species. The vertebrate immune system can respond to these allelic differences, and its ability to recognize allelic MHC variants is the underlying immunological cause for rejection of transplants that involve unrelated, genetically distinct individuals.

The two classes of MHC molecules are structurally similar in many respects, as are their interactions with peptides and the T-cell receptor. MHC molecules are particularly important for recognizing self tissue and distinguishing it from nonself (and thus possibly pathogenic) substances. In general, except for close relatives, any two individuals have a very low chance of sharing the same MHC variants. Any inter-individual differences in MHC molecules in a graft recipient and donor will be recognized by the recipient's immune system, which will treat the graft as foreign and eliminate it (graft rejection). The greater the similarity in the set of MHC alleles of a donor and a transplant recipient, the greater the chance that the transplant will be accepted. This is why surgeons look for an MHC-matched individual to donate an organ. If the tissue type (MHC alleles) of the donor does not exactly match that of the recipient, it is necessary to use drugs that suppress the immune responses of the recipient to prevent organ rejection.

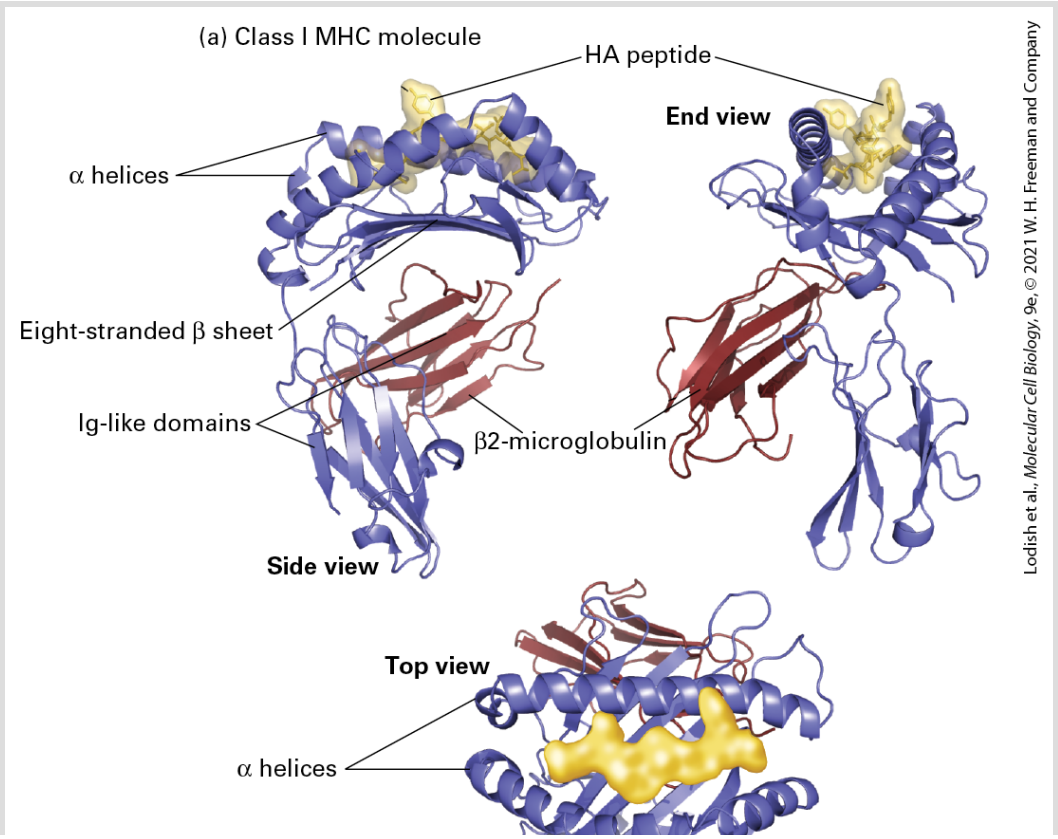
The cell-biological mechanisms by which the immune system develops the capacity to distinguish self from nonself (or nonpathogenic from pathogenic) are complex, yet worth understanding. Understanding the molecular and cellular basis of immunity has enormous practical consequences for medicine and public health. We will therefore consider these molecular and cellular mechanisms in detail, beginning with the structures of the MHC proteins and the way by which peptides become bound to them.

Class I MHC Molecules

Class I MHC molecules, which belong to the Ig superfamily, consist of two polypeptide subunits. The larger subunit, for which there are multiple independent gene copies in the MHC region of mammalian genomes, is a type I membrane glycoprotein (see Figure 13-10). The smaller β2-microglobulin subunit is not encoded by the MHC and corresponds in structure to an Ig domain. The larger subunits of class I MHC molecules in humans are encoded by the HLA-A, HLA-B, and HLA-C loci (see Figure 24-22). In the mouse, the larger subunits of class I MHC molecules are encoded by the H-2K, H-2D, and H-2L loci.

The three-dimensional structure of a class I MHC molecule shows two Ig-like domains (Figure 24-24a). These domains support an eight-stranded β sheet topped by two α helices. Jointly, the β sheet and the helices create a cleft, closed at both ends, in which a peptide

binds. The mode of peptide binding by a class I MHC molecule requires that the peptide be about 8–10 amino acids long, so that the ends of the peptide can be tucked into pockets that accommodate the charged amino and carboxyl groups at the termini of the peptide. Further, the peptide is anchored into the peptide-binding cleft by means of a small number of amino acid side chains, each of which is accommodated by a pocket in the MHC molecule that neatly fits that particular anchor amino acid residue (Figure 24-25a). On average, two such specificity pockets must be filled correctly to allow stable peptide binding, restricting binding to peptides with side chains that can fit into these pockets. In this manner, a given MHC molecule can accommodate a large number of peptides of diverse, yet circumscribed, sequence.



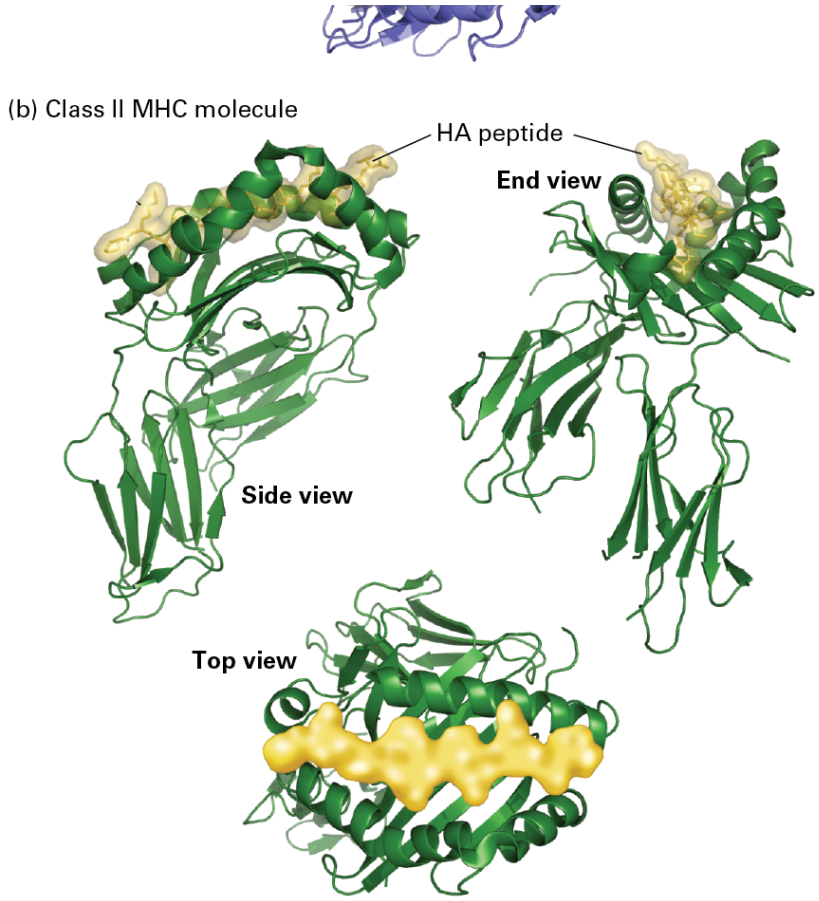
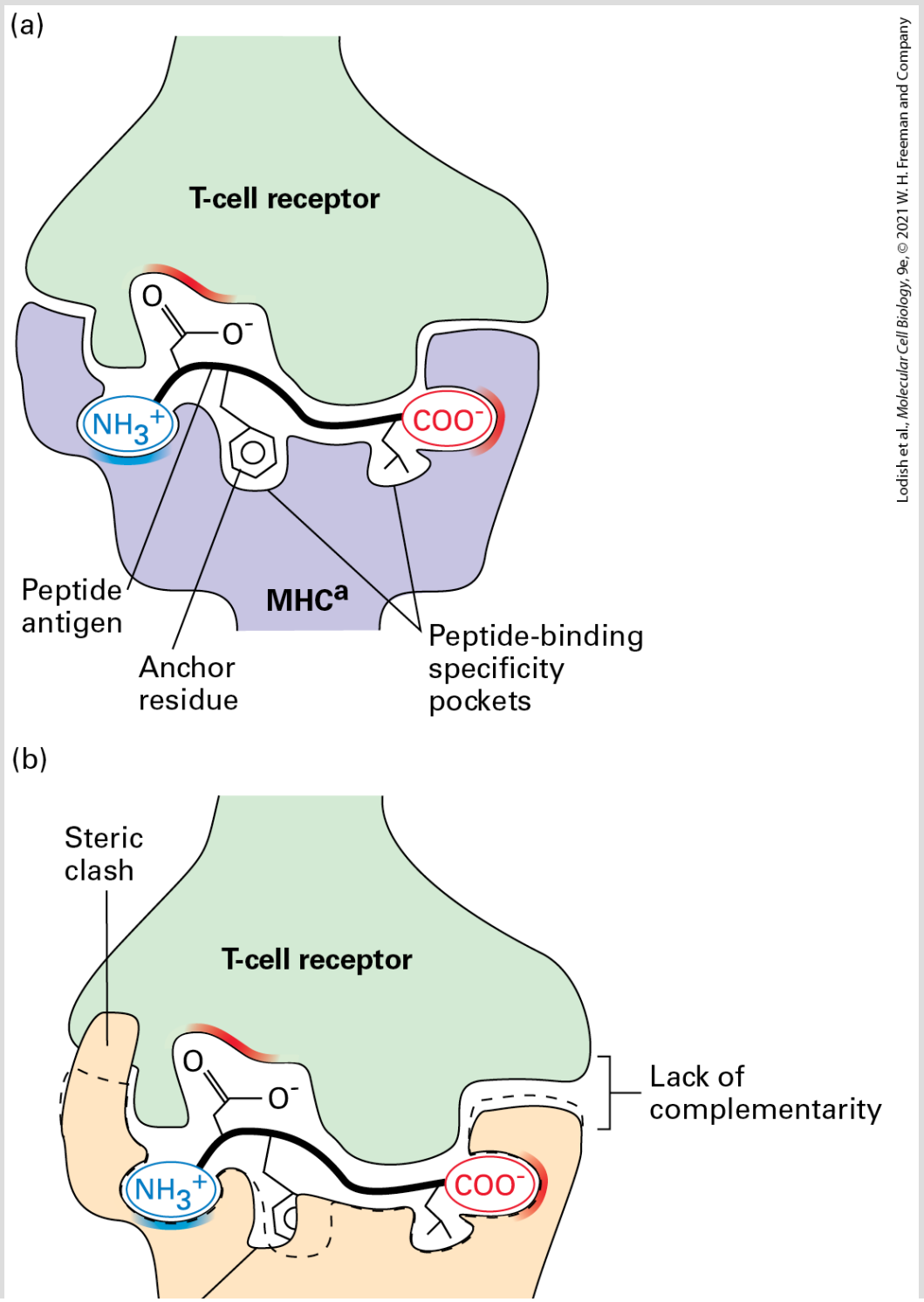


FIGURE 24-24 Three-dimensional structure of class I and class II MHC molecules. (a) Shown here is the structure of a class I MHC molecule with a bound antigenic (HA) peptide as determined by x-ray crystallography. The portion of a class I MHC molecule that binds a peptide consists of a β sheet composed of eight β strands and flanked by two α helices. The peptide-binding cleft is formed entirely from the MHC-encoded large subunit, which associates noncovalently with the small subunit (β2-microglobulin). (b) Class II MHC molecules are structurally similar to class I molecules but with several important distinctions. Both the α and β subunits of class II MHC molecules are MHC encoded and contribute to formation of the peptide-binding cleft. The peptide-binding cleft of class II MHC molecules accommodates a wider range of peptide sizes than that of class I molecules. Both class I and class II MHC proteins span the cell membrane, and each includes a transmembrane segment and a cytoplasmic tail (see Figures 24-21, 24-26, and 24-29). These are not included in the crystallographic analysis and are not shown in the figure.

[Part (a) data from D. N. Garboczi, 1996, *Nature* **384**:134, PDB ID 1ao7. Part (b) data from J. Hennecke, A. Carfi, and D. C. Wiley, 2000, *EMBO J*. **19**:5611, PDB ID 1 fyt.]



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

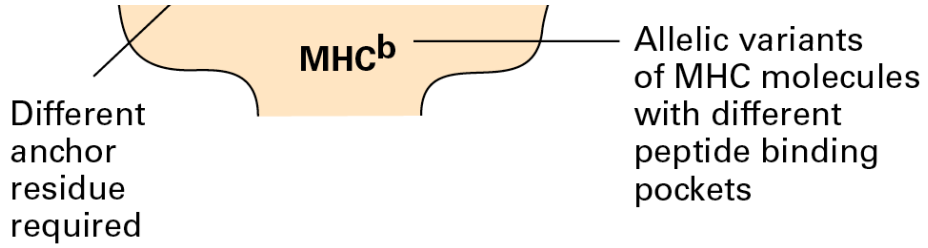


FIGURE 24-25 Peptide binding and restriction in class I MHC. (a) Peptides that bind to class I molecules are on average 8–10 residues in length, require proper accommodation of the termini, and include two or three residues that are conserved (anchor residues). Amino acid residues in class I molecules that distinguish one allele from another (polymorphic residues) occur in and around the peptide-binding cleft. The polymorphic residues in the MHC affect both the specificity of peptide binding and interactions with T-cell receptors. Successful recognition of an antigenic peptide–MHC complex by a T-cell receptor requires a good fit among the receptor, peptide, and MHC molecule. (b) Steric clash and a lack of complementarity between anchor residues and the MHC molecule, and a similar clash between the T-cell receptor and the MHC molecule, prevent proper binding. Each T cell in an individual expresses one type of T-cell receptor, one that is thus restricted to binding specific peptide-MHC complexes.

The polymorphic residues that distinguish one allelic MHC molecule from another are located mainly in and around the peptide-binding cleft. These residues therefore determine the architecture of the peptide-binding pocket and hence the specificity of peptide binding. Further, these polymorphic residues affect the surface of the MHC molecule that makes contact with the T-cell receptor. A T-cell receptor that can interact with one particular class I MHC allele will therefore, as a rule, not interact with unrelated MHC molecules because of their different surface architectures (Figure 24-25b); this is the molecular basis of MHC restriction. The CD8 molecule on cytotoxic T cells functions as a co-receptor,

binding to conserved portions of the class I MHC molecule. The presence of CD8 thus sets the class I MHC preference of any mature T cell that bears it.

Class II MHC Molecules

The two subunits (α and β) of class II MHC molecules are both type I membrane glycoproteins of the Ig superfamily. The typical mammalian MHC contains several loci that encode class II MHC molecules (see Figure 24-22). Like the large subunit of class I molecules, both the α and β subunits of class II molecules show genetic polymorphism.

The basic three-dimensional design of class II MHC molecules resembles that of class I MHC molecules. Two membrane-proximal Ig-like domains support a peptide-binding portion with a peptide-binding cleft (see Figure 24-24b). In class II MHC molecules, the α and β subunits contribute equally to the construction of the peptide-binding cleft. This cleft is open at both ends and thus supports the binding of peptides longer than those that bind to class I MHC molecules because the peptides can protrude from both ends of the cleft. The mode of peptide binding involves pockets that accommodate specific amino acid side chains as well as contacts between side chains of the MHC molecule and main-chain atoms of the bound peptide. As for class I MHC, class II MHC polymorphisms mainly affect residues in and around the peptide-binding cleft, so that peptide-binding specificity usually differs among different

allelic products. A T-cell receptor that interacts with a particular class II MHC molecule will not, as a rule, interact with a different class II MHC allelic variant, not only because of the difference in the peptide-binding specificity of the allelic MHC molecules, but also because of the polymorphisms that affect the residues that contact the T-cell receptor. Class I MHC is the basis for class II MHC restricted recognition of antigens. As we saw for class I MHC molecules and their CD8 co-receptor, the CD4 co-receptor recognizes conserved features on class II MHC molecules. Any mature T cell that bears the CD4 co-receptor uses class II MHC molecules for antigen recognition.

As we will see below, class II MHC molecules evolved to present peptides generated predominantly in endosomes and lysosomes. Peptide binding to a class II MHC molecule takes place in those organelles, and class II MHC molecules are targeted specifically to those locations after their synthesis in the endoplasmic reticulum. This targeting is accomplished by means of a chaperone called the *invariant chain*, a type II membrane glycoprotein (see Figure 13-10). The invariant chain, termed *Ii*, plays a key role in the early stages of class II MHC biosynthesis by forming in the endoplasmic reticulum a trimeric structure onto which three newly made class II MHC αβ heterodimers assemble. The final assembly product thus consists of nine polypeptides: $(\alpha\beta Ii)_3$. The interaction between Ii and the αβ heterodimer involves a stretch of Ii, called the *CLIP segment*, which occupies the class II MHC peptide-binding cleft. Once the $(\alpha\beta Ii)_3$

complex is assembled, the complex enters the secretory pathway and is diverted to endosomes and lysosomes at the *trans*-Golgi network (see Figure 14-2). The signals for this diversion are carried by the Ii cytoplasmic tail.

In Antigen Presentation, Protein Fragments Are Complexed with MHC Products and Posted to the Cell Surface

The process by which foreign materials enter the immune system is the key step that determines the eventual outcome of an immune response. A successful adaptive immune response, which includes the production of antibodies and the generation of helper and cytotoxic T cells, absolutely requires the involvement of professional APCs. These cells are what acquire antigen, process it and then display it in a form that can be recognized by T cells. The pathway that converts antigen into a form suitable for T-cell recognition is referred to as **antigen processing and presentation**.

The class I MHC pathway focuses predominantly on presentation of fragments of proteins synthesized by the cell itself (including pathogen-encoded proteins found inside the infected cells), whereas the class II MHC pathway is centered on materials acquired from outside the APC. Recall that all nucleated cells express class I MHC

products or can be induced to do so. This makes sense in view of the fact that a nucleated cell is capable of synthesizing nucleic acids and proteins and can thus in principle sustain replication of a viral pathogen. The ability to alert the immune system to the presence of an intracellular invader is inextricably linked to class I MHC-restricted antigen presentation. The distinction between the presentation of materials synthesized by an APC itself and the processing and presentation of antigen acquired from outside the cell is by no means absolute. Together, the class I and class II pathways of antigen processing and presentation sample all of the cellular and external compartments that need to be surveyed for the presence of pathogens.

Antigen processing and presentation in both the class I and class II pathways may be divided into six discrete steps that are useful in comparing the two pathways: (1) acquisition of antigen, (2) tagging the antigen for destruction, (3) proteolysis, (4) delivery of peptides to MHC molecules, (5) binding of peptide to an MHC molecule, and (6) display of the peptide-loaded MHC molecule on the cell surface. Here we describe the molecular details of each pathway.

The Class I MHC Pathway Presents Cytosolic Antigens

Figure 24-26 summarizes the six steps in the class I MHC pathway using a virus-infected cell as an example.

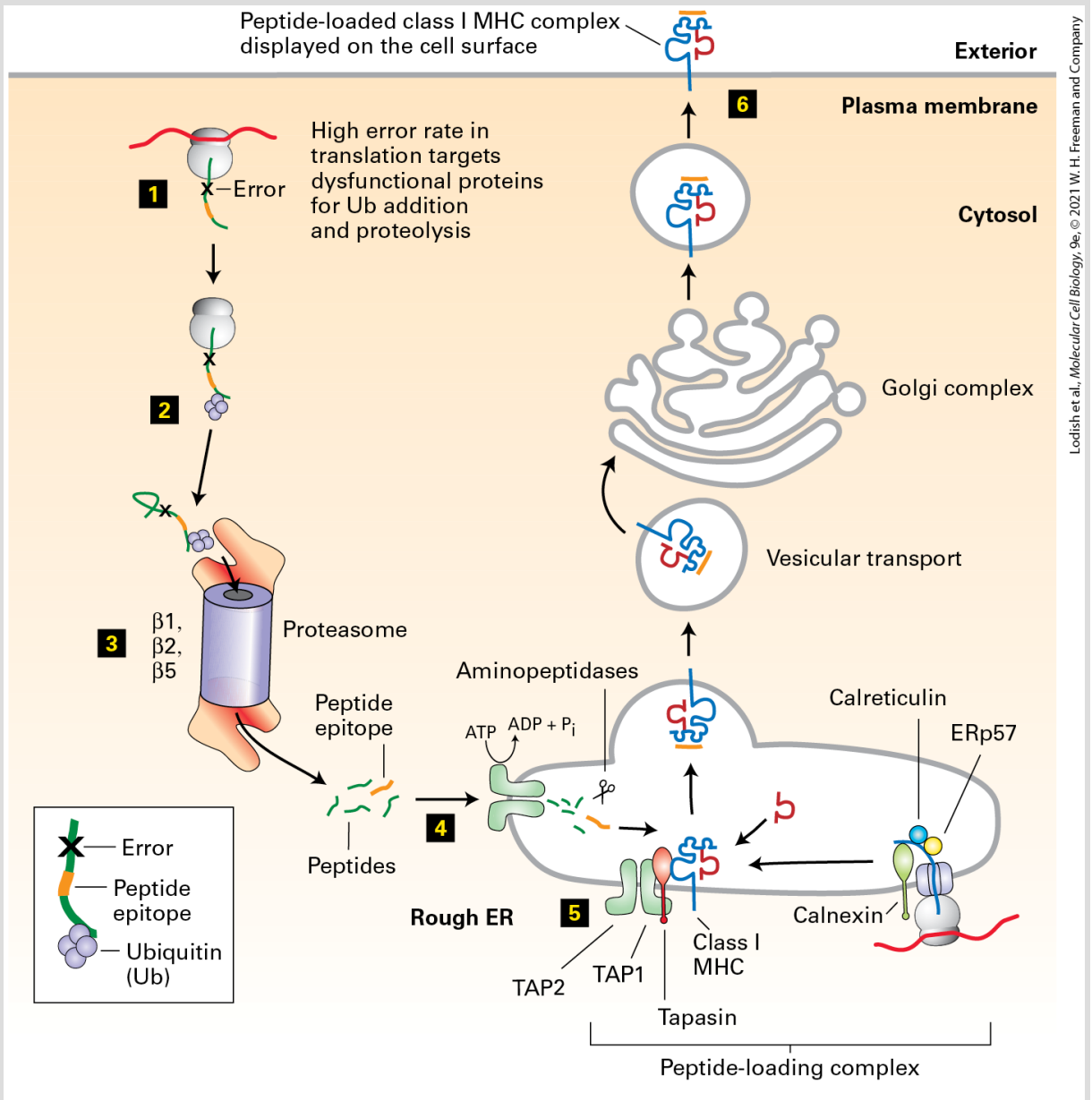


FIGURE 24-26 Class I MHC pathway of antigen processing and presentation. Step 1: Acquisition of antigen is synonymous with the production of proteins with errors (e.g., premature termination, misincorporation) or the production of pathogen-derived proteins. Step 2: Dysfunctional proteins are targeted for degradation by ubiquitinylation. Step 3: Proteolysis is carried out by the proteasome. In cells exposed to interferon γ, the

catalytically active β subunits of the proteasome are replaced by interferon-induced immune-specific β subunits. Step 4: Peptides are delivered to the interior of the endoplasmic reticulum (ER) via the dimeric TAP peptide transporter. Step 5: Each peptide is loaded onto a newly made class I MHC molecule within the peptide-loading complex. Step 6: The fully assembled class I MHC–peptide complex is transported to the cell surface via the secretory pathway. See text for details.

1. *Acquisition of Antigen:* In the case of a viral infection, acquisition of antigen is usually synonymous with the infected state. Viruses rely on the host protein synthesis apparatus to generate new viral proteins. Protein synthesis, unlike DNA replication, is an error-prone process, in which a fraction of newly initiated polypeptide chains are terminated prematurely or suffer from other errors (e.g., misincorporation of amino acids, frameshifts, improper or delayed folding). These mistakes in protein synthesis affect the host cell's own proteins and those specified by viral genomes equally. Such error-containing proteins must be rapidly removed so as not to clog up the cytoplasm, engage partner proteins in nonproductive interactions, or even act as dominant negative versions of a protein. Properly folded proteins may also sustain damage that leads to their unfolding, completely or in part, and necessitates their removal. These proteins are an important source of the peptides destined for presentation by class I MHC molecules. With the exception of a specialized process called *cross-presentation* (discussed below), the class I MHC pathway results in the formation of peptide-MHC complexes in which the peptides are derived from proteins synthesized by the class I MHC-bearing cell itself.

2. *Targeting Antigen for Destruction:* For the most part, polyubiquitinylation is responsible for targeting a protein for destruction (see Section 3.4). Polyubiquitinylation is a covalent modification that is tightly regulated.
3. *Proteolysis:* Polyubiquitinylated proteins are destroyed by proteolysis in proteasomes. The proteasome is a protease that engages its substrates and, without the release of intermediates, yields peptides in the size range of 3–20 amino acids as its final digestion products (see Figure 3-32). During the course of an inflammatory response and in response to interferon γ, the three catalytically active β subunits (β1, β2, β5) of the proteasome can be replaced by three immune-specific subunits: β1i, β2i, and β5i, subunits that are encoded in the MHC region of the genome. The net result of this replacement is the generation of an *immunoproteasome*, the output (length of peptide products) of which is matched to the requirements for peptide binding by class I MHC molecules. The immunoproteasome adjusts the average length of the peptides produced, as well as the sites at which cleavage occurs. Given the central role of the proteasome in the generation of the peptides presented by class I MHC molecules, proteasome inhibitors interfere potently with antigen processing via the class I MHC pathway.
4. *Delivery of Peptides to Class I MHC Molecules:* Protein synthesis, polyubiquitinylation, and proteasomal proteolysis all occur in the cytosol, whereas peptide binding by class I MHC molecules occurs in the lumen of the endoplasmic reticulum (ER). Thus peptides must cross the ER membrane to gain access to the

peptide-binding sites of newly made class I molecules, a process mediated by the heterodimeric TAP complex, a member of the ABC superfamily of ATP-powered pumps (see Figure 11-15). The TAP complex binds peptides on the cytoplasmic face of the ER and, in a cycle that includes ATP binding and hydrolysis, translocates them into the ER. The specificity of the TAP complex is such that it can transport only a subset of all cytosolic peptides, primarily those in the length range of 5–10 amino acids, that are compatible with the circumscribed length of peptides that can fit into the class I MHC molecules. The mouse TAP complex shows a pronounced preference for peptides that with C-terminal leucine, valine, isoleucine, or methionine residues match the binding preference of class I MHC molecules. The genes encoding the TAP1 and TAP2 subunits composing the TAP complex are located in the MHC region. Peptidases in the cytosol and in the lumen of the ER may further modify the products of proteasomal proteolysis.

5. *Binding of Peptides to Class I MHC Molecules:* Within the ER, newly synthesized class I MHC molecules are part of a multiprotein complex referred to as the *peptide-loading complex* (Figure 24-27). This complex includes two chaperones (calnexin and calreticulin) and the oxidoreductase ERp57. Another chaperone (tapasin) interacts with both the TAP complex and the class I MHC molecule about to receive peptide. The physical proximity of TAP and the class I MHC molecule is maintained by tapasin. Once peptide loading onto the class I MHC molecule has occurred, a conformational change releases the loaded

class I MHC molecule from the peptide-loading complex. This arrangement effectively ensures that only peptide-loaded class I MHC molecules are released from the ER and then transported to and displayed at the cell surface. The overall efficiency of this pathway is such that approximately 4000 molecules of a given protein must be destroyed to generate a single MHC-peptide complex carrying a peptide from that particular polypeptide.

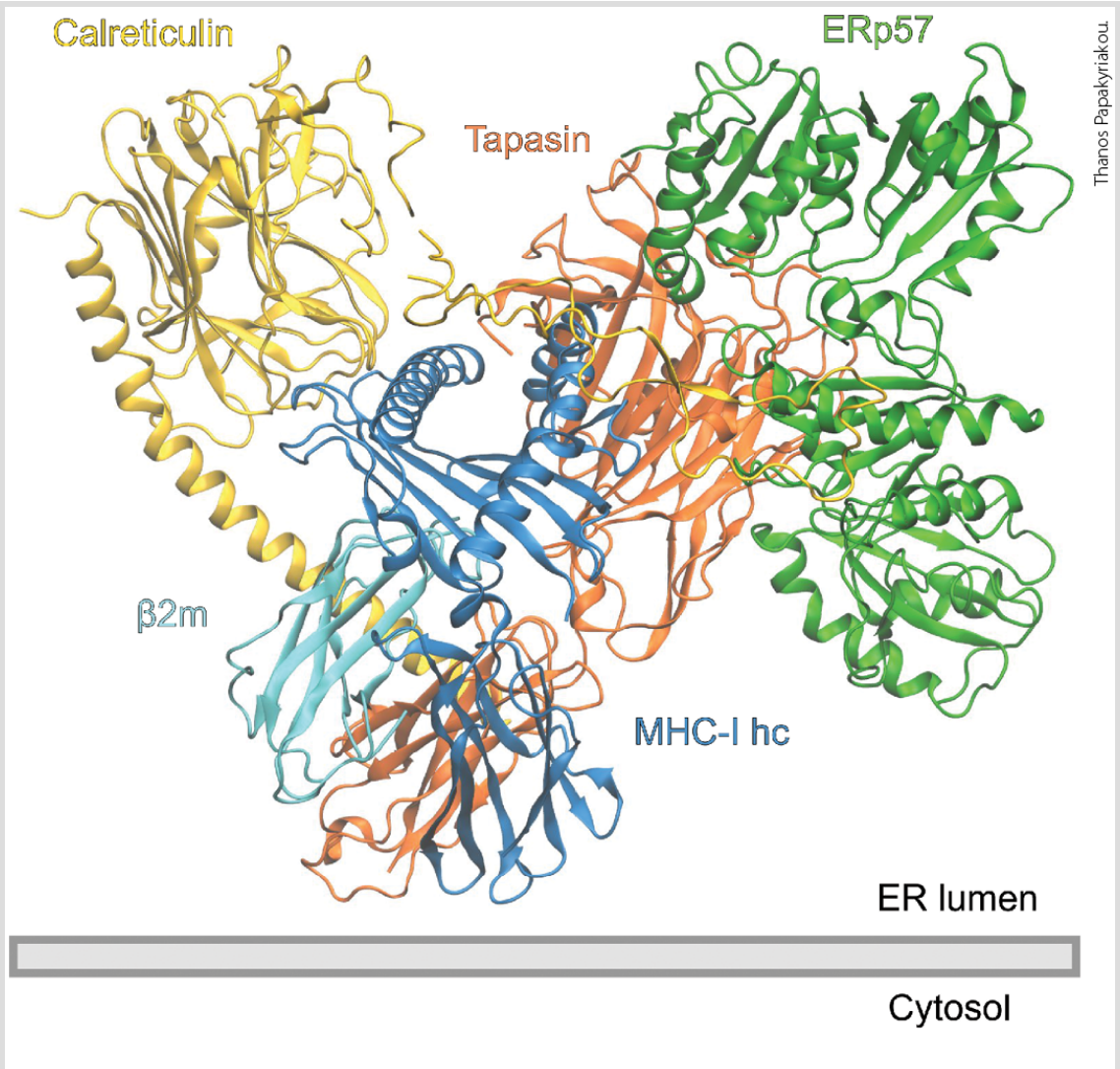


FIGURE 24-27 Class I MHC peptide-loading complex. The peptide-loading complex comprises not only the Class I MHC molecule that is about to receive a peptide, but

also the TAP-associated protein tapasin, the chaperone calreticulin, and the oxidoreductase ERp57. Tapasin maintains the Class I MHC molecule in a peptide-receptive state and also provides a physical connection with the *t*ransporter associated with *a*ntigen *p*resentation (TAP; not present in this structural diagram). Upon binding of peptide, the Class I MHC molecule is released from the peptide-loading complex and enters the secretory pathway to reach its destination, the cell surface.

6. *Display of Class I MHC–Peptide Complexes at the Cell Surface:* Once peptide loading is complete, the class I MHC–peptide complex is released from the peptide-loading complex and enters the constitutive secretory pathway (see Figure 14-2). Transfer from the Golgi to the cell surface is rapid and completes the biosynthetic pathway of a class I MHC–peptide complex.

The entire sequence of events in the class I pathway occurs constitutively in all nucleated cells, all of which express class I MHC molecules and the other required proteins or can be induced to do so. As we have seen, exposure to cytokines such as interferon γ can induce immune-specific proteasomal subunits to generate immunoproteasomes with enhanced ability to produce the appropriate peptides for presentation by class I MHC molecules. In the absence of a viral infection, protein synthesis and proteolysis continuously generate a stream of peptides that are loaded onto class I MHC molecules. Healthy, normal cells therefore display on their surfaces a representative selection of peptides derived from their own proteins. There may be several thousand distinct MHC-peptide combinations displayed at the surface of a typical nucleated

cell. The display of MHC–self-peptide complexes on the surfaces of normal, uninfected cells plays an essential role in the immune system. It is not until a virus makes its appearance that virus-derived peptides begin to make a contribution to the display of peptide-MHC complexes on cell surfaces.

As we noted above, a properly functioning immune system must be able to distinguish self (nonpathogenic) antigens from nonself (foreign, potentially pathogenic) antigens. The small organ called the *thymus* — located near the sternum at the level of the heart in humans — plays a critical role in controlling the ability of the immune system to identify self and nonself. Developing T cells in the thymus, referred to as *thymocytes*, calibrate their antigen-specific receptors to the sets of MHC-peptide complexes generated on thymic epithelial cells. The display of self peptides by self-MHC molecules in the thymus enables each developing T cell to learn which peptide-MHC combinations are self-derived and must therefore be ignored to avoid a self-destructive autoimmune reaction. T-cell development is thus driven by self-MHC molecules loaded with self peptides, a template on which a useful repertoire of T cells can be molded. Simply put, any T cell that bears a receptor that too strongly reacts with self-MHC–self-peptide complexes is potentially dangerous when it leaves the thymus, and it must be eliminated. This process of selection will be discussed below.

An exception to the usual mode of antigen presentation that is nonetheless crucial in the development of cytotoxic T cells is *cross-presentation*. This term refers to the acquisition by dendritic cells of apoptotic cell remnants, complexes composed of antigen bound to antibody, and possibly other forms of antigen, by phagocytosis. By a pathway that has yet to be fully understood, these materials escape from phagosomal or endosomal compartments into the cytosol, where they are then handled according to the steps described above, resulting in fragments loaded onto class I MHC proteins. Endosomal or lysosomal proteolysis may likewise generate peptides that find their way onto class I MHC proteins, either in the endocytic compartments themselves, or following delivery to the cytosol, where they may enter the canonical class I processing and presentation pathway. Dendritic cells are the most efficient at cross-presentation and so allow the loading of class I MHC molecules complexed with peptides that derive from cells other than the APC itself.

The Class II MHC Pathway Presents Antigens Delivered to the Endocytic Pathway

Although class I MHC and class II MHC molecules show a striking structural resemblance, the manner in which the two classes acquire peptide and their function in antigen recognition differ

greatly. Whereas the primary function of class I MHC molecules is to guide CD8-bearing cytotoxic T cells to their target (usually infected) cells, class II MHC molecules serve to guide CD4-bearing helper T cells to the cells with which they interact, primarily professional APCs. Activated helper T cells provide protection not only by helping B cells to produce antibodies, but also by means of the complex sets of cytokines they produce, which activate phagocytic cells to clear pathogens or help set up an inflammatory response.

As noted previously, class II MHC molecules are expressed primarily by professional APCs: dendritic cells and macrophages, which are phagocytic, and B cells, which are not. Hence the class II MHC pathway of antigen processing and presentation generally occurs only in these cells. The steps in this pathway are depicted in Figure 24-28.

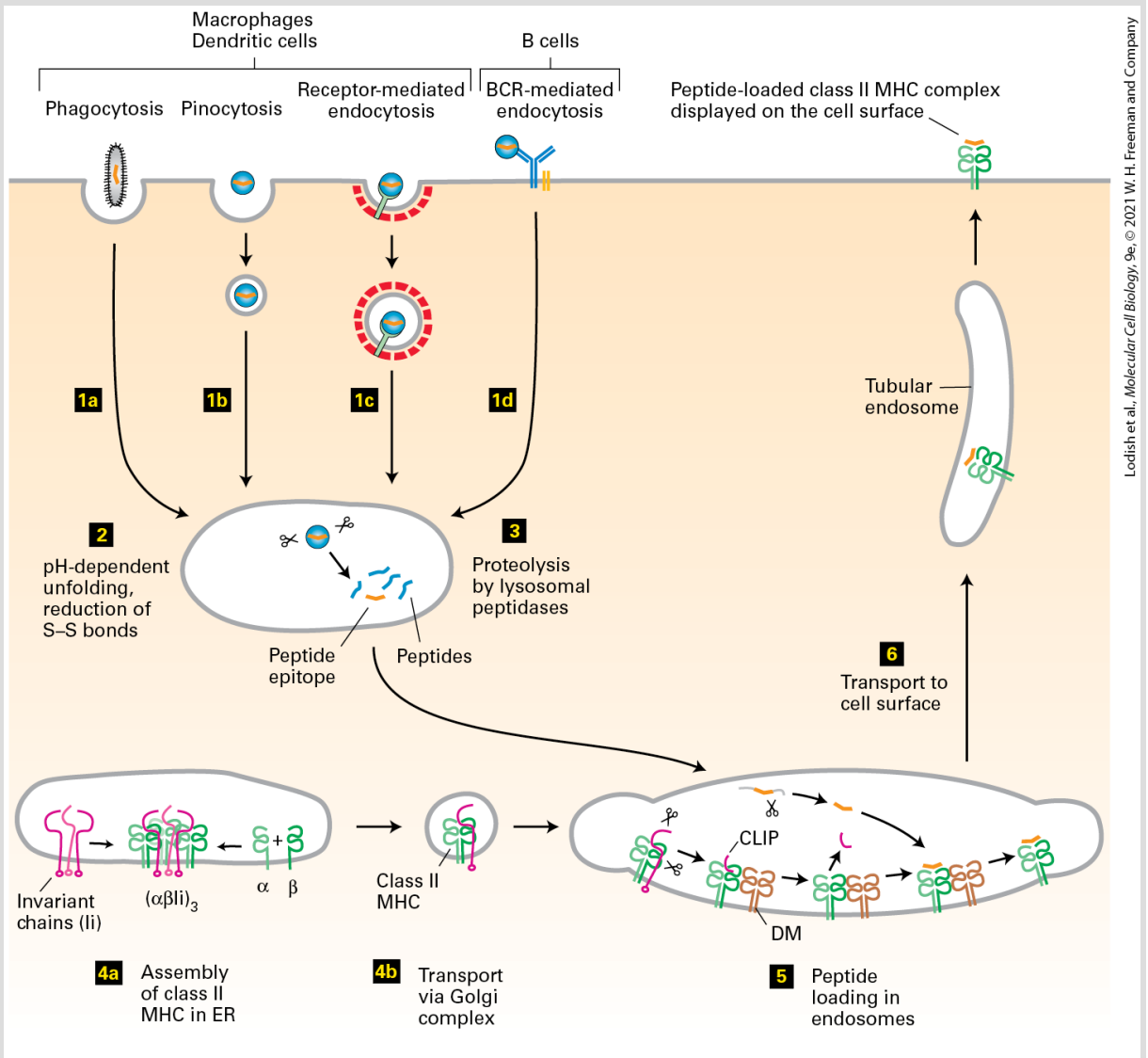
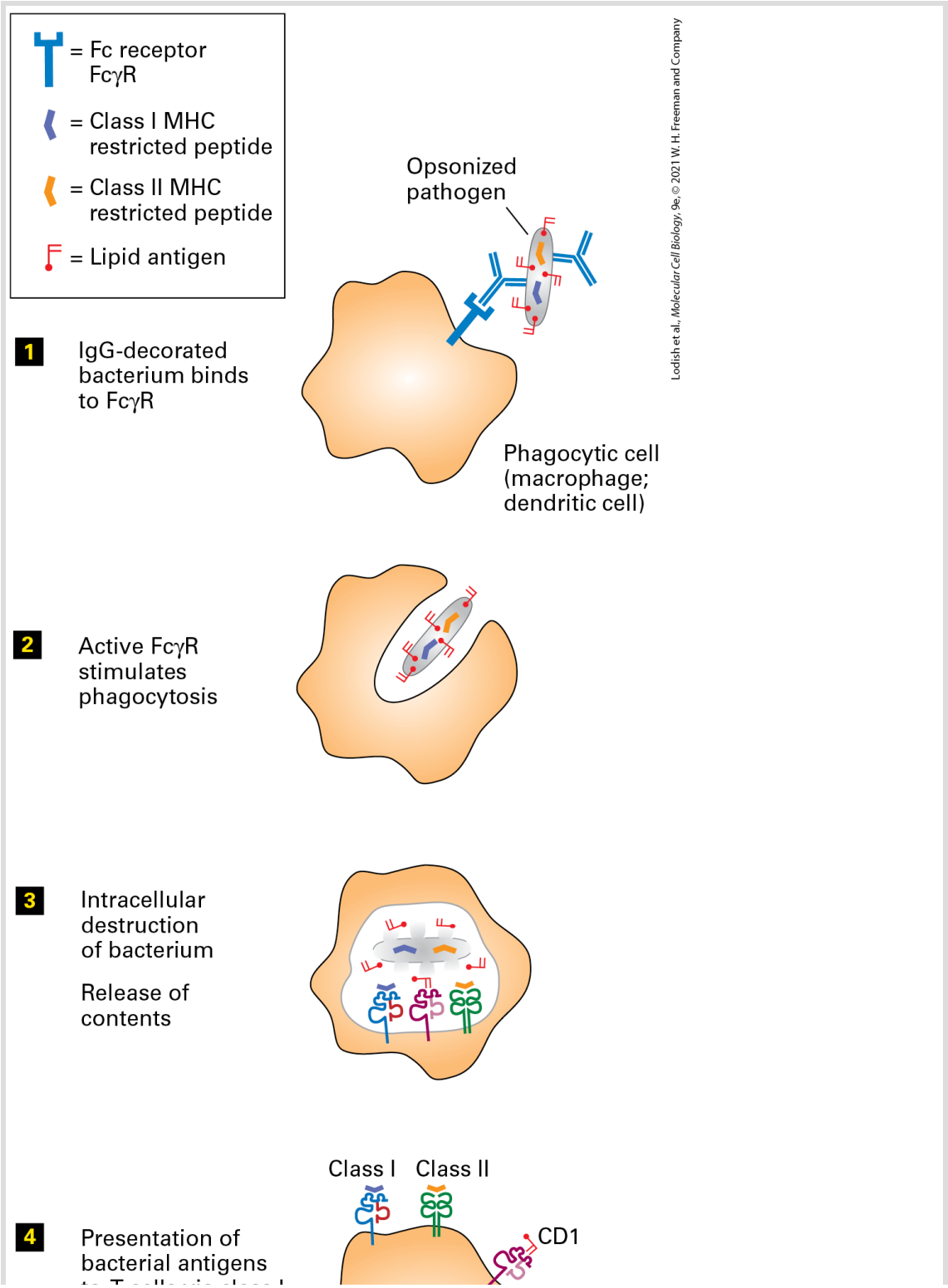


FIGURE 24-28 Class II MHC pathway of antigen processing and presentation. Step 1: Particulate antigens are acquired by phagocytosis and nonparticulate antigens by pinocytosis or endocytosis. Step 2: Exposure of antigen to the acidic and reducing environment of endosomes and lysosomes prepares the antigen for proteolysis. Step 3: The antigen is broken down by various proteases in endosomal and lysosomal compartments. Step 4: Class II MHC molecules, assembled in the ER from their subunits, are delivered to endosomal and lysosomal compartments by means of signals contained in the associated invariant (Ii) chain. This delivery targets late endosomes, lysosomes, and early endosomes, ensuring that class II MHC molecules are exposed to the products of proteolytic breakdown of antigen along the entire endocytic pathway. The Ii polypeptide is degraded in this compartment, followed by removal of Ii remnants in exchange for other peptides destined for presentation. Step 5: Peptide loading is accomplished with the

assistance of DM, a class II MHC–like chaperone protein. Step 6 : Peptide-loaded class II MHC molecules are transported to the cell surface. See text for details.

1. *Acquisition of Antigen:* In the class II MHC pathway, antigen is acquired by pinocytosis, phagocytosis, or receptor-mediated endocytosis. Pinocytosis, which is rather nonspecific, involves the delivery, by a process of membrane invagination and subsequent vesicle fusion, of a volume of extracellular fluid and the molecules dissolved therein. Phagocytosis, the ingestion of particulate materials such as bacteria, viruses, and remnants of dead cells, involves extensive remodeling of the actin-based cytoskeleton to accommodate the incoming particle. Although phagocytosis may be initiated by specific receptor-ligand interactions, these are not always required: even latex particles and other particulates such as glass beads can be ingested very efficiently by macrophages. Pathogens decorated (opsonized) by antibodies and certain complement components are targeted to macrophages and dendritic cells, which recognize them by means of cell-surface receptors for complement components or for the Fc portion of immunoglobulins. The cells then phagocytose them (Figure 24-29). Macrophages and dendritic cells also express several types of less-selective receptors (e.g., C-type lectins, Toll-like receptors, scavenger receptors) that recognize molecular patterns in both soluble and particulate antigens. These cells then internalize the bound antigens by receptor-mediated endocytosis. B cells, which are not phagocytic, can also acquire antigens by receptor-mediated endocytosis using their antigen-specific B-cell receptors (Figure

24-30). Finally, cytosolic antigens may enter the class II MHC pathway via autophagy (see Figure 14-34).



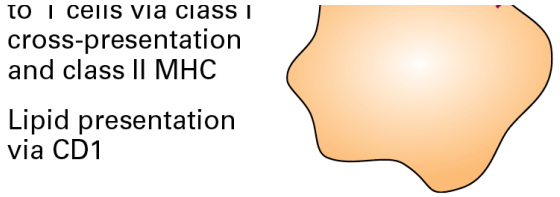


FIGURE 24-29 Presentation of opsonized antigen by phagocytic cells. By means of Fc receptors such as FcγR displayed on their cell surface, specialized phagocytic cells such as macrophages or dendritic cells can bind and ingest pathogens that have been decorated with antibodies (opsonization). After digestion of the phagocytosed particle (e.g., immune complex, bacterium, virus), some of the peptides produced, including fragments of the pathogen (orange), are loaded onto class II MHC molecules (green). Class II MHC–peptide complexes displayed at the surface allow activation of T cells whose receptors are specific for these MHC-peptide combinations. Lipid antigens are delivered to the class I MHC–like molecule CD1 (pink), whose binding site is specialized to accommodate lipids. Certain pathogen-derived peptides (purple) may be delivered to class I MHC products (blue) by means of cross-presentation. The molecular mechanisms that underlie cross-presentation remain to be clarified.

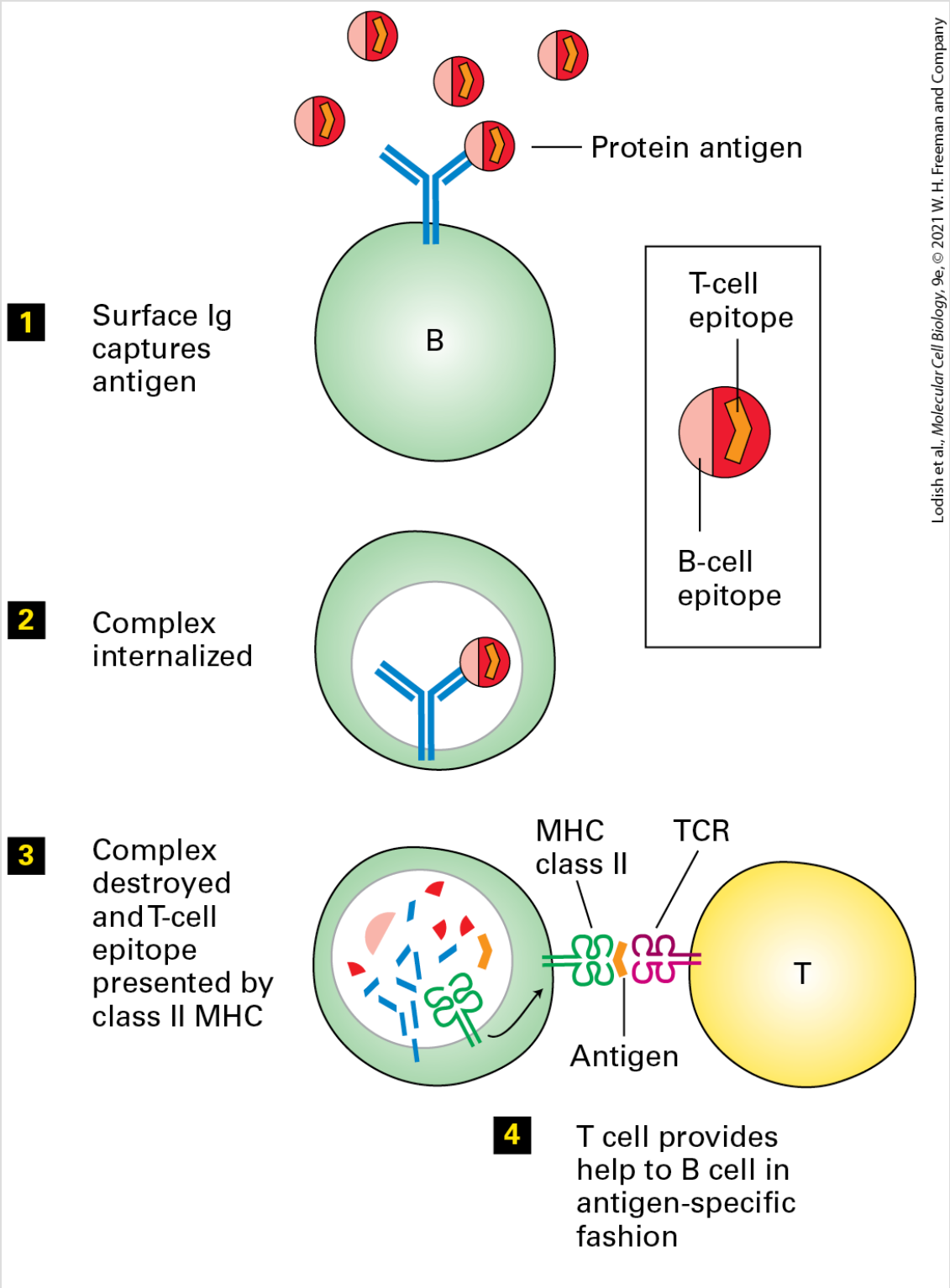


FIGURE 24-30 Antigen presentation by B cells. B cells bind antigen, even if present at low concentration, to their B-cell receptors, or surface Ig. The immune complex that results is internalized and then delivered to endosomal or lysosomal compartments,

where it is degraded. Peptides liberated from the immune complex, including fragments of the protein antigen, are displayed as class II MHC–peptide complexes at the cell surface. Helper T cells specific for the displayed complex can now provide help to the B cell, allowing it to proliferate and eventually form antibody-secreting plasma cells. This help is MHC restricted and antigen specific.

2. *Targeting Antigen for Destruction:* Proteolysis is required to convert intact protein antigens into peptides of a size suitable for binding to class II MHC molecules. Protein antigens are targeted for degradation by progressive unfolding, brought about by the drop in pH as proteins progress along the endocytic pathway. The pH of the extracellular environment is around pH 7.2, while in early endosomes it is between pH 6.5 and 5.5; in late endosomes and lysosomes the pH may drop to pH 4.5. ATP-powered V-class proton pumps in the endosomal and lysosomal membranes are responsible for this acidification (see Figure 11-9). Proteins that are stable at neutral pH tend to unfold when they are exposed to extremes of pH through rupture of hydrogen bonds and destabilization of salt bridges. Furthermore, the environment in the endosomal or lysosomal compartment is a reducing one, in which lysosomes attain a concentration of reducing equivalents in the millimolar range. Reduction of the disulfide bonds that stabilize many extracellular proteins can also be catalyzed by a thioreductase inducible by exposure to interferon γ. The combined action of low pH and reducing environment prepares the antigens for proteolysis.

3. *Proteolysis:* Degradation of proteins in the class II MHC pathway is carried out by a large set of lysosomal proteases, collectively referred to as *cathepsins*, which are either cysteine or aspartyl proteases. A wide range of peptide fragments is produced, including some that can bind to class II MHC molecules. The lysosomal proteases operate optimally at the acidic pH within lysosomes. Consequently, agents that inhibit the activity of the V-class proton pumps that maintain their acidification interfere with antigen processing, as do inhibitors of lysosomal proteases.
4. *Encounter of Peptides with Class II MHC Molecules:* Recall that most class II MHC molecules synthesized in the endoplasmic reticulum are directed to late endosomes. The peptides generated by proteolysis thus reside in the same intracellular space as the class II MHC molecules themselves — they do not have to cross a membrane, as is the case for peptides destined to bind to class I MHC molecules (see Figure 24-26). To allow peptides and class II MHC molecules to meet, the $(\alpha\beta Ii)_3$ complex is transported from the endoplasmic reticulum via the secretory pathway to endosomal compartments.
5. *Binding of Peptides to Class II Molecules:* The $(\alpha\beta Ii)_3$ complex delivered to endosomal compartments is incapable of binding peptide because the peptide-binding cleft in the class II molecule is occupied by the invariant chain (Ii). For the same reason, newly assembled $(\alpha\beta Ii)_3$ complexes do not compete for class I MHC–destined peptides delivered to the ER via TAP: their peptide-binding site is already occupied by Ii. Recall that the ER is where both class I and class II MHC molecules are assembled.

The presence of Ii in the nascent class II MHC complex ensures that class II MHC molecules do not bind peptide in the ER. The same proteases in endosomes and lysosomes that act on internalized antigens and degrade them into peptides also act on the $(\alpha\beta Ii)_3$ complex, resulting in removal of the Ii molecule from the complex, with the exception of a small portion of Ii called the *CLIP segment*. Because it is firmly lodged in the class II MHC peptide-binding cleft, CLIP is resistant to proteolytic attack. The class II MHC molecules themselves are also resistant to unfolding and proteolytic attack under the harsh conditions that prevail in the endocytic pathway. Next the CLIP segment is removed from the αβ heterodimer by the chaperone DM, and the newly vacated peptide-binding cleft of the class II MHC molecule may now bind the peptides that are abundantly present in the endocytic pathway. Although the DM protein is MHC encoded and structurally very similar to class II molecules, it does not itself bind peptides. However, newly formed class II MHC–peptide complexes are themselves susceptible to further editing by DM, which may dislodge the peptide already bound, until the class II molecule acquires a peptide that binds so strongly that it cannot be removed by DM. The resulting class II MHC–peptide complexes are extremely stable, with estimated half-lives far in excess of 24 hours.

6. *Display of Class II MHC–Peptide Complexes at the Cell Surface:* The newly generated class II MHC–peptide complexes are localized mostly in late endosomal compartments, which include multivesicular endosomes (or bodies) (see Figure 14-32). Endosomal compartments then elongate and ultimately deliver

class II MHC–peptide complexes to the surface by membrane fusion. These events are tightly regulated: tubulation and delivery of class II MHC molecules to the surface are enhanced in dendritic cells and macrophages following their activation by signals generated in response to infection, such as bacterial lipopolysaccharide, which is detected by Toll-like receptors on the surfaces of these professional APCs, as well as inflammatory cytokines, such as interferon γ, produced by CD4-expressing helper T cells.

For professional APCs, the above steps are constitutive — they happen all the time — but they can be modulated by exposure to microbial agents and cytokines. In addition to the pathways described here for class I and class II MHC products, there is a category of class I MHC–related molecules, the CD1 proteins, that are specialized in the presentation of lipid antigens. The structure of a CD1 molecule resembles that of a class I MHC molecule: a larger subunit complexed with β2-microglobulin. Many species of bacteria produce lipids whose chemical structures are not found in their mammalian hosts. These lipids can serve as antigens when presented by CD1 molecules (see Figure 24-29), to which they bind via a lipid-binding pocket that is conceptually similar to that of most MHC molecules. Signals in the cytoplasmic tail of the larger CD1 subunit target these molecules to endosomal or lysosomal compartments, where loading with antigenic lipids occurs. The CD1-lipid complexes engage a relatively rare class of T cells, referred to as *NKT cells*, as well as *γδ T cells*, both described below. NKT cells

fulfill an important role in cytokine production and help initiate and orchestrate adaptive immune responses via their cytokine outputs. Other proteins structurally related to Class I MHC products such as the MR1 protein can bind small metabolites and present them to a special subset of T lymphocytes called *MAIT cells*.

KEY CONCEPTS OF SECTION 24.4

The MHC and Antigen Presentation

- The MHC, discovered as the genetic region responsible for acceptance or rejection of grafts, encodes many different proteins involved in the immune response. Two of these proteins, class I and class II MHC molecules, are highly polymorphic, occurring in many allelic variations (see Figure 24-22).
- The function of the class I and class II MHC proteins is to bind peptide antigens and display them on the surfaces of cells so that the antigen–MHC protein complex can interact with antigen-specific T-cell receptors on T cells. When an antigen–MHC protein complex on an antigen-presenting cell binds to its complementary T-cell receptor on a T cell, the T cell is activated to assume effector functions, such as the production of cytokines or the ability to kill a virus-infected cell. Class I MHC molecules are found on most nucleated cells, whereas the expression of class II MHC molecules is confined largely to professional APCs such as dendritic cells, macrophages, and B cells.
- The organization and structure of class I and class II MHC molecules is similar and includes a peptide-binding cleft that is specialized for binding a wide variety of peptides (see Figure 24-24).
- Different allelic variants of MHC molecules bind different sets of peptides because the differences that distinguish one allele from another include residues that define the architecture of the peptide-binding cleft (see Figure 24-25). Allelic variation also includes residues in the MHC molecule that directly contact the corresponding T-cell receptor. Thus different allelic variants of an MHC molecule, even if they bind the identical peptide, do not usually react with the same T-cell receptor. This phenomenon is called *MHC restriction*.
- Peptide binding by class I and class II MHC molecules occurs in different intracellular compartments: class I molecules bind predominantly to cytosolic materials, whereas

class II molecules bind to extracellular materials internalized by phagocytosis, pinocytosis, or receptor-mediated endocytosis.

- The process by which protein antigens are acquired, processed into peptides, and converted into surface-displayed MHC-peptide complexes is referred to as *antigen processing and presentation*. This process operates continuously in cells that express the relevant MHC molecules, yet can be modulated in the course of an immune response.
- Antigen processing and presentation can be divided into six discrete steps: (1) acquisition of antigen; (2) targeting of the antigen for destruction; (3) proteolysis; (4) encounter of peptides with MHC molecules; (5) binding of peptides to MHC molecules; and (6) display of the peptide-loaded MHC molecules on the cell surface (see Figures 24-25 and 24-27).

24.5 T Cells, T-Cell Receptors, and T-Cell Development

T lymphocytes recognize antigen through specific interactions of their receptors with MHC molecules. The diverse, antigen-specific T-cell receptors made by individual T cells that are entrusted with this task are structurally related to the F(ab) portion of immunoglobulins. To generate a large repertoire of antigen-specific T-cell receptors, T cells rearrange the genes encoding the T-cell receptor subunits by mechanisms of somatic recombination essentially identical to those used by B cells to rearrange immunoglobulin genes. The development of T cells, like that of B cells, is strictly dependent on successful completion of these somatic gene rearrangements to yield a functional T-cell receptor. Genetic ablation of RAG activity eliminates all of the T cells as well as all of the antibody-producing B cells. As noted previously, each T cell expresses a single type of T-cell receptor. In this section, we describe the receptor subunits that mediate antigen-specific recognition, how they pair up with membrane glycoproteins essential for signal transduction, and how these complexes recognize MHC-peptide combinations.

As pointed out in the preceding section, an individual's T cells recognize peptide antigens only when they are bound to the polymorphic MHC molecules present in that individual. In the course of T-cell development, T cells must learn the identity of these

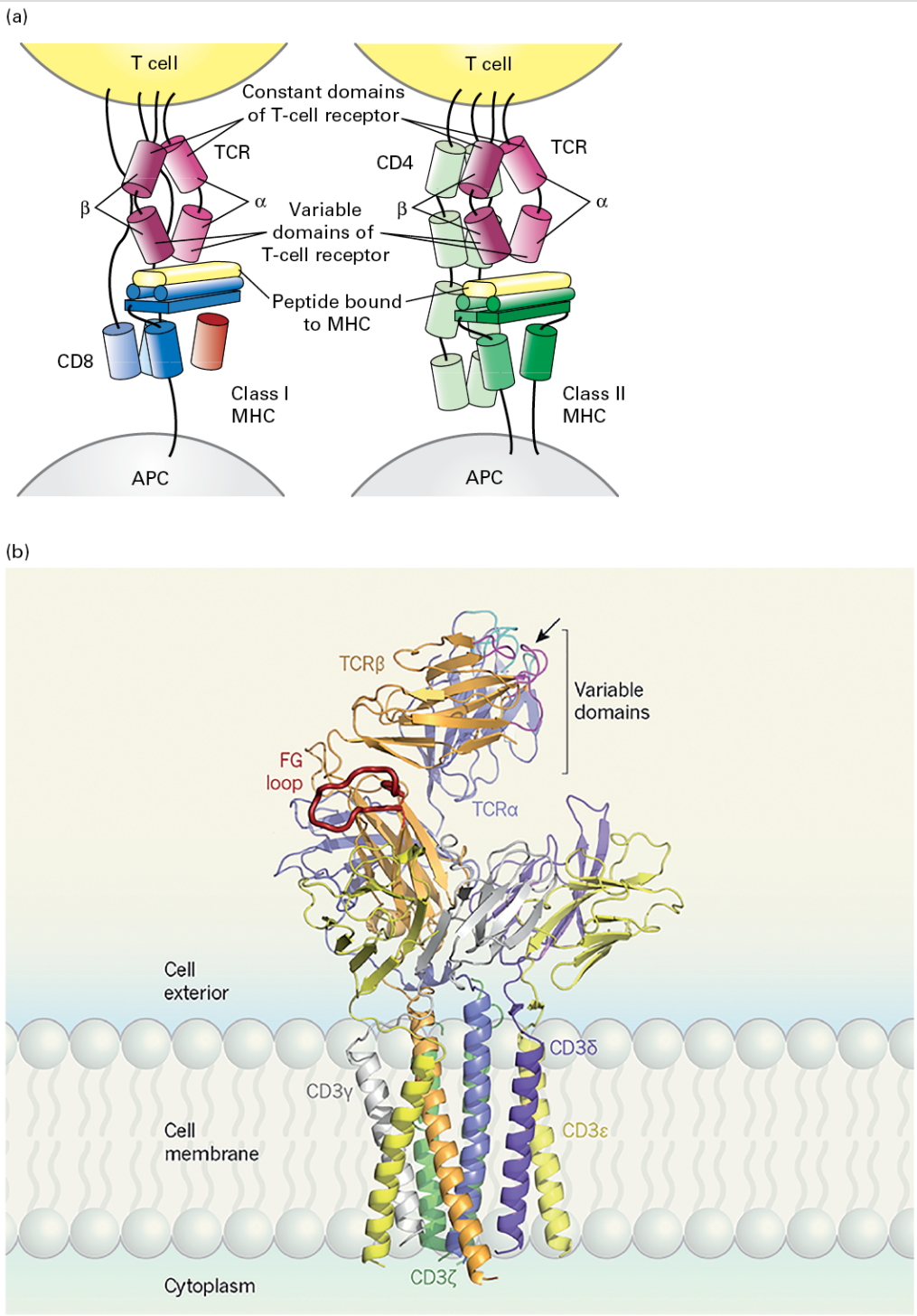
self MHC molecules and receive instructions about which MHC-peptide combinations to ignore, so as to avoid autoimmunity: the potentially catastrophic reactions of newly generated T cells with the individual's own tissues.

The Structure of the T-Cell Receptor Resembles the F(ab) Portion of an Immunoglobulin

Much as B cells use the B-cell receptors on their surfaces to recognize antigens and generate intracellular signals that lead to clonal expansion, T cells depend on their **T-cell receptors (TCRs)** to initiate their participation in immune responses. T cells that have been activated via these antigen-specific receptors proliferate and acquire the capacity to kill antigen-bearing target cells (in the case of cytotoxic T cells) or to secrete cytokines that will assist B cells in their differentiation into antibody-producing plasma cells (in the case of helper T cells). The TCR recognizes antigenic peptides bound to MHC molecules.

The TCR is composed of two core glycoprotein subunits (Figure 24-31), each of which is encoded by a somatically rearranged gene. The receptor is usually composed of an α and a β subunit, but certain TCRs are alternatively composed of a γ and a δ subunit. The structure of these subunits is similar to that of the F(ab) portion of

an immunoglobulin: at the N-terminal end is a variable region, followed by a constant region and a transmembrane segment. The cytoplasmic tails of the TCR subunits are short and do not directly interact with cytoplasmic signal transduction molecules. Instead, the TCR associates with the CD3 complex, a set of membrane glycoproteins composed of γ, δ, ε, and ζ chains. The structure of this complex has been determined by cryoelectron microscopy (Figure 24-31b). The extracellular domains of the CD3 subunits are similar to immunoglobulin domains, and the cytoplasmic domain in each contains an ITAM (immunoreceptor tyrosine-based activation motif), which recruits adapter proteins upon phosphorylation of its tyrosine residues. The ζ chain is integrated into the CD3-TCR complex as a disulfide-bonded homodimer, and each ζ chain contains three ITAMs (see Figure 24-33).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 24-31 Structure of the T-cell receptor and its co-receptors. (a) The antigen-specific T-cell receptor (TCR) is composed of two chains, the α and β subunits, which are produced by V-J and V-D-J recombination, respectively (see Figure 24-32). The α and β subunits must associate with the CD3 complex (not

shown; see panel b and Figure 24-33) to allow the transduction of signals. The formation of a full TCRαβ–CD3 complex is required for surface expression. The TCR further associates with a co-receptor, CD8 (light blue) or CD4 (light green), which allows interaction with conserved features of class I MHC or class II MHC molecules, respectively, on antigen-presenting cells. (b) Structure of the extracellular portion of the TCR-CD3 complex as determined by cryoelectron microscopy. The subunits of the CD3 complex are color-coded. The extracellular portions of the T-cell receptor α(TCRα) and β(TCRβ) subunits as well as those of CD3γ,δ, and ε all possess the basic fold of immunoglobulin domains. The cytoplasmic portions of the CD3 complex subunits were not resolved by cryoelectron microscopy and are therefore absent from the figure. Unlike the CD3γ, δ, and ε subunits, the CD3ζ chain has a minimal extracellular domain. The intracellular portions of the CD3 complex contain the immunoreceptor tyrosine-based activation motifs (ITAMs) essential for signal transduction via the TCR-CD3 complex.

[Part (b) republished with permission from Nature Publishing Group, from E. L. Reinherz, 2019, "The Structure of a T-Cell Mechanosensor," *Nature* **573**(7775):502–504; permission conveyed through Copyright Clearance Center, Inc.]

TCR Genes Are Rearranged in a Manner Similar to Immunoglobulin Genes

Virtually all antigen-specific receptors generated by somatic recombination contain a subunit that is the product of V-D-J recombination (e.g., Ig heavy chain; TCR β chain) and another that is the product of V-J recombination (e.g., Ig light chain; TCR α chain). The mechanism of V-D-J and V-J recombination for TCRs is essentially identical to that described for immunoglobulin genes

and requires all the component proteins composing the nonhomologous end-joining machinery: RAG1, RAG2, Ku70, Ku80, the catalytic subunit of DNA-dependent protein kinase, XRCC4, DNA ligase IV, and Artemis. Similarly *r*ecombination *s*ignal *s*equences (RSSs) are required (Figure 24-32).

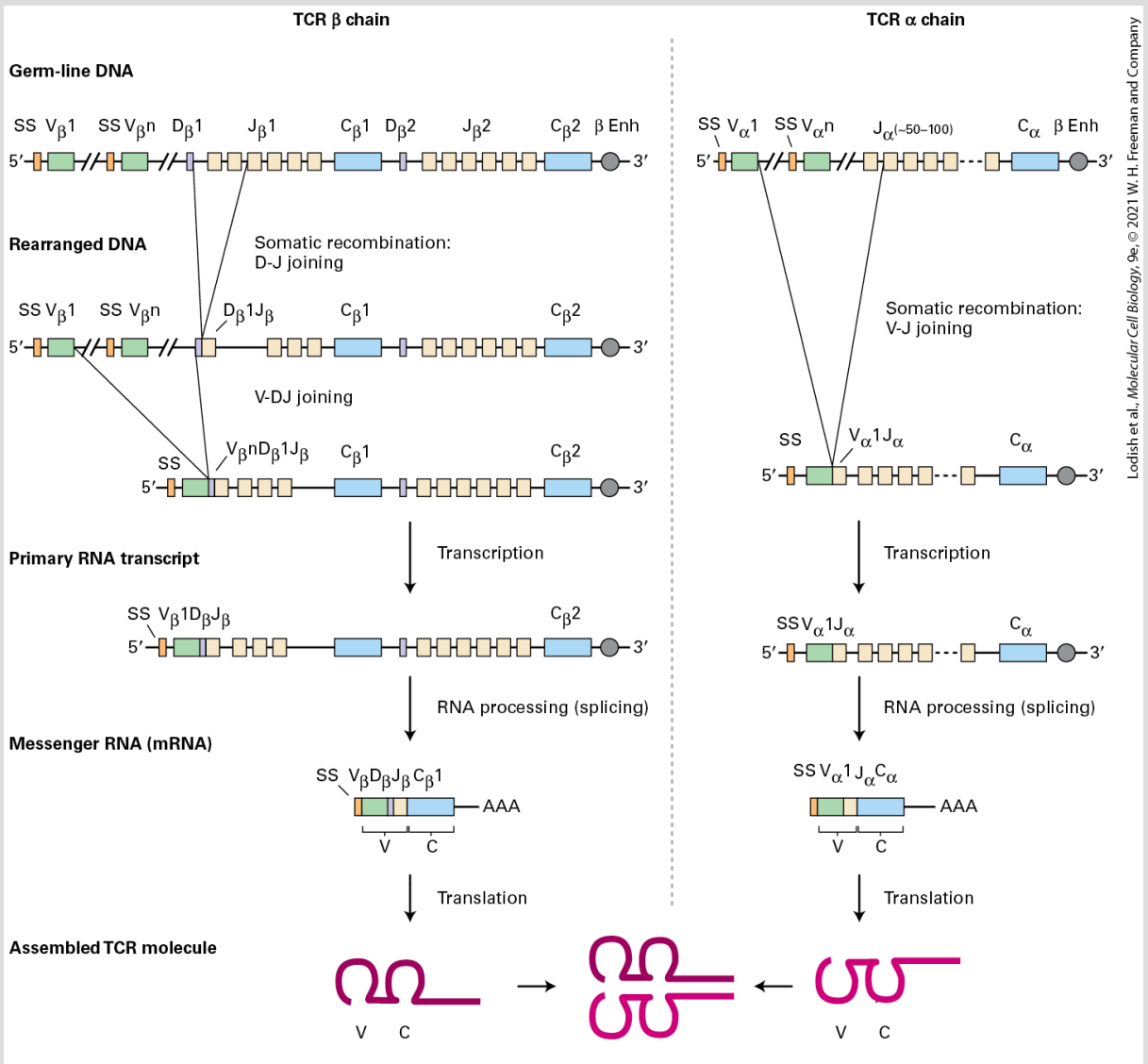


FIGURE 24-32 Organization and recombination of TCR loci. The organization of TCR loci is in principle similar to that of immunoglobulin loci (see Figure 24-15). *Left:* The TCR β-chain locus includes a cluster of V segments, a cluster of D segments, and several J segments,

downstream of which are two constant regions. The arrangement of the recombination signals is such that not only is D-J joining allowed, but also D-D joining followed by joining to a V segment. Direct V-J joining in the TCR β locus is not observed even though the arrangement of the recombination signal sequences would be consistent with such a possibility. *Right:* The TCR α-chain locus is composed of a cluster of V segments and a large number of J segments. SS = exon encoding signal sequence; Enh = enhancer.

A number of noteworthy features characterize the organization and rearrangement of the TCR loci. First, the organization of the RSSs is such that D-to-D rearrangements are allowed, unlike the case for Ig. Second, terminal deoxynucleotidyltransferase (TdT) is active at the time the TCR genes are rearranged, and therefore additional nucleotides, termed *N nucleotides*, can be present at the V-D and D-J junctions in all rearranged TCR genes. Third, in humans and mice, the TCR δ locus is embedded within the TCR α locus. This organization results in complete excision of the interposed δ locus when TCR α rearrangement occurs, so a choice of the TCR α locus for rearrangement precludes use of the δ locus, which is lost by deletion. T cells that express the αβ receptor and those that express the γδ receptor are considered separate lineages with distinct functions. Among the T-cell γδ receptors are some capable of recognizing the CD1 molecule, which as we learned is specialized for the presentation of lipid antigens. The γδ T cells are programmed to home to distinct anatomic sites (e.g., the epithelium lining, the genital tract, or the skin) and probably play a role in host defense against pathogens commonly found at these sites. They may also have anti-tumor activity.

Deficiencies in the key components of the recombination apparatus, such as the RAG recombinases, preclude rearrangement of TCR genes. As we have seen for B cells, development of lymphocytes is strictly dependent on the rearrangement of the antigen-receptor genes. A deficiency in either RAG1 or RAG2 thus prevents both B-cell and T-cell development. Mice with homozygous RAG gene knockouts are frequently used to assess the roles of B and T cells in physiological and pathophysiological processes.

Many of the Variable Residues of TCRs Are Encoded in the Junctions Between V, D, and J Gene Segments

The diversity created by somatic recombination of TCR genes is estimated to exceed 10^{10} unique receptors. Combinatorial use of different V, D, and J gene segments makes an important contribution to this diversity, as do the mechanisms of junctional imprecision and N-nucleotide addition already discussed for immunoglobulin gene rearrangements. The net result is a degree of variability in the V regions that at least matches that of the immunoglobulins (see Figure 24-13). Indeed, each of the TCR's variable regions includes three hypervariable regions (CDRs), equivalent to those in the BCR. Unlike immunoglobulin genes, however, the TCR genes do not undergo somatic hypermutation. Therefore, TCRs exhibit nothing equivalent to the affinity

maturation of antibodies during the course of an immune response, nor is there the option of class-switch recombination or the use of alternative polyadenylation sites to create soluble and membrane-bound versions of the receptors.

The crystal structures of a number of TCRs bound to class I MHC–peptide or class II MHC–peptide complexes have been determined. These structures show variation in how the TCR docks with the MHC-peptide complex, but the most extensive contacts in the somatically diverse CDR3 region are made with the central peptide-containing portion of the complex, with the germ line–encoded CDR1 and CDR2 contacting the α helices of the MHC molecules. Many of the TCRs for which a structure has been solved dock diagonally across the peptide-binding portion of the MHC-peptide complex. As a result, the TCR makes extensive contacts with the peptide as well as with the α helices of the MHC molecule to which it binds. The positions at which allelic MHC molecules differ from one another are frequently those residues that directly contact the TCR, thus precluding tight binding of unrelated allelic MHC products.

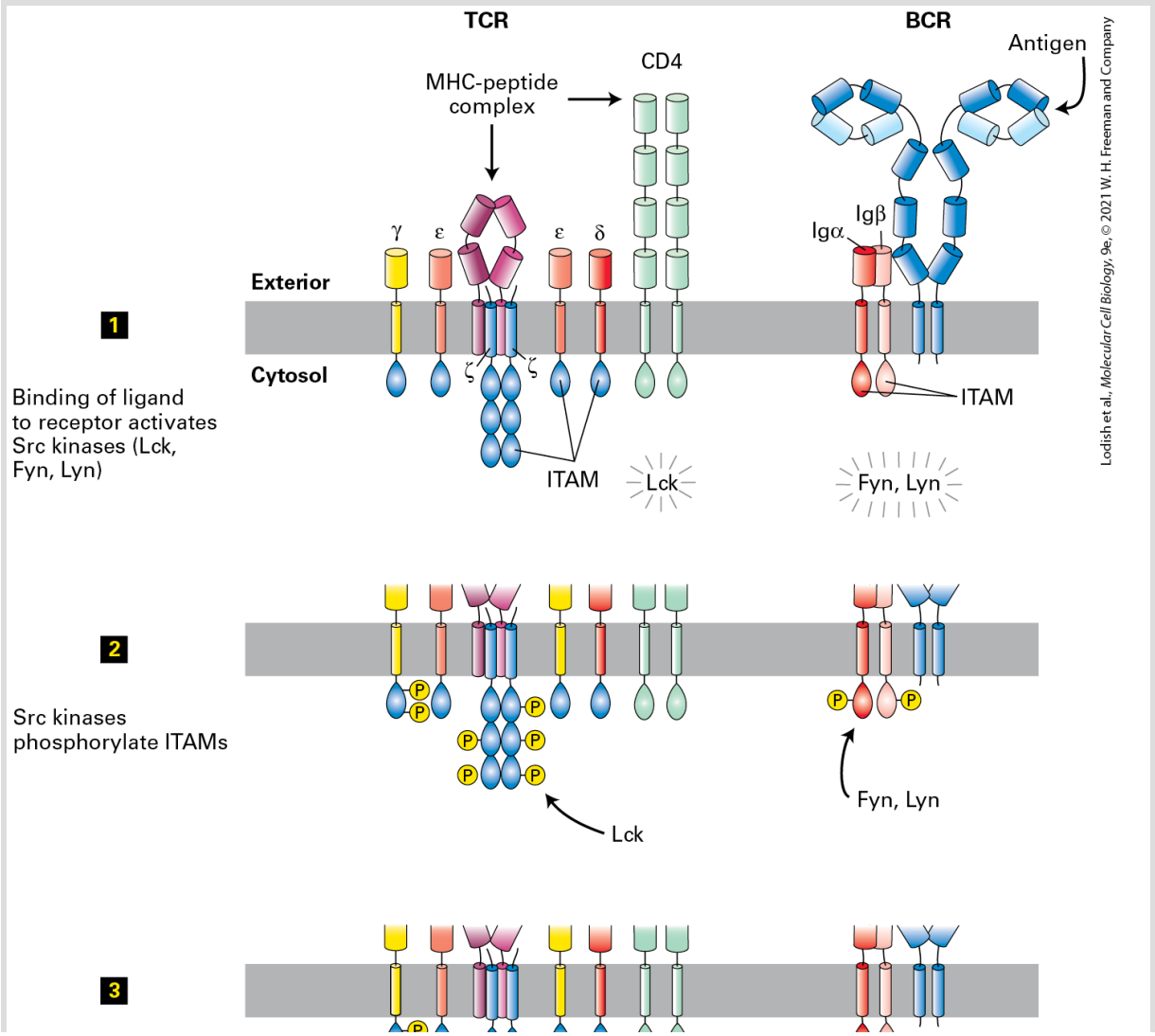
Amino acid differences that distinguish one MHC allele from another also affect the architecture of the peptide-binding cleft. Even if the MHC residues that interact directly with the TCR were shared by two allelic MHC molecules, their peptide-binding specificity would probably differ because of amino acid differences in the peptide-binding cleft. Consequently, the TCR contact residues provided by bound peptide, which are essential for stable

interaction with a TCR, would be absent from the wrong MHC-peptide combination. A productive interaction with the TCR would then be unlikely to occur. As mentioned earlier, these molecular features are the underpinnings of MHC restriction.

Signaling via Antigen-Specific Receptors Triggers Proliferation and Differentiation of T and B Cells

The immune responses mediated by T cells and B cells are initiated once TCRs or BCRs bind their respective ligands and are thus activated. The ligands for TCRs are MHC-peptide complexes expressed on the surfaces of APCs. The ligands for BCRs are antigens that bind to the receptors without the need for MHC intervention or a presenting cell. TCR and BCR activation by their antigens is similar to the activation of signaling receptors like tyrosine kinase receptors that we have considered previously (see Chapter 16). In all cases, multiple phosphotyrosine signal transduction cascades are engaged. Several integral membrane proteins, as well as soluble cytosolic enzymes, participate in TCR and BCR signaling. In some cases, these membrane-associated proteins can be thought of as auxiliary subunits of the receptors. Examples of how such auxiliary proteins participate in signaling are shown in Figure 24-33. The cytosolic portions of the antigen-specific receptors do not protrude much beyond the cytosolic leaflet of the

plasma membrane, and are incapable of recruiting downstream signaling molecules. Instead, the antigen-specific receptors associate with auxiliary subunits that contain ITAMs. Engagement of the antigen-specific receptors by ligand initiates a series of receptor-proximal events: kinase activation, phosphorylation of one or two tyrosine residues in the ITAM sequence, and subsequent binding to these phosphotyrosines of SH2 domains of adapter proteins that serve as scaffolds for recruitment of other downstream signaling molecules.



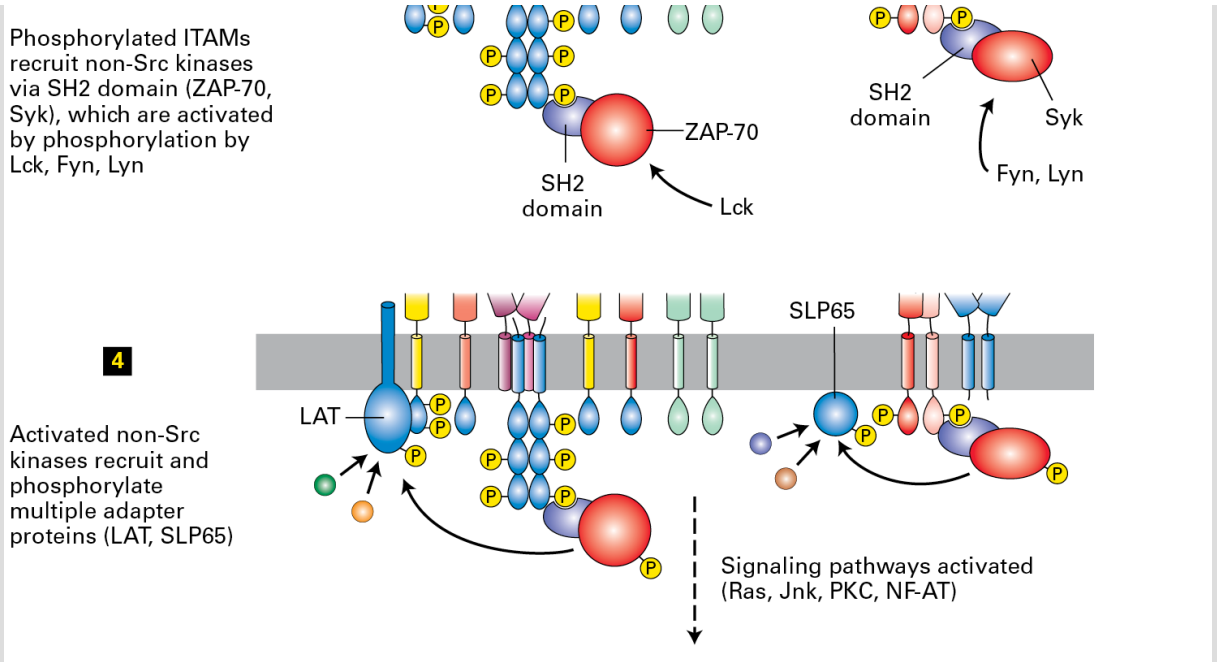


FIGURE 24-33 Signal transduction from the T-cell receptor (TCR) and B-cell receptor (BCR). The signal transduction pathways used by the antigen-specific receptors of T cells (*left*) and B cells (*right*) are conceptually similar. The initial stages are depicted in this figure; downstream signaling events lead to changes in gene expression that result in proliferation and differentiation of the antigen-stimulated lymphocytes. See text for further discussion.

One major downstream signal transduction pathway is outlined in Figure 24-33. Engagement of the antigen-specific receptors by ligand activates Src-family cytosolic tyrosine kinases (e.g., Lck in helper T cells; Lyn and Fyn in B cells; Figure 24-33, step **1**); these kinases are found in close proximity to or physically associated with the antigen-specific receptors. The active kinases phosphorylate tyrosine residues in the ITAMs (step **2**). In their phosphorylated forms, the ITAMs bind to SH2 domains of non-Src-family tyrosine kinases (ZAP-70 in T cells, Syk in B cells), leading to activation of these kinases (step **3**). The ITAMs bind other adapter proteins, which are activated by the non-Src-family tyrosine kinases. As with

other phosphotyrosine-triggered signaling pathways (see Chapter 16) this leads to activation of phosphoinositide-specific phospholipase Cγ and several PI-3 kinases. Subsequent downstream events parallel those described in Chapter 16 for signaling from receptor tyrosine kinases. Antigen-specific receptors on B and T cells are perhaps best characterized as modular receptor tyrosine kinases, with the ligand recognition units and kinase domains carried by separate molecules. Ultimately, signaling via these T cell and B cell antigen-specific receptors initiates transcription programs that determine the fate of the activated lymphocyte: proliferation and differentiation.

T cells depend critically on the cytokine interleukin 2 (IL-2) for clonal expansion. Following antigen stimulation of a T cell, one of the first genes to be turned on is that for IL-2 (see Figure 16-19). The T cell responds to its own initial burst of IL-2 and proceeds to make more IL-2, an example of autocrine stimulation and part of a positive feedback loop. An important transcription factor required for the induction of IL-2 synthesis is the NF-AT protein (*n*uclear *f*actor of *a*ctivated *T* cells). This protein is sequestered in the cytoplasm in phosphorylated form and cannot enter the nucleus unless it is de-phosphorylated first. The phosphatase responsible is calcineurin, a Ca^{2+}-activated enzyme. The initial rise in cytosolic Ca^{2+} leading to activation of calcineurin results from mobilization of ER-resident Ca^{2+} stores triggered by hydrolysis of $PI(4,5)P_2$ and the concomitant generation of IP_3 (see Figure 15-28, steps **2** – **4**).

The immunosuppressant drug cyclosporine inhibits calcineurin activity through formation of a cyclosporine-cyclophilin complex, which binds and inhibits calcineurin. If de-phosphorylation of NF-AT is suppressed, NF-AT cannot enter the nucleus and induce transcription of the *IL-2* gene. This precludes clonal expansion of antigen-stimulated T cells and so leads to immunosuppression, arguably the single most important intervention that contributes to the success of organ transplantation involving unrelated donors and recipients (individuals who are genetically different and therefore express different MHC products), referred to as *allogeneic* tissue transplantation. Although the success of transplantation varies with the organ used, the availability of strong immunosuppressants such as cyclosporine has expanded the possibilities of clinical transplantation enormously. ■

T Cells Capable of Recognizing MHC Molecules Develop Through a Process of Positive and Negative Selection

Rearrangement of the gene segments that are assembled to encode a functional T-cell receptor is a stochastic event, completed on the part of the T cell without any prior knowledge of the MHC molecules with which its receptors must ultimately interact. As in somatic recombination of Ig heavy-chain loci in B cells, the first gene segments to be rearranged in the TCR β chain are the D and J

elements; a V segment is then joined to the newly recombined DJ (see Figure 24-32). At this stage of T-cell development, productive rearrangement allows the synthesis of the TCR β chain, which is incorporated into the pre-TCR through association with a surrogate α subunit termed *pre-Tα*. This dimeric pre-TCR fulfills a function strictly analogous to that of the pre-BCR in B-cell development: it tells the T cell that it has successfully completed a productive β chain rearrangement, with no need for further rearrangements in the β chain gene on the homologous chromosome. The pre-TCR allows clonal expansion of the pre-T cells that successfully underwent TCR β chain rearrangement, and it imposes allelic exclusion to ensure that, as a rule, a single functional TCR β subunit is generated for a given T cell and its descendants. RAG expression subsides until the expansion phase of the pre-T cells is complete, after which it is re-initiated to allow rearrangement of the TCR α locus, ultimately leading to the generation of T cells with a fully assembled TCR containing a rearranged α and β subunit. Figure 24-34 illustrates the analogous steps in the development of T and B cells.

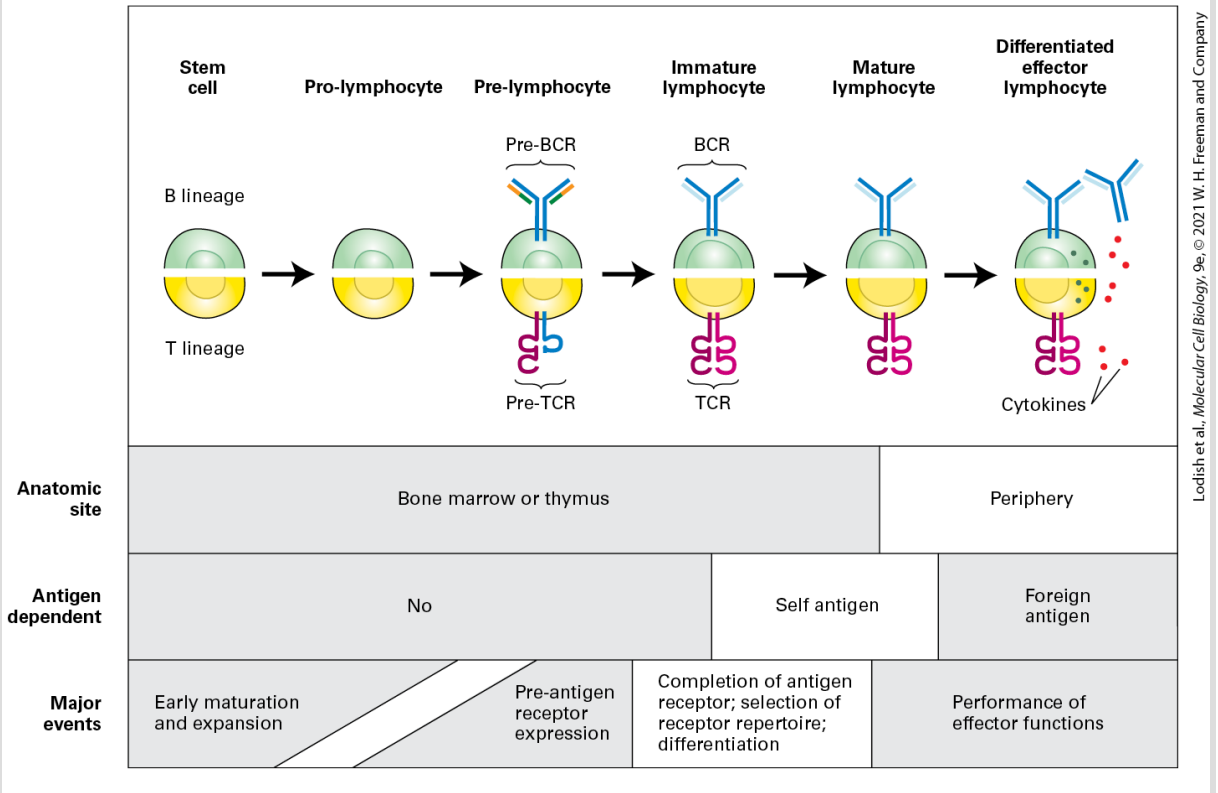


FIGURE 24-34 Comparison of T-cell and B-cell development. Cell fate decisions are executed by receptors containing either the newly rearranged μ chain (pre-BCR) or the newly rearranged β chain (pre-TCR). The pre-BCRs and pre-TCRs serve similar functions: signaling clonal expansion of cells that have successfully undergone rearrangement and allelic exclusion. This phase of lymphocyte development does not require antigen recognition. Both the pre-BCR and pre-TCR include subunits unique to each receptor type and absent from the antigen-specific receptors found on mature lymphocytes: VpreB and λ5 (orange, green) for the pre-BCR; pre-T α (blue) for the pre-TCR. Upon completion of the expansion phase, expression of the gene encoding the remaining subunit of the antigen-specific receptor begins: Ig light chain (light blue) for the BCR; TCR α chain (light red) for the TCR. Lymphocyte development and differentiation occur at distinct anatomic sites, and only fully assembled antigen-specific receptors (BCR, TCR) recognize antigen. Mature lymphocytes are strictly dependent on antigen recognition for their activation.

How is the emerging repertoire of T cells, with its diverse set of pre-TCRs, further sorted so that the only surviving T cells are ones

whose receptors have a productive interaction with self-MHC–peptide complexes? The random nature of the gene rearrangement process and the enormous variability engendered as a consequence produces a large and diverse set of TCRs. Yet the vast majority of this set of TCRs cannot interact productively with the host MHC products and are therefore useless. Thymocytes bearing useless receptors are not exported to the periphery: they wither on the vine and die. The immune system has developed selection processes to eliminate T cells that make TCRs incapable of a productive interaction with peptide complexes to the organism's own MHC (self MHC). Selection in the thymus also removes T cells with TCRs that strongly interact with self-MHC–self-peptide complexes, because such T cells may be self-reactive, damaging normal healthy tissue (autoimmunity). T cells whose TCRs recognize peptide-MHC complexes in the thymus with an affinity that falls between these two extremes receive survival signals and are positively selected. Some may by chance have the ability to bind strongly to an MHC protein complexed to a peptide derived from a foreign protein. Recall that antigen processing and presentation are constitutive processes, so that in the thymus, all self-MHC molecules are necessarily occupied with peptides derived from self proteins. These combinations of self peptides complexed to class I and class II MHC molecules constitute the substrate used by the set of newly generated T-cell receptors to determine what is self and ought to be ignored.

The heterogeneity of peptide-MHC complexes presented by different cells in the thymus to T cells undergoing selection makes it

highly probable that the T-cell receptor interprets signals not only by the strength and duration of the interaction with a particular peptide-MHC complex, but also in additive fashion, meaning the summation of the binding energies of the different MHC–self-peptide combinations it encounters. These help determine the outcome of the self–nonself selection. This process is called the *avidity model of T-cell selection*.

Newly made T cells are killed by apoptosis only if the appropriate self antigen is represented adequately in the thymus in the form of MHC-peptide complexes. How does the immune system ensure that T cells generated in the thymus learn to ignore self antigens that are not normally expressed at that location? Proteins that are expressed in tissue-specific fashion or after the development of the thymus, such as insulin in the β cells of the pancreas or the components of the myelin sheath in the nervous system, fit this category. A factor called *AIRE (a*uto*i*mmune *re*gulator) allows expression of such tissue-specific antigens in a subset of epithelial cells in the thymus. How AIRE accomplishes this is not known, but it is widely suspected that AIRE directly regulates the transcription of the relevant genes in the thymus and at select sites in secondary lymphoid organs. Defects in AIRE lead to a failure to express these tissue-specific antigens in the thymus. In individuals who do not express AIRE, developing T cells fail to receive the full set of instructions in the thymus that lead to the elimination of potentially self-reactive T cells. As a consequence, these individuals show a bewildering array

of autoimmune responses, causing widespread tissue damage and disease.

T Cells Commit to the CD4 or CD8 Lineage in the Thymus

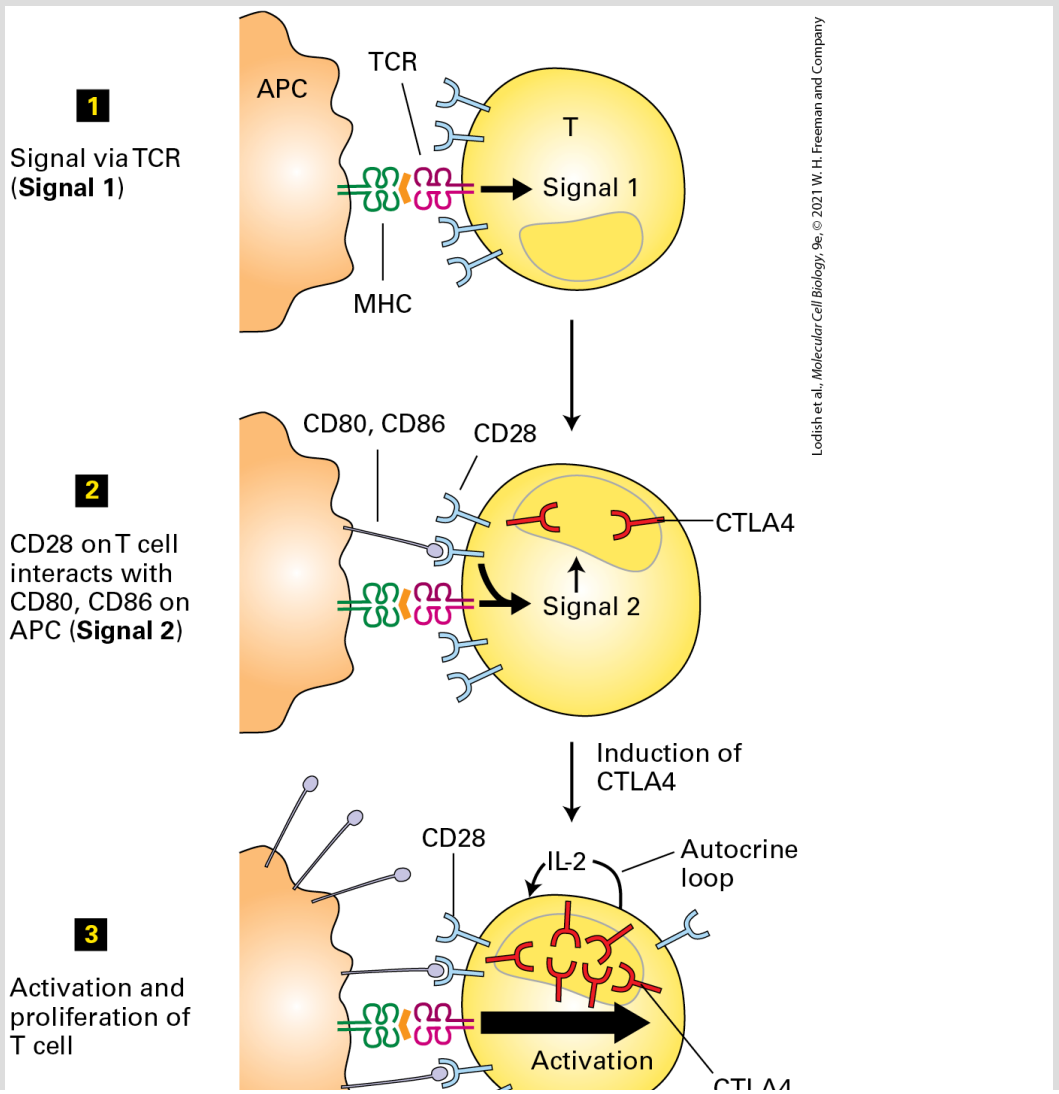
TCR gene rearrangement coincides with the acquisition of co-receptors. A key intermediate in T-cell development is a thymocyte that expresses both of the TCR co-receptors, CD4 and CD8, as well as a functional TCR-CD3 complex. These cells, called *double-positive* ($CD4CD8^+$) *cells*, are found only as developmental intermediates in the thymus. As the T cells mature, they lose either CD4 or CD8 to become single-positive cells. The choice of which co-receptor (CD4 or CD8) to express determines whether a T cell will recognize class I or class II MHC molecules. The question of how a $CD4CD8^+$ cell is instructed to become a CD8 (class I MHC-restricted) T cell or a CD4 (class II MHC-restricted) T cell is not entirely settled, but we know that the transcription factors ThPOK and Runx3 play fundamental roles. ThPOK and Runx3 are regulated by TCR signaling. Cells that transiently express ThPOK will commit to the CD4 lineage and repress Runx3 expression. On the other hand, if ThPOK expression is not induced, Runx3 expression is high, and cells commit to the CD8 lineage. In mice, a loss-of-function mutation in the *ThPOK* gene abrogates CD4 T-cell development, and all thymocytes become CD8-expressing T cells.

Another type of CD4 T cell also develops in the thymus, named *natural* (or *thymically derived*) *regulatory T cells (Treg)*. Their function differs from that of CD4 helper T cells, as will be described below. The thymus also gives rise to other less numerous types of T cells, such as invariant natural killer T cells (iNKT) that express the NK cell marker NK1.1 and are selected based on their ability to bind the nonclassic MHC molecule CD1, which presents lipid antigens. Also generated are intraepithelial lymphocytes that will colonize the mucosal surfaces of the intestine. After the final stages of maturation, T cells of all types are exported to the peripheral lymphoid organs.

T Cells Require Two Types of Signals for Full Activation

All T cells require a signal via their TCR for activation, but that signal is not sufficient: the T cell also needs co-stimulatory signals that are present on an antigen-presenting cell. To perceive these co-stimulatory signals, T cells carry on their surface several additional receptors, of which the CD28 molecule is the best known example. It is these regulatory signals that are the basis of cancer immunotherapies using monoclonal antibodies that target certain of these signaling proteins: CTLA4, PD-1, and PD-L1. These therapies are discussed in Section 24.6.

The CD28 protein on a T cell interacts with CD80 and CD86, two surface glycoproteins on the professional APC with which the T cell interacts. Expression of CD80 and CD86 increases when these APCs have themselves received the proper stimulatory signals, for example, by engagement of their Toll-like receptors (TLRs). The signals delivered to T cells via CD28 synergize with signals that emanate from the TCR when bound to its cognate self-MHC–peptide antigen complex, all of which are required for full T-cell activation (Figure 24-35, steps **1**–**3**).



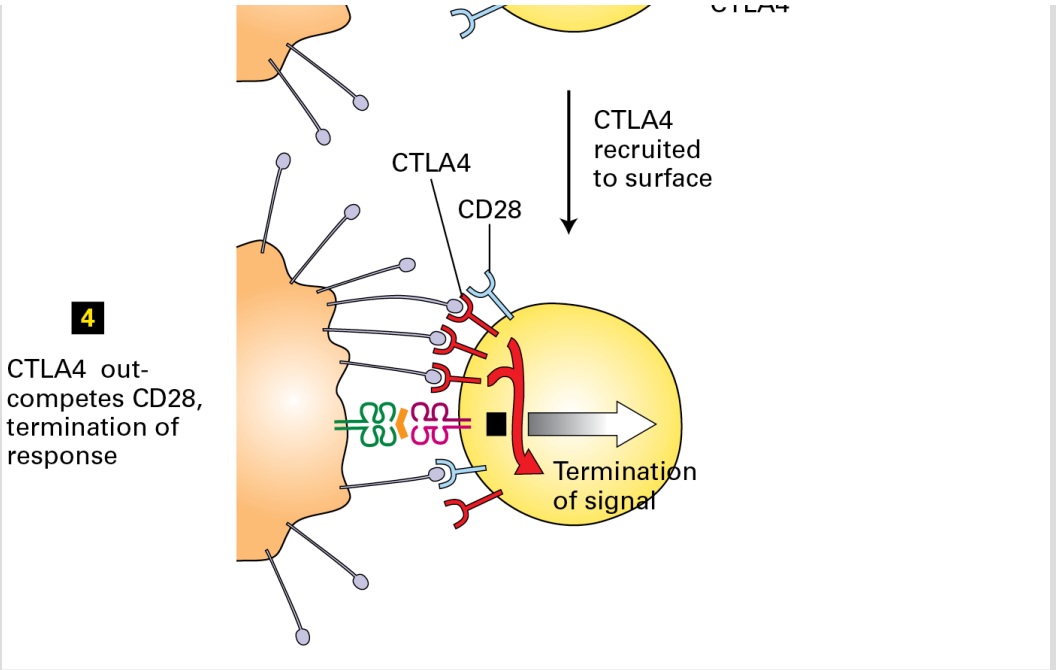


FIGURE 24-35 Signals involved in T-cell activation and its termination. The two-signal model of full T-cell activation: signal 1 is recognition of an MHC-peptide complex by the T-cell receptor (step 1), and signal 2 is recognition by CD28 of co-stimulatory molecules (CD80, CD86) on the surface of an antigen-presenting cell (step 2). Full activation, in turn, leads to increased expression of CTLA4 on the surface of T cells (step 3). CTLA4 binds CD80 and CD86, leading to inhibition of the T-cell response (step 4). Because the affinity of CTLA4 for CD80 and CD86 is greater than that of CD28, T-cell activation is eventually terminated. Note that after signal 1, if co-stimulation is not provided, the newly engaged T cell becomes unresponsive (anergic).

T cells, once activated, also express receptors that provide an attenuating or inhibitory signal upon recognition of these very same co-stimulatory molecules, providing negative feedback regulation. The CTLA4 protein, whose expression in T cells is induced only upon activation, competes with CD28 for binding of CD80 and CD86. Because the affinity of CTLA4 for the CD80 and C86 proteins is higher than that of CD28, the inhibitory signals provided through

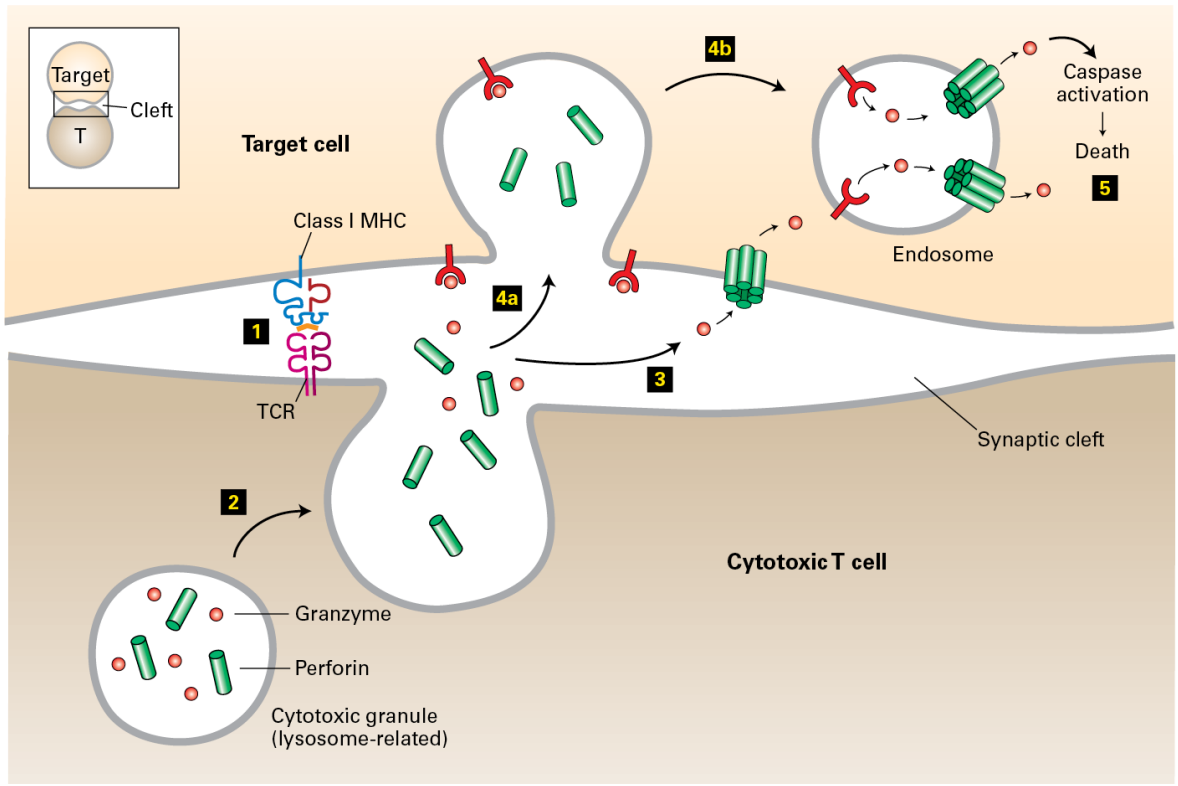
CTLA4 will ultimately overwhelm the stimulatory signals coming via CD28. Co-stimulatory molecules can thus be stimulatory or — as was discovered later without adjusting the nomenclature — inhibitory, and they therefore provide an important means of controlling the activation status and duration of a T-cell response.

Cytotoxic T Cells Carry the CD8 Co-Receptor and Are Specialized for Killing

As we have seen, cytotoxic T cells (CTLs) generally express on their surfaces the TCR co-receptor glycoprotein called *CD8*. These $CD8^+$ T cells kill target cells that display their cognate class I MHC–peptide combinations and do so with exquisite sensitivity: a single MHC-peptide complex suffices to allow a properly activated CTL to kill the target cell that bears it.

The mechanism of killing by CTLs involves two classes of proteins that act synergistically: perforins and granzymes (Figure 24-36). Perforins exhibit homology to the terminal components of the complement cascade composing the membrane attack complex (see Figure 24-5). Like the attack complex, perforins attach to a membrane and form pores up to 20 nm across within the membrane. The pores cause electrolytes and other small solutes to leak from the cell, contributing to cell death. Granzymes are

delivered to and are presumed to enter the target cell, probably via the pores generated by perforin. Granzymes are serine proteases that activate caspases and so propel the target cell on a path of programmed cell death (apoptosis; see Chapter 22). Perforins and granzymes are packaged into cytotoxic granules, which are stored inside the cytotoxic T cell. Upon binding of the T-cell receptor to its cognate class I MHC–antigen complex, signal transduction from the TCR leads to release of the cytotoxic granules and their contents into the extracellular space that is formed between the cytotoxic T cell and the target cell, called the *synaptic cleft*. How the T cell avoids being killed upon release of granzymes and perforins into the synapse is unknown. Natural killer cells also exert cytotoxic activity and likewise rely on perforins and granzymes to kill their targets (see Figure 24-6).



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 24-36 Perforin- and granzyme-mediated cell killing by cytotoxic T cells. Upon recognition of a target cell (step 1), a cytotoxic T cell forms a tight antigen-specific contact with the target cell. Tight contact results in the formation of a synaptic cleft, into which the contents of cytotoxic granules, including perforins and granzymes, are released (step 2). Perforins form pores in the membranes onto which they adsorb, and granzymes are serine proteases that enter through the perforin pores (step 3). Perforins are believed to act not only at the surface of the target cell, but also at the surface of its endosomal compartments after the perforin molecules have been internalized from the cell surface (step 4). Once in the cytoplasm, the granzymes activate caspases, which initiate programmed cell death (step 5).

T Cells Secrete an Array of Cytokines That Provide Signals to Other Immune-System Cells

Many lymphocytes and other cells in lymphoid tissues produce cytokines. These small secreted proteins control lymphocyte activity by binding to specific cytokine receptors on the surface of a lymphocyte and initiating a transcriptional program that allows the lymphocyte to either proliferate or differentiate into an effector cell ready to exert cytotoxic (cytotoxic T cells), helper (helper T cells), or antibody-secreting activity (B cells). Cytokines that are produced by or act primarily on leukocytes are called **interleukins**. More than 35 interleukins have been recognized and molecularly characterized. Interleukin receptors share structural similarity with each other and with receptors for erythropoietin; those interleukins whose structures are most closely related can be recognized by their cognate receptors. The interleukin-2 receptor (Figure 16-19b) is particularly well characterized. Interleukin 2 (IL-2), a T-cell growth factor, is one of the first cytokines produced when T cells are stimulated. IL-2 acts as an autocrine (self-acting) and paracrine (acting on neighboring cells) growth factor and drives clonal expansion of activated T cells.

Interleukin 4 (IL-4), which is produced by helper T cells, induces activated B cells to proliferate and to undergo class-switch recombination and somatic hypermutation. Interleukin 7 (IL-7), produced by stromal cells in the bone marrow, is essential for development of T and B cells. Both IL-7 and IL-15 play a role in the maintenance of *memory cells*, which are antigen-experienced T cells that may be called upon when re-exposure to antigen occurs; these

memory cells then rapidly proliferate and deal with the reinvading pathogens. The receptors for IL-2, IL-4, IL-7, and IL-15 all rely on a common subunit for signal transduction, the common γ chain (γ_c), with α (IL-2, IL-15) and β subunits (IL-2, IL-4, IL-7, IL-15) providing ligand specificity (Figure 16-19b). Genetic defects in the γ_c result in nearly complete failure of lymphocyte development, illustrating the importance of these cytokines (and their receptors) not only during the effector phase of an immune response, but also in the course of lymphocyte development, where IL-7 in particular plays a key role. By knocking out the gene that encodes γ_c, immunodeficient mice can be generated that accept human tissues and tumors as grafts. Such mice have become very useful models to study tumor immunology and to create a setting in which a partial human immune system can be installed in mice, for example, to study HIV, which does not infect mouse lymphocytes.

The mechanism of signal transduction by cytokine receptors through the JAK/STAT pathway is described in Chapter 16 (reviewed in Figure 16-19a). Among the many genes under the control of interleukins and the STAT pathway are those that encode suppressors of cytokine signaling, or SOCS proteins. These proteins, which are themselves induced by cytokines, bind to the activated form of JAKs and target them for proteasomal degradation (see Figure 16-21b).

Helper T Cells Are Divided into Distinct Subsets Based on Their Cytokine Production and Expression of Surface Markers

CD4-expressing T cells are helper T cells that provide assistance to B cells and guide their proliferation and differentiation into plasma cells. This function requires both the production and secretion of cytokines such as IL-4 as well as direct contact between the helper T cell and the B cell to which it provides help.

A second class of helper T cell has as its major function secreting the cytokines that contribute to the establishment of an inflammatory environment. Multiple subtypes of such inflammatory T cells are categorized based on the spectrum of different cytokines they produce and their respective roles in regulating immune responses. Whereas all activated T cells can produce IL-2, other cytokines are produced only by particular helper T-cell subsets.

Helper T cells classified as T_H1 cells secrete interferon γ and tumor necrosis factor (TNF), which can activate macrophages and stimulate an inflammatory response. Referred to also as *inflammatory T cells*, T_H1 cells nonetheless play an important role in antibody production, notably facilitating the production of complement-fixing antibodies such as IgG1 and IgG3. T_H2 cells

secrete IL-4 and IL-10, and through production of IL-4, play an important role in B-cell responses that involve class switching to the IgG1 and IgE isotypes (see Section 24.3). Recall that in B cells, the induction of activation-induced deaminase (AID) prepares the B cell for class-switch recombination and somatic hypermutation. This induction is a consequence of the precise mixture of cytokines produced by helper T cells and the binding of CD40, a surface membrane protein on the activated T cell, to a protein on the B-cell surface, called *CD40 ligand* or *CD40L*.

Conventional helper T cells can also differentiate into T_H17 cells, which produce IL-17, and into *induced regulatory T cells* (induced Tregs, distinct from the natural Tregs generated in the thymus). Both types of Treg cells attenuate immune responses by exerting a suppressive effect on other types of T cells. Natural Tregs restrain the activity of potentially self-reactive T cells and are important in maintaining peripheral tolerance (the absence of an immune response to self antigens), whereas induced Tregs are believed to regulate excessively strong immune responses against foreign antigens. T_H17 cells are important in defense against bacteria (extracellular bacteria in particular) and also play a pathogenic role in autoimmune diseases.

Innate Lymphoid Cells Regulate Inflammation and the Overall Immune

Response

The immune response against a pathogen exerts a selective pressure that allows the emergence of pathogen variants no longer recognized by antibodies or T cells. Many viruses with comparatively small genomes, such as the virus that causes influenza or human immunodeficiency virus, replicate by means of error-prone polymerases that generate mutant descendants with amino acid substitutions, some of which escape immune recognition because the proper epitope is no longer expressed as a consequence of such mutations. Viruses with large genomes such as herpesviruses use other means of immune escape, for example, through down-regulation of Class I MHC products. This renders the infected cell refractory to killing by CD8 T cells. A similar situation applies to many cancerous cells, which often show starkly reduced expression of Class I MHC products. Natural killer (NK) cells can sense the absence of Class I MHC products because NK cells express receptors that recognize Class I MHC products and shut off NK cell function. A lack of Class I MHC products on the surface of a virus-infected cell or on a tumor cell relieves the inhibition mediated by such receptors and allows the NK cell to kill the target. Because of this, the combined action of CD8 T cells and NK cells affords a level of protection against virus-infected and tumor cells that is more effective than either cell type can confer on its own.

Although NK cells were discovered several decades ago, they are now referred to as a member of the group of *innate lymphoid cells* or *ILCs*. Such cells resemble conventional lymphocytes in many of their functional properties such as the types of cytokine they can produce, but they do not make use of gene rearrangements to create the receptors characteristic of the cells of the adaptive immune system: the B and T cells. ILCs are important in regulating inflammation and in regulatory control of the immune response more generally. They are poised to rapidly secrete cytokines when activated and their preferred location at mucosal sites suggests a role in the control of various infections. The group 1 ILC lineage comprises ILCs such as NK cells that produce type 1 cytokines, notably interferon-γ and TNF. ILC2s express interleukin-5 (IL-5) and IL-13, not unlike conventional T_H2 CD4 T cells. ILC2s, together with T_H2 CD4 T cells, protect against helminth (parasitic worms) infection. ILC3s were first defined as intestinal lymphoid cells that express the NK cell–activating receptor NKp46 but are otherwise quite distinct from NK cells (ILC1s). ILC3s express IL-17A and IL-22. Lymphoid tissue-inducer (LTi) cells are an ILC subset that appears to be closely related to ILC3s, but their exact relationship remains controversial. Together, LTi cells and ILC3s have been classified as group 3 ILCs.

Leukocytes Move in Response to Chemotactic Cues Provided by

Chemokines

Interleukins tell lymphocytes what to do by eliciting a transcriptional program that allows lymphocytes to acquire specialized effector functions. Chemokines, on the other hand, tell leukocytes where to go. Many cells emit chemotactic cues in the form of chemokines. When tissue damage occurs, resident fibroblasts produce a chemokine, IL-8, that attracts neutrophils to the site of damage. The regulation of lymphocyte traffic within lymph nodes is essential for dendritic cells to attract T cells and for T cells and B cells to meet. These trafficking steps are all controlled by chemokines.

There are approximately 40 distinct chemokines and more than a dozen chemokine receptors. One chemokine may bind to more than one receptor, and a single receptor can bind several different chemokines. This flexibility creates the possibility of generating a combinatorial code of chemotactic cues of great complexity. This code is used to guide the navigation of leukocytes from where they are generated, in the bone marrow, into the bloodstream for transport to their target destination.

Some chemokines direct lymphocytes to leave the circulation and take up residence in lymphoid organs. These migrations contribute to the population of lymphoid organs with the required sets of lymphocytes. Because these movements occur as part of normal lymphoid development, such chemokines are referred to as

homeostatic chemokines. Those chemokines that serve the purpose of recruiting leukocytes to sites of inflammation and tissue damage are referred to as *inflammatory chemokines*.

Chemokine receptors are G protein–coupled receptors that function as an essential component of the regulation of cell adhesion and cell migration. Leukocytes that travel through blood vessels do so at high speed and are exposed to high hydrodynamic shear forces. For a leukocyte to traverse the endothelium and take up residence in a lymph node or seek out a site of infection in tissue, it must first slow down, a process that requires interactions of glycoprotein surface receptors called *selectins* with their ligands on the surfaces of leukocytes, which are mostly carbohydrate in nature. If chemokines are adsorbed to the extracellular matrix and if the leukocyte possesses a receptor for those chemokines, activation of its chemokine receptor elicits a signal that allows integrins carried by the leukocyte to undergo a conformational change. This change results in an increase in the affinity of the integrin for its ligand and causes firm arrest of the leukocyte. The leukocyte may now exit the blood vessel by a process known as *extravasation* (see Figure 20-40).

KEY CONCEPTS OF SECTION 24.5

T Cells, T-Cell Receptors, and T-Cell Development

- The antigen-specific T-cell receptors are dimeric proteins consisting of α and β subunits or γ and δ subunits. T cells occur in at least two major classes defined by their expression of the glycoprotein co-receptors CD4 and CD8 (see Figure 24-31).

- Cells that use class I MHC molecules as the molecular guideposts for antigen recognition (restriction elements, in immunological parlance) carry CD8; those that use class II MHC molecules carry CD4. These classes of T cells are functionally distinct: CD8 T cells are cytotoxic T cells; CD4 T cells provide help to B cells and are an important source of cytokines.
- Genes encoding the TCR subunits are generated by somatic recombination of V and J segments (α chain) and of V, D, and J segments (β chain); their rearrangement obeys the same rules as does rearrangement of Ig genes in B cells (see Figure 24-32). Rearrangement of TCR genes occurs when the lymphocytes are present in the thymus and only in those cells destined to become T lymphocytes.
- A complete T-cell receptor includes not only the α and β subunits responsible for antigen and MHC recognition, but also the accessory subunits referred to as the *CD3 complex*, which is required for signal transduction. Each subunit of the CD3 complex carries in its cytoplasmic tail one or three ITAM domains; when phosphorylated, these ITAMs recruit adapter proteins involved in signal transduction (see Figure 24-33).
- In the course of T-cell development, the TCR β locus is rearranged first. If that locus encodes a functional β subunit, it is incorporated with the pre-Tα chain into a pre-TCR (see Figure 24-34). Like the pre-BCR, the pre-TCR mediates allelic exclusion, that is, the expression of a functionally rearranged T-cell receptor encoded by only one of the two alleles and proliferation of those cells that successfully underwent TCRβ rearrangement.
- Developing T cells that fail to recognize self-MHC molecules die for lack of survival signals. T cells that interact too strongly with self-peptide–self-MHC complexes encountered during development are instructed to die (negative selection); those that have intermediate affinity for self-peptide–self-MHC complexes are allowed to mature (positive selection) and are exported from the thymus to the periphery.
- Innate lymphoid cells (ILCs) share properties with lymphocytes, notably the types of cytokines they release, without reliance on antigen-specific receptors, unlike T lymphocytes.
- T cells are instructed where to go (cell migration) through chemotactic signals in the form of chemokines. Receptors for chemokines are G protein–coupled receptors that show some promiscuity in terms of their binding of chemokines. The complexity of chemokine–chemokine receptor binding allows precise regulation of leukocyte traffic, both within lymphoid organs and in the periphery.

24.6 Collaboration of Immune-System Cells in the Adaptive Response

An effective adaptive immune response requires the presence of B cells, T cells, and antigen presenting cells (APCs). For B cells to execute class-switch recombination and somatic hypermutation — prerequisites for production of high-affinity antibodies — they require help from activated T cells. These T cells, in turn, can be activated only by professional APCs such as dendritic cells. Dendritic cells sense the presence of pathogens through Toll-like receptors and other pattern-recognition receptors, such as the cell-surface mannose-binding protein and other related lectins that can recognize polysaccharides and carbohydrate determinants. The interplay between components of the innate and adaptive immune systems is therefore a very important aspect of adaptive immunity. This layered, interwoven nature of innate and adaptive immunity ensures a rapid early response of immediate protective value and primes the adaptive immune system for a specific response to any persisting pathogen. In this section, we describe how these various elements are activated and how the relevant cell types interact.

Toll-Like Receptors Perceive a Variety of Pathogen-Derived Macromolecular

Patterns

An important part of the innate immune system is its ability to immediately detect the presence of a microbial invader and respond to it. This response includes direct elimination of the invader, but it also prepares the mammalian host for a proper adaptive immune response, particularly through activation of professional APCs. These APCs are positioned throughout the epithelia (airways, gastrointestinal tract, genital tract), where contact with pathogens is most likely to occur. In the skin, a network of dendritic cells called *Langerhans cells* makes it virtually impossible for a pathogen that breaches this barrier to avoid contact with these professional APCs. Dendritic cells and other professional APCs detect the presence of bacteria and viruses through members of the Toll-like receptor (TLR) family. These proteins are named after the *Drosophila* protein Toll because of the structural and functional homology between them. *Drosophila* Toll was discovered because of its important role in dorsal/ventral patterning in the fruit fly, but related receptors are now recognized as capable of triggering an innate immune response in insects as well as in vertebrates.

TLR Structure

Toll itself and all TLRs possess a sickle-shaped extracellular domain, composed of *leucine-rich repeats*, that is involved in ligand recognition. The cytoplasmic portion of a TLR contains a domain

responsible for the recruitment of adapter proteins to enable signal transduction. The signaling pathways engaged by TLRs have many of the same components (and outcomes) as those used by receptors for the cytokine IL-1 (Figure 24-37; see also Figure 16-31).

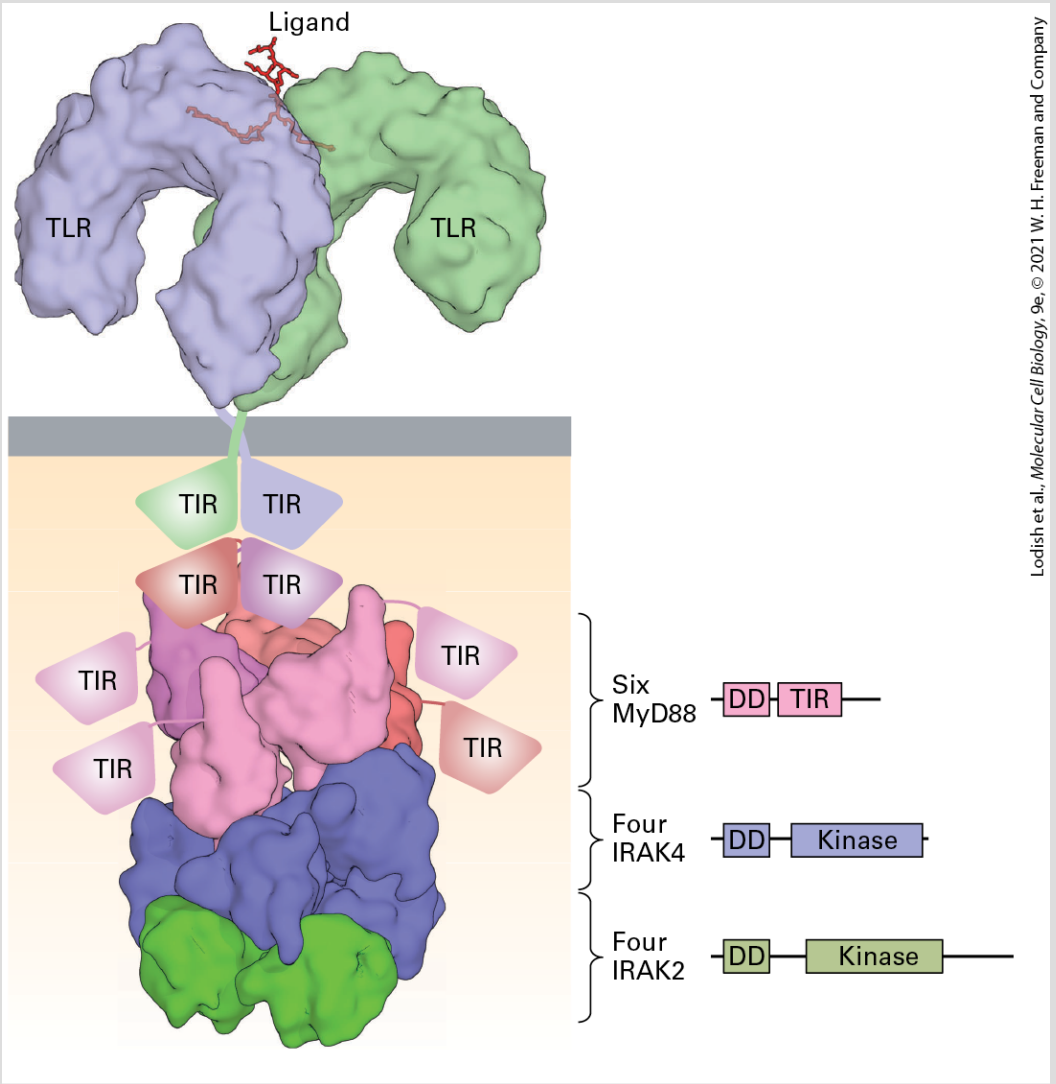


FIGURE 24-37 Toll-like receptor activation. The extracellular portions of TLRs recognize ligands of diverse chemical nature (nucleic acids, lipopolysaccharides). The cytoplasmic portions of the TLRs, called *TIR* (BToll/IL-1β receptor homology) domains, associate with the adapter protein MyD88, present in six copies per complex, and recruit two types of kinases, both members of the IRAK family. These complex interactions are maintained by TIR

domains and death domains (DD) as shown in the figure. The assembled complex on the cytoplasmic side is referred to as the *myddosome*. As depicted in Figure 16-31, among the proteins bound to IRAK2 is TRAF6, an E3 ubiquitin ligase. Because of extensive TRAF6-TRAF6 interactions, hundreds of copies of this TLR complex become cross-linked together, forming a giant signalsome in the plasma membrane.

[Republished with permission from Annual Reviews, from J. Y. Kang and J.-O. Lee, 2011, "Structural Biology of the Toll-Like Receptor Family," *Annu. Rev. Biochem.* **80**(1):917–941; permission conveyed through Copyright Clearance Center, Inc.]

The *Drosophila* Toll protein interacts with its ligand, Spaetzle, itself the product of a proteolytic conversion initiated by components of the cell walls of fungi that prey on *Drosophila*. In the fly, activation of Toll unleashes a signaling cascade that ultimately controls the transcription of genes that encode antimicrobial peptides. The activated receptor at the cell surface communicates with the transcriptional apparatus by means of a series of adapter proteins that activate downstream kinases interposed between Toll, the TLRs, and the transcription factors that are activated by them. A key step is the ubiquitin-dependent proteasomal degradation of the Cactus protein. Its removal allows the protein Dif to enter the nucleus and initiate transcription. This pathway is highly homologous in its operation and structural composition to the NF-κB pathway in mammals (see Figure 16-30).

Diversity of TLRs

There are approximately a dozen mammalian TLRs that can be activated by various microbial products and are expressed by a variety of cell types; TLR function is crucial for the activation of dendritic cells and macrophages. Neutrophils also express TLRs. The microbial products recognized by TLRs include macromolecules found in the cell envelopes of bacteria, such as lipopolysaccharides, flagellins (subunits of bacterial flagella), and bacterial lipopeptides. Direct binding of at least some of these macromolecules to TLRs has been demonstrated in crystallographic analyses of the relevant complexes. The presence of distinct classes of microbial molecules is sensed by distinct TLRs: for example, TLR4 binds lipopolysaccharides; heterodimers of TLR1 and 2 and TLR2 and 6 bind lipopeptides; and TLR5 binds flagellin. Recognition of all bacterial envelope components by these TLRs occurs at the cell surface.

A second set of TLRs — TLR3, TLR7, and TLR9 — sense the presence of pathogen-derived nucleic acids. They do so not at the cell surface but rather within the endosomal compartments where these receptors reside. Mammalian DNA is methylated at many CpG dinucleotides, whereas microbial DNA generally lacks this modification. TLR9 is activated by unmethylated, CpG-containing microbial DNA. Similarly, double-stranded RNA molecules, replication intermediates of RNA viruses, are present in these virus-infected cells and lead to activation of TLR3. Finally, TLR7 responds to the presence of certain single-stranded RNAs. Thus the full set of mammalian TLRs allows the recognition of a variety of

macromolecules that are diagnostic for the presence of bacterial, viral, or fungal pathogens and parasites such as malaria.

Inflammasome

A variety of intracellular receptors for RNA and DNA have been described that recognize viral RNA and are structurally distinct from TLRs. The list of cytoplasmic receptors capable of recognizing DNA, both pathogen-derived and host DNA-derived, continues to grow. Several of these receptors participate in the assembly of the inflammasome (Figure 24-38), whose major function is the conversion of the enzyme precursor procaspase-1 to the active caspase-1. Caspase-1 is a protease that converts pro-IL-1β into active IL-1β, a cytokine that elicits a strong inflammatory response. Mature IL-1β lacks a signal sequence and is released from producer cells via an unconventional secretory route that involves the formation of pores composed of gasdermin proteins (Figure 24-39). Gasdermins can be converted by caspases from their inactive precursors to the active form to yield pores of a size that allows passage of proteins. In so doing, these pores not only contribute to the release of IL-1β but also to certain forms of cell death called *pyroptosis* ("fiery death," referring to the inflammation that occurs). The core components of inflammasomes are proteins with leucine-rich repeats, members of the neuronal inhibitors of apoptosis (NALP) family of proteins, and the NOD proteins, so named because of the presence of a *n*ucleotide *o*ligomerization *d*omain. Ipaf-1, a protein related to the Apaf-1

molecule involved in apoptosis (see Chapter 22), recruits an adapter protein, ASC, to mediate complex formation with procaspase-1. Assembly of this multisubunit complex allows the conversion of procaspase-1 to active caspase-1 and of pro-IL-1β to IL-1β. Many seemingly unrelated substances can induce assembly and activation of an inflammasome, including silica, uric acid crystals, and asbestos particles. Accordingly, inhibition of the inflammasome signaling cascade, or blocking of the receptor for IL-1β, has shown therapeutic promise for a variety of inflammatory diseases.

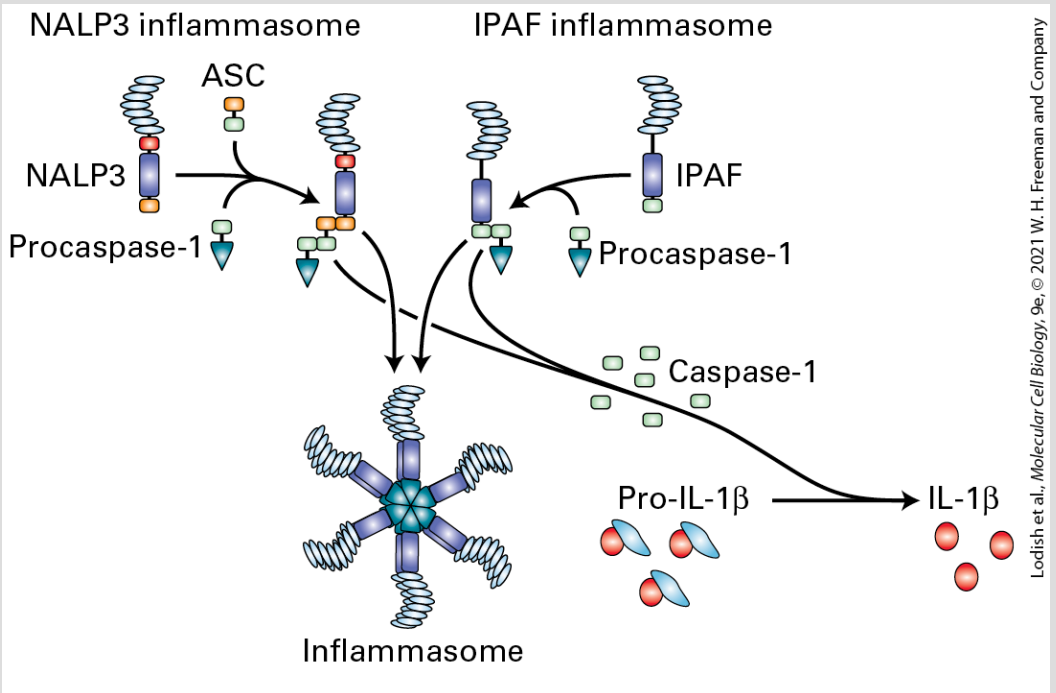


FIGURE 24-38 The inflammasome. The inflammasome is a type of complex that senses the presence of cytoplasmic, pathogen-derived nucleic acids. It can also be activated by other danger signals including particulate matter such as uric acid crystals or asbestos. There are close to two dozen proteins that can participate in the formation of these complexes to yield inflammasomes of different composition, two of which are represented here. The fully assembled inflammasome activates caspase-1, the enzyme that converts pro-IL-1β into

the active, cleaved cytokine IL-1β. NALP3, also referred to as *NLRP* = a member of the protein family characterized by the presence of NACHT, LRR, and PYD domains; ASC = apoptosis-associated Speck-like protein containing a CARD (caspase recruitment domain).

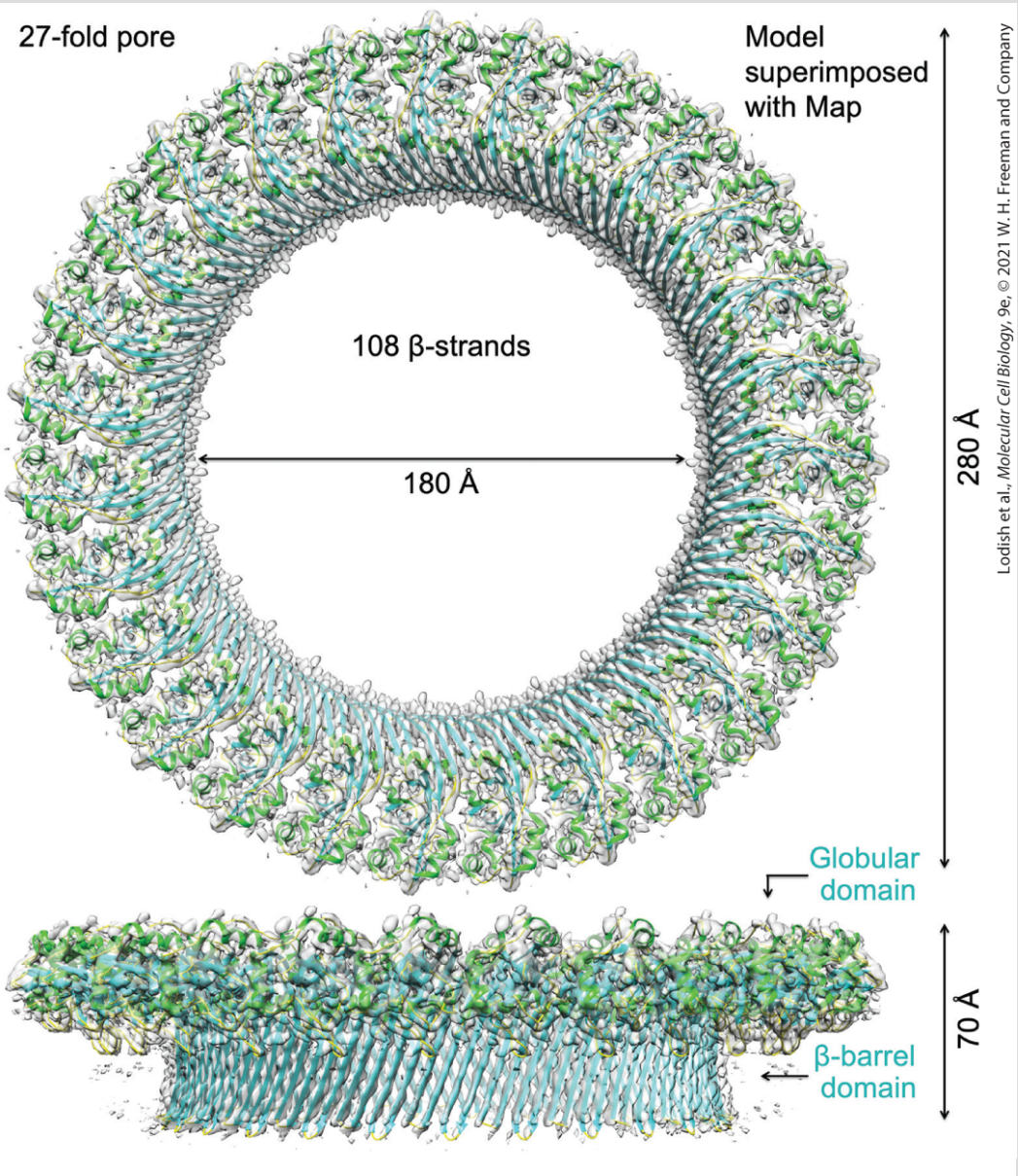


FIGURE 24-39 Gasdermin pore structure. Structure of the gasdermin A3 pore determined by cryoelectron microscopy. Gasdermin subunits are converted by caspases upon activation of inflammasomes. The activated subunits self-assemble into a multimeric structure, shown here for 27 subunits of the

gasdermin A3 complex in top and side views. The β barrel domain interacts with the fatty acyl chains of the lipid bilayer and spans the width of the membrane. The resulting pore allows the escape of cellular content, thus contributing to cell death. The large pores formed by an assembled gasdermin complex may facilitate the release into the extracellular environment of intracellularly activated cytokines such as IL-1, a protein that lacks a signal sequence typical of secretory proteins.

[Republished with permission from Springer Nature, from J. Ruan et al., 2018, "Cryo-EM Structure of the Gasdermin A3 Membrane Pore," *Nature*, **557**(7703):62–67; permission conveyed through Copyright Clearance Center, Inc.]

TLR Signaling Cascade

As shown in Figure 24-37, engagement of mammalian TLRs leads to recruitment of the adapter protein MyD88, which in turn allows the binding and activation of IRAK (*i*nterleukin 1 *r*eceptor-*a*ssociated *k*inase) proteins. After IRAK phosphorylates TNF-receptor–associated factor 6 (TRAF6), several downstream kinases come into play, leading to release of active NF-κB, a transcription factor, for translocation from the cytoplasm to the nucleus, where NF-κB activates various target genes (see Figure 16-30). These target genes include those encoding the cytokines IL-1β and IL-6, which as we saw contribute to inflammation, as well as the genes for TNF and IL-12. Expression of type I interferons, small proteins with antiviral effects, is also turned on in response to TLR signaling.

Cell responses to TLR signaling are quite diverse. For professional APCs, these responses include production of cytokines as well as the up-regulation of co-stimulatory molecules, the surface proteins important for full activation of T cells that have yet to encounter antigen (referred to as “naive” T cells). TLR signaling allows dendritic cells to migrate from where they encounter a pathogen to the lymph nodes draining that area, where they can interact with naive lymphocytes. Not all activated TLRs evoke an identical response. In dendritic cells, each activated TLR controls production of a particular set of cytokines. For each engaged TLR, the combination of co-stimulatory molecules and the cytokine profile induced by TLR engagement creates a unique activated-dendritic-cell phenotype. The identity of the microbial antigen encountered by a dendritic cell determines the pattern of the TLRs that will be activated. This pattern, in turn, shapes the differentiation pathways of activated dendritic cells, influencing the cytokines produced, the surface molecules displayed, and the chemotactic cues to which the dendritic cells respond. The mode of activation of a dendritic cell and the cytokines it produces in response create a unique local microenvironment in which T cells differentiate. Within this microenvironment, the neighboring T cells acquire the functional characteristics required to fight the infectious agent that led to engagement of the TLRs in the first place.

Engagement of Toll-Like Receptors Leads to Activation of Antigen-Presenting Cells

Professional APCs engage in continuous endocytosis, and in the absence of pathogens, they display at their surface class I and class II MHC molecules loaded with peptides derived from self proteins. In the presence of pathogens, the TLRs on these cells become activated, inducing the APCs to become motile. The cells detach from the surrounding extracellular matrix and start to migrate in the direction of the draining lymph node, following the directional cues provided by chemokines. An activated dendritic cell, for example, reduces its rate of antigen acquisition, up-regulates the activity of endosomal and lysosomal proteases, and increases the transfer of class II MHC–peptide complexes from the loading compartments to the cell surface. Finally, activated professional APCs up-regulate expression of the co-stimulatory molecules CD80 and CD86 (see Figure 24-35), which will allow these APCs to activate T cells more effectively. The initial contact of a professional APC with a pathogen thus results in its migration to the draining lymph node in a state that is fully capable of activating a naive T cell. Antigen is displayed in the form of peptide-MHC complexes, co-stimulatory molecules are abundantly present, and cytokines are secreted that assist in setting up the proper differentiation program for the T cells to be activated.

Antigen-laden dendritic cells engage antigen-specific T cells, which respond by proliferating and differentiating. The cytokines produced in the course of this priming reaction determine whether a CD4-expressing T cell will polarize toward an inflammatory or a classic helper T cell phenotype. If engagement occurs via class I MHC molecules, a CD8-expressing T cell may develop from a precursor cytotoxic T cell into a fully active cytotoxic T cell. Activated T cells are motile and move through the lymph node in search of B cells or enter the circulation to execute effector functions elsewhere in the body.

Production of High-Affinity Antibodies Requires Collaboration Between B and T cells

To generate the high-affinity antibodies that are necessary for tight binding to antigens and effective neutralization of pathogens, B cells require assistance from T cells. B-cell activation requires both a source of antigen to bind to the B-cell receptor (BCR) and the presence of activated antigen-specific T cells (Figure 24-40).

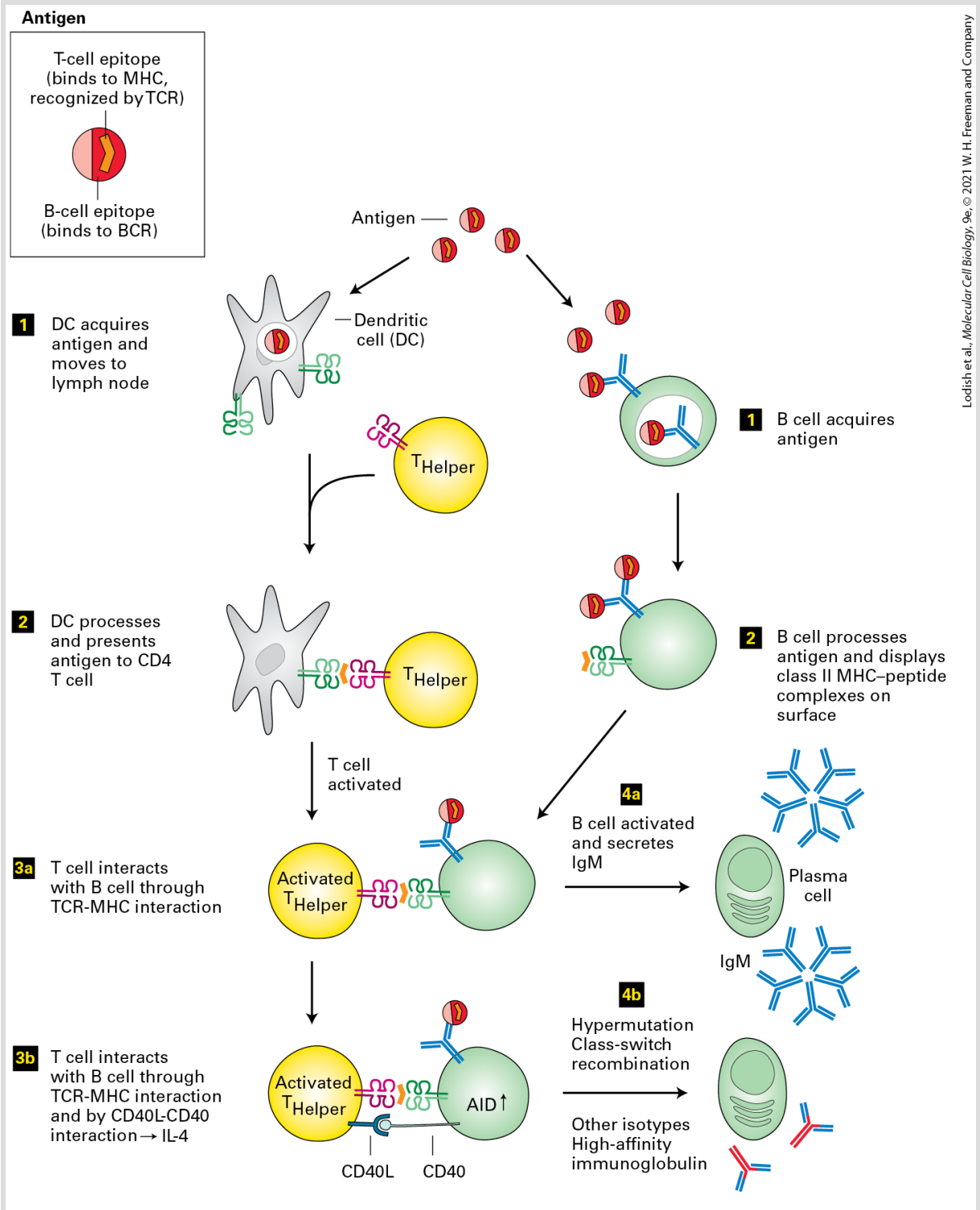


FIGURE 24-40 Collaboration between T and B cells is required to initiate the production of antibodies. *Left:* Activation of T cells by means of antigen-loaded, antigen-presenting dendritic cells (DCs). *Right:* Antigen acquisition by and subsequent activation of B cells to proliferate and differentiate into antibody-secreting plasma cells. Step 1: Professional

antigen-presenting cells (dendritic cells, B cells) acquire antigen. Step **2**: Professional APCs internalize and degrade the antigen and present the resulting antigen-derived peptides, bound to MHC proteins. T-cell activation occurs when dendritic cells present antigen to T cells with receptors that bind the antigen peptide-MHC complex. Step **3a**: Activated T cells engage B cells that express a BCR (surface-bound immunoglobulin) for the antigen and that (like other APCs) present antigen to T cells through peptide-MHC complexes displayed on the surface of the B cell. Step **3b**: T cells that are persistently activated initiate expression of the CD40 ligand (CD40L), a prerequisite for B cells becoming fully activated and turning on the enzymatic machinery (including AID) to initiate class-switch recombination and somatic hypermutation. Step **4a**: A B cell that receives the appropriate instructions from CD4 helper T cells becomes an IgM-secreting plasma cell. Step **4b**: A B cell that receives signals from activated CD4 helper T cells in the form of CD40-CD40L interactions and the appropriate cytokines can switch to other immunoglobulin isotypes and engage in somatic hypermutation, eventually differentiating into a plasma cell.

Soluble antigen reaches the lymph node from the periphery by transport through the afferent lymphatic vessels (see Figure 24-7). Bacterial growth is accompanied by the release of microbial products that can serve as antigens. If the infection is accompanied by local tissue destruction, activation of the complement cascade results in the killing of bacteria and the concomitant release of bacterial proteins, which are also delivered via the lymphatic vessels to the draining lymph node. Antigens covalently modified by proteins of the complement system are superior to their unmodified counterparts in the activation of B cells through engagement of complement receptors on those cells, which serve as co-receptors for the BCR.

B cells that acquire antigen via their BCRs internalize the antibody-antigen complex by endocytosis and proteolytically degrade the

antigen into peptides for presentation to T cells via the class II MHC pathway (see Figure 24-28). B cells that have successfully engaged antigen thus convert it into a call for T-cell help in the form of a class II MHC–peptide complex expressed on the cell surface (Figure 24-40, step **2**). Importantly, the sequence of amino acids on the antigen molecule recognized by the B-cell receptor (a membrane-anchored immunoglobulin) may be quite distinct from the peptide ultimately displayed on the B-cell surface in association with a class II MHC molecule. As long as the B-cell epitope on the antigen and the class II MHC–presented peptide — a T-cell epitope — are physically part of the same antigen, whether linked covalently or held together through noncovalent interactions, successful B-cell differentiation and antibody production can be initiated by T-cell help.

This concept of linked recognition — namely, the engagement of antigen by the B cell's BCR and the display of antigen-derived fragments to T cells by class II MHC molecules — explains why there is a minimum size for molecules that can be used to successfully elicit a high-affinity antibody response, as we will see below. Such immunogenic molecules must fulfill several criteria: they must contain the epitope that binds to the B-cell receptor, they must undergo endocytosis and proteolysis, and a proteolyzed fragment of the protein must bind to the allelic class II MHC molecules available in order to be presented as a class II MHC–peptide complex, which serves as a call for T-cell help.

Often investigators would like to generate an antibody (either polyclonal or monoclonal) that can recognize a short peptide fragment from a larger protein. Such antibodies can be used for a variety of experiments, including detection of a target molecule by immunofluorescence or immunoprecipitation. These antibodies are called *anti-peptide antibodies*. If the peptide alone is used as an immunogen [injected into an animal (e.g., a rabbit, goat, or mouse) to generate antibodies], it probably will not successfully induce robust antibody formation, even though there may be B cells with BCRs that can bind tightly to the peptide. The reason is that it is unlikely that those B cells will be able to generate a complex of a class II MHC with that same peptide and recruit helper T cells to drive proliferation and affinity maturation. For this reason, synthetic peptides used to elicit antibodies are conjugated to carrier proteins to improve their immunogenicity; the carrier proteins serve as the source of peptides for presentation via class II MHC products. Only through recognition of such a class II MHC–peptide complex via its TCR can a T cell provide the help necessary for the B cell to run its complete course of differentiation leading to robust, high-affinity antibody production.

This concept applies equally to B cells capable of recognizing particular modifications on proteins or peptides. Antibodies that recognize the phosphorylated form of a kinase are commonly raised by immunization of experimental animals with the phosphorylated peptide in question, conjugated to a carrier protein. An appropriately specific BCR recognizes the phosphorylated site on the peptide of interest, internalizes the phosphorylated peptide and carrier, and generates a complex set of peptides by endosomal proteolysis of the carrier protein. Among these peptides, there should be at least one that can bind to the class II MHC molecules carried by that B cell. If properly displayed at the surface of the B cell, this class II MHC–peptide complex becomes the call for T-cell help, which is provided by helper T cells equipped with receptors capable of recognizing the complex of class II MHC molecule and carrier-derived peptide. ■

The helper T cell identifies, via its TCR, an antigen-experienced B cell by means of the class II MHC–peptide complex the B cell displays (Figure 24-40, step **3a**). The B cell also displays co-stimulatory molecules and receptors for cytokines produced by the activated T cell (e.g., IL-4). After interacting with T cells, these B cells proliferate. Some of them differentiate into plasma cells; others are set aside and become memory B cells. The first wave of antibodies they produce is always IgM (Figure 24-40, step **4a**). Class switching to other isotypes and somatic hypermutation (necessary for the generation and selection of high-affinity antibodies) require the persistence of antigen or repeated exposure to antigen. In

addition to cytokines, B cells require cell-cell contacts to initiate these processes. These contacts involve the CD40 protein on B cells and CD40L on T cells (Figure 24-40, step **3b**); these proteins are members of the TNF–TNF receptor family. Work on HIV suggests that extensive somatic hypermutation is a prerequisite for the generation of broadly neutralizing antibodies — antibodies that can neutralize a broad selection of highly variable HIV isolates. More insight into the control of somatic hypermutation will be required to understand the nature of the antigen(s) capable of eliciting such desirable antibodies as a prophylactic strategy.

Vaccines Elicit Protective Immunity Against a Variety of Pathogens

Arguably the most important practical application of immunological principles is the vaccine. Vaccines are materials that are designed to be innocuous but that can elicit an immune response for the purpose of providing protection against a challenge with the virulent version of a pathogen (Figure 24-41). It is not always known why vaccines are as successful as they are, but in many cases, the ability to raise antibodies that can neutralize a pathogen (viruses) or that show microbicidal effects (bacteria) are good indicators of successful vaccination.

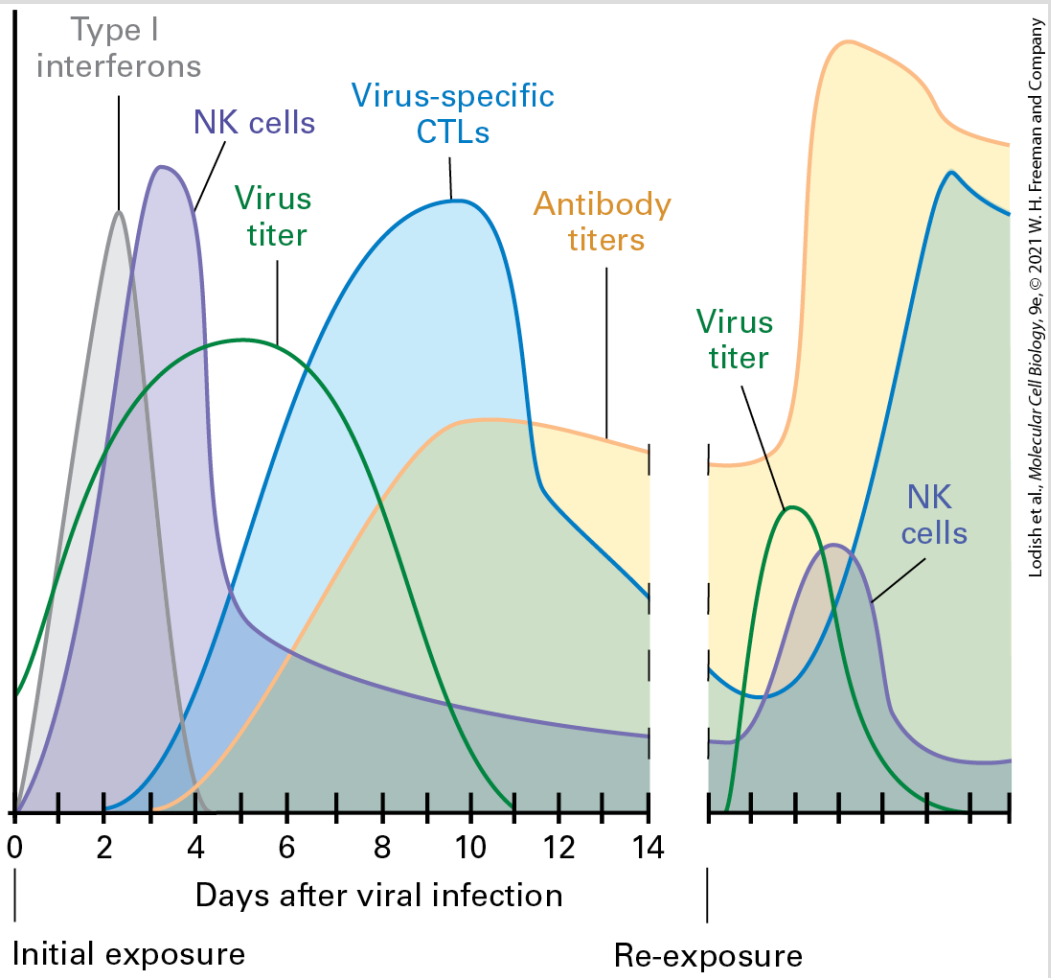


FIGURE 24-41 Time course of a viral infection. The initial antiviral response, seen when the number of infectious particles rises, includes activation of natural killer (NK) cells and production of type I interferons. These responses are part of the innate immune response. The production of antibodies and the activation of cytotoxic T cells (CTLs) follow, eventually clearing the infection. Re-exposure to the same virus leads to more rapid and more pronounced production of antibodies and to more rapid activation of cytotoxic T cells. A successful vaccine induces an immune response similar in some respects to that following initial exposure to a pathogen, but without causing significant symptoms of disease. If a vaccinated person is subsequently exposed to the same pathogen, the adaptive immune system is primed to respond quickly and strongly.

Several strategies can lead to a successful vaccine. Serial passage of a pathogen in tissue culture or from animal to animal often leads to *attenuation*; the molecular basis of attenuation is not well understood. Vaccines may be composed of live attenuated variants of more virulent pathogens. The attenuated version of the pathogen causes a mild form of the disease or causes no symptoms at all. However, by recruiting all the component parts of the adaptive immune system, such live attenuated vaccines can elicit protective levels of antibodies. These antibody levels may wane with advancing age because the lymphocytes responsible for immunological memory may have a finite life span, so repeated immunizations (booster injections) are often required to maintain full protection. Live attenuated vaccines are in use against flu, measles, mumps, and tuberculosis. In the latter case, an attenuated strain of the mycobacterium that causes the disease is used (Bacille Calmette-Guerin; BCG). Although live attenuated poliovirus was used as a vaccine until recently, its use was discontinued because the risk of reemergence, by mutation, of more virulent strains of the poliovirus outweighed the benefit. Currently, killed poliovirus is used as the vaccine of choice in the United States and Europe, although live attenuated poliovirus vaccines continue to be used elsewhere.

Vaccines based on the cowpox virus, a close relative of the human variola virus that causes smallpox, have been used successfully to eradicate smallpox; the first successful elimination of an infectious disease. Attempts to achieve a similar feat for polio are nearing

completion, but socioeconomic and political factors or armed conflict often complicate the administration of vaccines, leading to reemergence of the disease, as seen recently in Asia.

Spurious claims that vaccines may cause autism have led to skepticism among some people regarding the benefits of childhood vaccinations against serious childhood diseases such as rubella and mumps. The number of vaccinated individuals drops when parents choose not to vaccinate their children. As a consequence, the type of protection referred to as *herd immunity* wanes and exposes entire populations to the resurgence of diseases easily controlled by childhood vaccination. There is no evidence that childhood vaccinations are in any way connected to the onset or severity of autism or related disorders.

The other major type of vaccine is called a *subunit vaccine*. Rather than attenuated or killed pathogen, only one or several of the pathogen's components are used to elicit immunity. In certain cases, this approach is sufficient to afford lasting protection against a challenge with the live, virulent source of the antigen used for vaccination. It has been successful in preventing infections with the hepatitis B virus. The commonly used flu vaccines are composed mainly of the envelope proteins neuraminidase and hemagglutinin (see Figure 3-10); these vaccines elicit neutralizing antibodies. For the vaccine against human papillomavirus HPV 16, a virus that causes cervical cancer, viruslike particles composed of the virus's

capsid structural proteins but devoid of the viral DNA are generated; these particles are noninfectious, yet in many respects mimic the intact virion. The HPV vaccine now licensed for use in humans is expected to reduce the incidence of cervical cancer in susceptible populations by perhaps as much as 80 percent, the first example of a vaccine that prevents a particular type of cancer.

From a public health perspective, cheaply produced and widely distributed vaccines are formidable tools for preventing or even eradicating communicable diseases. Current efforts are aimed at producing vaccines against diseases for which no other suitable therapies are available (Ebola virus, coronaviruses such as the SARS-CoV-2 virus responsible for the 2020 pandemic) or where socioeconomic conditions have made the distribution of drugs problematic (malaria, HIV). With a more complete understanding of how the immune system operates, it should be possible to improve on the design of existing vaccines and extend these principles to diseases for which no successful vaccines are currently available. A remaining challenge is the massive genetic variation that pathogens can acquire: the error-prone reverse transcriptase of HIV introduces mutations with every successive cycle of viral replication, creating untold numbers of variants. Viable variants that carry such mutations may escape detection by the immune system. The design of effective vaccines, not only against HIV but also influenza virus, must therefore be focused on those structural elements that do not tolerate mutations and that can be seen by the adaptive immune system. ■

The Immune System Defends Against Cancer

The immune system not only defends against the immediate consequences of infection with pathogens but may also help in warding off cancer. As we have seen, the adaptive immune system is purged of many self-reactive B and T cells by negative selection. Self-reactive lymphocytes that escape this process are usually silenced because they are not provided with the appropriate co-stimulatory signals. Conditions that lead to severe immunosuppression, such as a genetic lesion in the RAG somatic recombination machinery or immunodeficiency caused by infection with HIV, confer an increased risk of cancer, not only for cancers caused by transforming viruses but also for those elicited by carcinogens. This observation establishes a role for the immune response in keeping precancerous cells in check.

Recall that B and T cells require activation not only via their antigen-specific receptors, but also by a second, co-stimulatory signal (e.g., binding of CD28 on T cells to CD80 or CD86 on APCs; see Figure 24-35). Withholding of this co-stimulatory signal silences, or anergizes, any self-reactive lymphocyte that escaped deletion in the course of T-cell selection. Because tumor cells are exceedingly similar to the progenitors that give rise to them (see Chapter 25), with only those few mutations ("driver mutations") required to cause cancer, it is not immediately obvious how immune recognition aids in the

eradication of (pre) malignant cells before they have chance to grow into larger tumors.

Somatic Mutations Allow Immune Recognition of Tumor Cells

Nonetheless, somatic mutations in the developing tumor cell — even those that are adventitious and do not directly contribute to causing cancer — can create so-called neo-antigens that may be recognized by antigen-specific receptors on T cells. Chemical mutagens, as experienced by heavy smokers who expose their lungs to tobacco products, not only cause mutations in genes that then drive tumorigenesis, but also cause mutations in other genes (passenger mutations), providing a rich spectrum of altered gene products to which the developing immune system was never exposed. If there is no immune tolerance for these mutagen-induced neo-antigens, they may serve as targets for recognition by the host's immune system.

Often the deregulation of gene expression that accompanies a transformed phenotype results in re-expression of differentiation antigens characteristic of a much earlier developmental state. If these differentiation antigens were expressed at a stage of development when the immune system had not yet fully matured, immune tolerance for such differentiation antigens may not have been established. These antigens may therefore be targets for immune recognition. Finally, the levels of certain gene products may no longer be properly regulated in cancer cells and may begin

to exceed a threshold required for immune recognition, notwithstanding the fact that they are proteins normally made by the host, albeit at much reduced levels.

In summary, because cancer can be considered a disease caused first and foremost by mutations, whose effects are modified by epigenetic events (see Chapter 25), there is the potential for immune recognition of cancer cells. Furthermore, in much the same way that an immune response against a virus or a bacterium can result in the outgrowth of variants that are no longer recognized by the immune system, selective pressure exerted by the immune system may also lead to variants of cancer cells that have lost expression of a possible tumor antigen. For example, many colon cancers show reduced levels, if not complete loss of expression, of class I MHC products, and are thus rendered invisible to cytotoxic T cells.

Antibodies Against CTLA4, PD-1, and PD-L1 Are the Basis of Cancer Immunotherapy

The tumor microenvironment is composed of stromal cells: fibroblasts and myeloid-derived cells, including macrophages. Lymphocytes are known to invade tumors, as do neutrophils. The interplay between tumor cells and the microenvironment in which these cells reside can create immunosuppressive conditions that preclude a successful anti-tumor immune response, even if the tumor cells themselves are sufficiently antigenically distinct to be recognized as such. Important players in establishing an

immunosuppressive environment are molecules now referred to as *immunological checkpoints*, such as CTLA4, the expression of which on T cells increases as T cells undergo full activation and maturation. Normally, CTLA4 would play a role in terminating an immune response, but its expression on the surface of tumor-specific T cells would compromise their anti-tumor activity. Moreover, the thymus and peripheral lymphoid compartments produce Treg cells, which are capable of suppressing the activity of other T cells. An abundance of Treg cells would keep other T cells from attacking a tumor. By the same logic, these Treg cells may keep potentially self-reactive T cells in check as a means of preventing the onset or reducing the severity of autoimmune disease.

Two key inhibitors of immune responses are PD-1 on T cells and PD-L1 on T-cell targets. Binding of PD-L1 on a tumor cell to PD-1 on a T cell that recognizes a tumor neo-antigen inhibits T-cell functions, including its ability to kill the tumor cell. A spectacular breakthrough in the treatment of cancer is the use of antibodies that target the inhibitory T-cell proteins CTLA4 and PD-1 proteins as well as PD-L1, often present on certain tumor cells. Some 30 percent of patients with metastatic melanoma, refractory to other forms of therapy, respond to treatment with antibodies against CTLA4, PD-1, or PD-L1, which has resulted in complete remissions and even cures. Similar approaches are underway to treat different forms of lung and renal cancer. Treatment with anti-CTLA4 apparently broadens the repertoire of cytotoxic T cells capable of recognizing tumor antigens as well as suppressing the activity of Treg cells. Treatment with anti-PD-1 or anti PD-L1 enhances T-cell recognition of tumors. It is perhaps ironic that smokers with the heaviest exposure to tobacco products may benefit the most from these forms of treatment because of the high mutational load in their cancers. ■

KEY CONCEPTS OF SECTION 24.6

Collaboration of Immune-System Cells in the Adaptive Response

- Antigen-presenting cells such as dendritic cells require activation by means of signals delivered to their Toll-like receptors and other pattern recognition receptors. These receptors are broadly specific for macromolecules produced by bacteria and viruses.

Their engagement activates the NF-κB signaling pathway, whose outputs include the synthesis of inflammatory cytokines (see Figure 24-37).

- Upon activation, dendritic cells become migratory and move to lymph nodes, ready for their encounter with T cells. Activation of dendritic cells also increases their display of MHC-peptide complexes and expression of co-stimulatory molecules required for initiation of a T-cell response.
- B cells require the assistance of T cells to execute their full differentiation program into plasma cells. Antigen-specific help is provided to B cells by activated T cells, which recognize class II MHC–peptide complexes on the surfaces of B cells. These B cells generate the relevant MHC-peptide complexes by internalizing antigen via BCR-mediated endocytosis, followed by antigen processing and presentation via the class II MHC pathway (see Figure 24-40).
- In addition to cytokines produced by activated T cells, B cells require cell-cell contact to initiate somatic hypermutation and class-switch recombination. This contact involves CD40 on B cells and CD40L on T cells.
- Important applications of the immunological concept of collaboration between T and B cells include vaccines. The most common forms of vaccines are live attenuated viruses or bacteria, which can evoke a protective immune response without causing pathology, and subunit vaccines.
- The adaptive immune system can sometimes distinguish between normal cells and their cancerous counterparts. What complicates immune-system detection of cancer cells are the often relatively minor differences between normal and transformed cells.
- Immunological checkpoints dampen the activity of antigen-specific T cells under normal circumstances as a means of turning off or controlling an immune response.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter.

- Perspectives for the Future
- Classic Experiment 24.1: Two Genes Become One: Somatic Recombination of Immunoglobulin Genes
- Analyze the Data
- Chapter References
- Additional study tools, including videos, animations, and quizzes

Key Terms

antibody
antigen
antigen-presenting cell (APC)
antigen processing and presentation
autoimmunity
B cell
B-cell receptor (BCR)
chemokine
complement
cytokine
dendritic cell

diversity
epitope
Fc receptor (FcR)
immunity
immunoglobulin
immunoglobulin fold
inflammasome
inflammation
interleukin
isotype
lymphocyte
macrophage
major histocompatibility complex (MHC)
memory
monoclonal antibodies
natural killer (NK) cell
neutrophil
opsonization
paratope
phagocyte
T cell
T-cell receptor (TCR)
tolerance
Toll-like receptor (TLR)
transcytosis

Review the Concepts

1. Describe the ways in which each of the following pathogens can disarm their host's immune system or manipulate it to their own advantage:
   a. Pathogenic strains of *Staphylococcus aureus*, a Gram-positive bacterium that replicates extracellularly and secretes cytotoxic products (toxins)
   b. Enveloped viruses such as HIV (see Section 5.7)
2. Trace the movement of leukocytes as they perform their functions throughout the body.
3. Identify the major mechanical and chemical defenses that protect internal tissues from microbial attack.
4. Compare and contrast the classical pathway of complement activation with the alternative pathway.
5. What is opsonization? What is the role of antibodies in this process?
6. In B cells, what mechanism ensures that only rearranged V genes are transcribed?
7. What prevents further rearrangement of immunoglobulin heavy-chain gene segments in a pre-B cell once a productive heavy-chain rearrangement has occurred?
8. How and why do B cells undergo a class switch from producing IgM antibodies to any of the other Ig isotypes?
9. What biochemical mechanism underlies affinity maturation of the antibody response?

10. Compare and contrast the structures of class I and class II MHC molecules. What kinds of cells express each class of MHC molecule? What are the functions of these cells?
11. Describe the six steps in antigen processing and presentation via the class I MHC pathway.
12. Describe the six steps in antigen processing and presentation via the class II MHC pathway.
13. What prevents self-reactive T cells from leaving the thymus?
14. Explain why T-cell–mediated autoimmune diseases are associated with particular alleles of class II MHC genes.
15. How are antigen-presenting cells and helper T cells involved in B-cell activation?
16. Outline the events in the innate and adaptive immune responses, from when a pathogen invades to clearance of the pathogen.
17. Define passive immunization and give an example.
18. How would you design a vaccine that protects against HIV infection without the possibility of infecting the patient?
19. The annual flu shot is composed of either live attenuated influenza virus or influenza subunits (the envelope proteins neuraminidase and hemagglutinin). How does the annual flu shot protect you against infection?
20. Design a laboratory protocol to develop a monoclonal or polyclonal antibody against a protein of interest (see Section 4.1).
21. Consider a person without any functioning plasma cells. What effects would this condition have on the person's

adaptive immune system? Innate immune system?

CHAPTER 25

Cancer

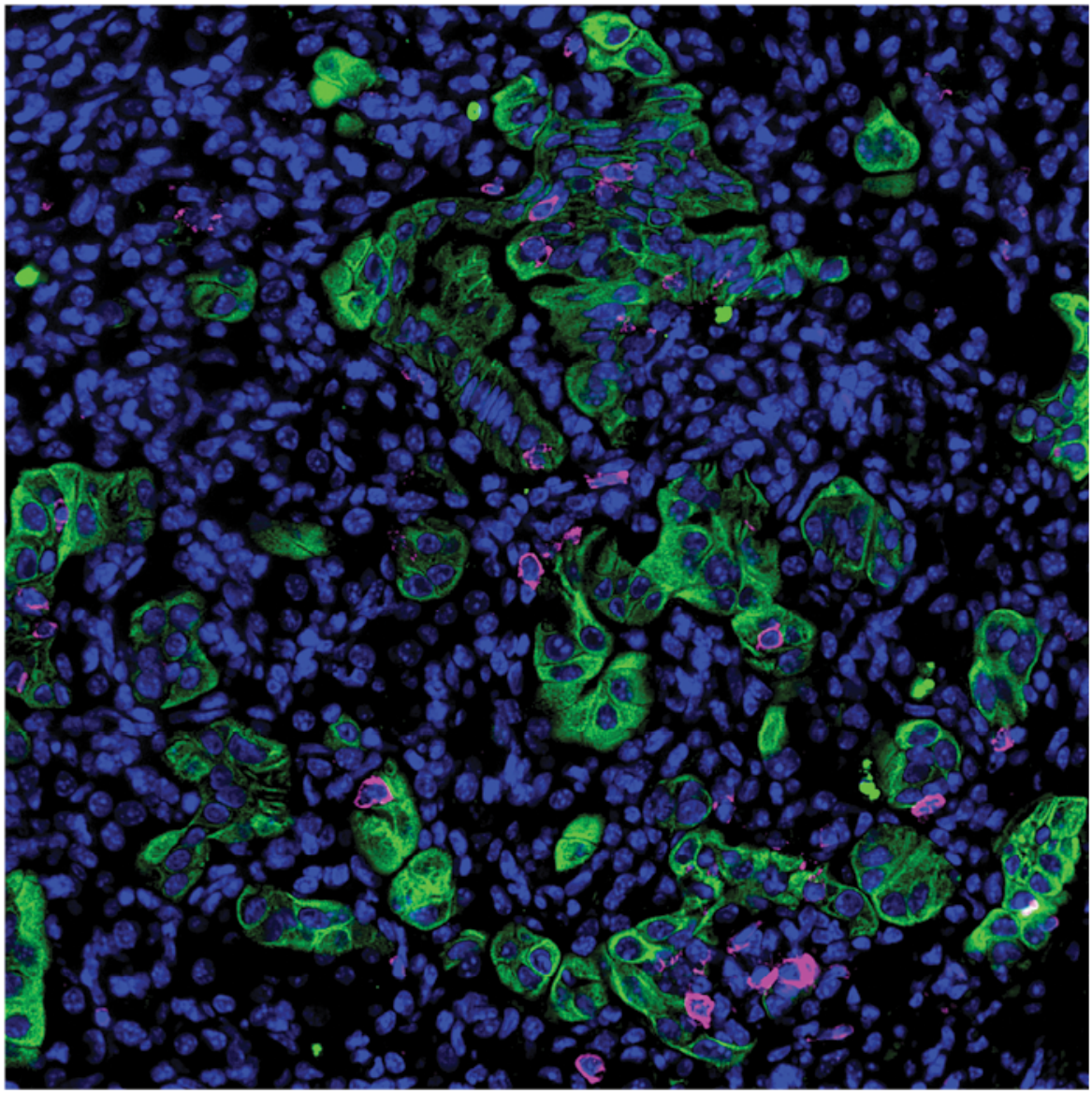


Image provided by Elen Torres and Stefani Spranger, MIT.

Cells of the immune system often interact with cancer cells within the tumor microenvironment. This section of a lung adenocarcinoma shows all cell nuclei (blue), tumor

cells (green) and $CD8^+$ T cells (magenta). Note the tendency of the T cells to be in contact with tumor cells.

OUTLINE

25.1 How Tumor Cells Differ from Normal Cells
25.2 Genetic and Genomic Basis of Cancer
25.3 Dysregulation of Cell Growth and Developmental Pathways Initiates Tumorigenesis
25.4 Evasion of Programmed Cell Death and Immune Surveillance Processes

Cancer, described in the simplest terms, arises as a disease when normal somatic cells of the body acquire a set of mutations that cause them to escape from the orderly patterns of tissue development and to grow and spread through the body. Because there are many ways that loss of normal developmental regulation can lead to uncontrolled growth, cancer is actually many different diseases depending upon the origin of the cancer cells. Tumors derived from epithelial tissues such as endoderm (gut epithelium) or ectoderm (skin and neural epithelia) are classified as **carcinomas**, whereas tumors derived from mesodermal tissues (muscle, bone, cartilage and connective tissue precursors) are classified as **sarcomas**. Many tumors form solid masses, but tumors of the blood and bone marrow, such as the **leukemias** and **lymphomas**, can grow as individual cells circulating in the blood. The classification of cancer types is further subdivided according to the exact cell type of origin and sometimes even according to the types of mutations that have caused dysfunction in a particular cell type. There are about 200 types of cancer based on histological examination of tumor

cells; the number of distinct types of cancer that are recognized is increasing as differences between cancer cells can be discerned at a molecular level.

Despite the wide variety of different types of cancer, all cancers are thought to arise by the same process, called **oncogenesis** or **tumorigenesis**, which begins with a single mutation in a somatic cell that eventually develops into a clone of cancer cells possessing many different oncogenic mutations. Although it is possible to imagine how one of the approximately 10^{12} dividing cells in the body could be struck by even a very rare mutational event, it seems inconceivable that a clone of cells could by chance alone acquire a whole set of highly specific mutations needed for all of the phenotypic properties of a cancer cell. The answer to this puzzle is that oncogenesis is the grim result of a version of natural selection in which cells that proliferate a bit more rapidly than their neighbors, or do not undergo apoptosis when they should, have greater opportunity to undergo further genetic changes. Most tumors develop slowly, acquiring additional oncogenic mutations by selection over a span of many years.

Figure 25-1 shows an outline for understanding how cancer progresses from an initiating mutation to a full tumor by the acquisition of successive mutations in a clone of cells successively acquiring oncogenic mutations. Usually, by the time a tumor is detected and its genome can be analyzed, so many mutations have

accumulated that it is not possible to identify with certainty the original mutation that started a cell down the path to cancer. Nevertheless, two general types of mutations are known to be capable of initiating oncogenesis. The first type of initiating mutation may affect one of the many signal transduction pathways that regulate cell growth. Such mutations could either inappropriately activate a growth-promoting pathway or could inactivate a growth-inhibiting pathway. In either case, the mutant cells would grow more rapidly than their well-regulated neighbors and thus provide more opportunity for those cells to acquire additional oncogenic mutations. The second type of initiating mutation impairs the ability of the cell to repair damage to DNA caused by a **mutagen** or by an error in DNA replication. Such mutations could be in the DNA repair enzymes themselves or could prevent the arrest of the cell at one of the cell cycle checkpoints to allow the repair to take place (see Chapter 20). As described in Chapter 5, DNA replication and cell division are necessary for damage to DNA to be converted to a mutation in the base sequence. This fundamental relationship between cell division and mutagenesis is one of the reasons why cells in actively dividing tissues are the most susceptible to acquiring oncogenic mutations and therefore generally the most prone to form tumors.

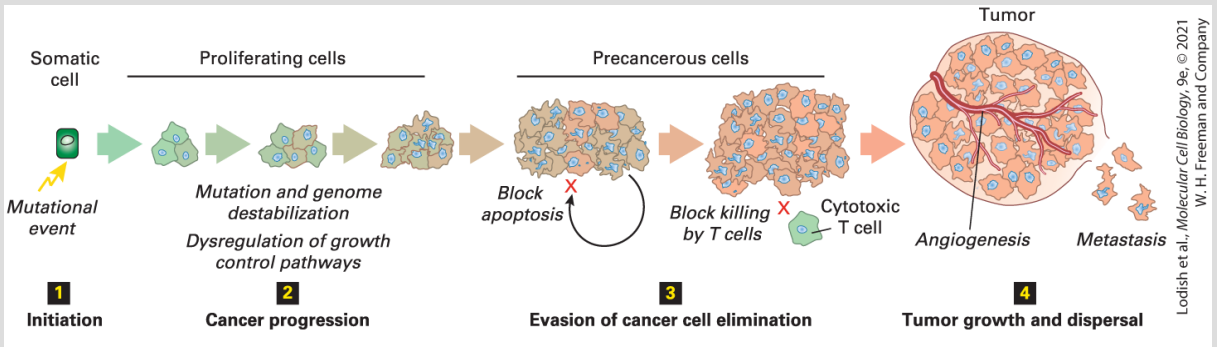


FIGURE 25-1 General framework for the evolutionary progression of cancer. The progression of cancer from a normal somatic cell to a full-grown tumor can be thought to take place in four stages. Step 1: Initiation. Except in rare cases in which cancer is caused by a singular genetic event, such as infection by a tumor virus, we usually cannot pinpoint the initiating event but a somatic mutation is usually suspected. Step 2: Cancer progression. Precancerous cells acquire mutations that dysregulate growth-control pathways, causing inappropriate cell proliferation. In addition, they acquire mutations that cause genome instability and propensity to acquire yet more mutations, mainly by loss of the DNA damage checkpoint. Most cancers progress by a process of natural selection, sequentially acquiring multiple such mutations in different pathways. Step 3: Evasion of cancer cell elimination. Precancerous cells are normally removed from the body by either apoptosis or immune-surveillance processes. To progress beyond this stage, tumor cells must acquire additional somatic mutations that allow them to evade these systems. Step 4: Tumor growth and dispersal. Continued growth of a tumor requires a blood supply and tumor cells must acquire yet more changes that promote angiogenesis. To spread throughout the body by metastasis, cells from a solid tumor must acquire the ability to migrate from their original site and adhere at a new location in the body.

The clonal descendants of an initial mutant are susceptible to acquiring additional oncogenic mutations of either a type that promotes cell proliferation or that inhibits DNA repair. These additional mutations in turn increase the propensity for yet more oncogenic mutations. Each iteration of this cycle of mutation and selection for more rapidly growing cells produce cellular

descendants exhibiting successively more uncontrolled growth and more defects in the genome. Tumors are usually detected only after they have reached 1 cm^3 in size, by which time their cells have undergone many cycles of error-prone replication and their genomes will carry many different single-base changes due to errors in replication and chromosome rearrangements due to erroneous repair of breaks in the DNA. The base sequence of cancer cell genomes usually reveals a profusion of mutations; it is often difficult to distinguish mutations that drive the oncogenic process from mutations that arose fortuitously but do not themselves contribute to oncogenic progression.

Other oncogenic mutations enable cancer cells to evade the two major systems that cause removal of aberrant cells from the body. Mammalian cells are normally able to monitor the integrity of their genome, and if they suffer a major, irreparable aberration in genome structure will enter a cell death program (see Chapter 22) that will cause them to be eliminated before they can become cancerous. The precancerous cells that do survive to become tumors have lost one or more of the key elements of the programmed cell death pathway. Another class of oncogenic mutations often found in tumors has allowed the cancer cells to evade surveillance by T cells of the immune system. As a cancer cell accumulates an increasing burden of mutations in the genome, the proteins of the cell will increasingly carry amino acid substitutions, truncations, and domain rearrangements that will eventually cause a cancer cell to

appear as foreign to the immune system. Specific mutations in cancer cells allow them to avoid immune detection and destruction.

The growth of a mass of rapidly dividing cells will reach a limit simply because the cells on the inside of the mass are restricted by diffusion for the availability of nutrients and oxygen. For a tumor to grow beyond the dimensions of several millimeters requires that the tumor cells be supplied with blood. Accordingly, a fourth type of oncogenic mutation found in solid tumors enables the tumor cells to signal for formation of blood vessels by secreting certain growth factors. In addition to the transformed cancer cells, tumors usually recruit numerous cells that are not mutated and that form differentiated structures in response to signals emanating from the cancer cells. Because tumors contain multiple cell types and structures, they more closely resemble developed organs than a clonal mass of cells. Finally, due to yet other oncogenic mutations, tumor cells often have acquired the ability to detach from the primary tumor, to enter circulating blood or lymphatic systems and adhere to a new site and form a secondary tumor in a different tissue. The capability of a tumor to spread throughout the body in this way is known as **metastasis** and is a hallmark of the deadliest forms of cancer.

This basic outline of the origin, progression, growth, and spread of cancer shows why cancer is such a difficult disease to treat. Because cancer cells are derived from normal somatic cells, it is generally not possible to kill cancer cells selectively without to some extent

killing normal cells as well. Most cancer therapy aims at killing rapidly dividing cells that cannot repair their DNA efficiently. For example, alkylating agents that cause cross-links in DNA are among the most commonly used anticancer chemotherapeutic compounds that preferentially kill cells that are actively undergoing DNA replication but that have lost the ability to repair cross-linked and broken DNA. Even the most effective anticancer compounds cause tremendous collateral damage to actively dividing, normal cells, such as the dividing cells in the hair follicles, bone marrow, and immune system, causing hair loss, anemia, and susceptibility to infections.

Another avenue to cancer therapy is to reverse the effects of oncogenic mutations by, for example, inhibiting the product of an overexpressed oncogene. This type of targeted therapy is effective at treating rare forms of cancer that arise because of a single oncogenic event. As we will see, the vast majority of cancers have acquired multiple oncogenic mutations in different pathways for cell growth, DNA repair, and programmed cell death and there are no prospects yet for a magic bullet that can simultaneously reverse the dysregulation of multiple pathways.

Currently, our understanding of how cancer develops at the cellular level offers the most promise for early diagnosis and prevention of the disease. All of the known oncogenic mutations as well as their physiological consequences can be thought of as molecular markers that could allow early detection of the presence of precancerous

cells. In the early stages of progression of the disease, precancerous cells are likely to be the most vulnerable to correction by a drug targeted to a specific pathway or by elimination by heightened immune surveillance. We know that most cancers develop in stages over a long period of time; detection of the first steps of oncogenesis should afford significant opportunities to intervene.

In this chapter, we first introduce the properties of tumor cells, illustrating how the genome, cellular metabolism, regulation of growth and proliferation, and morphology all can be altered in cancer. We will also see how the interaction of cancer cells with their environment enables development of large tumors and the spread of cancer by metastasis. We then discuss the genetic and genomic basis of cancer, how cancer can originate with inherited mutations or by somatic mutations caused by carcinogens and how the breakdown of genome maintenance functions contributes to tumorigenesis. Next we consider the general types of genetic changes affecting both growth-promoting and growth-inhibiting processes that can result in excess cell proliferation. We conclude the chapter with a discussion of how cancer cells acquire mutations that allow them to escape from programmed cell death and immune surveillance. We will end with promising new ways to treat mature tumors based on activation of the immune system to recognize and destroy cancer cells, known as *cancer immunotherapy*.

25.1 How Tumor Cells Differ from Normal Cells

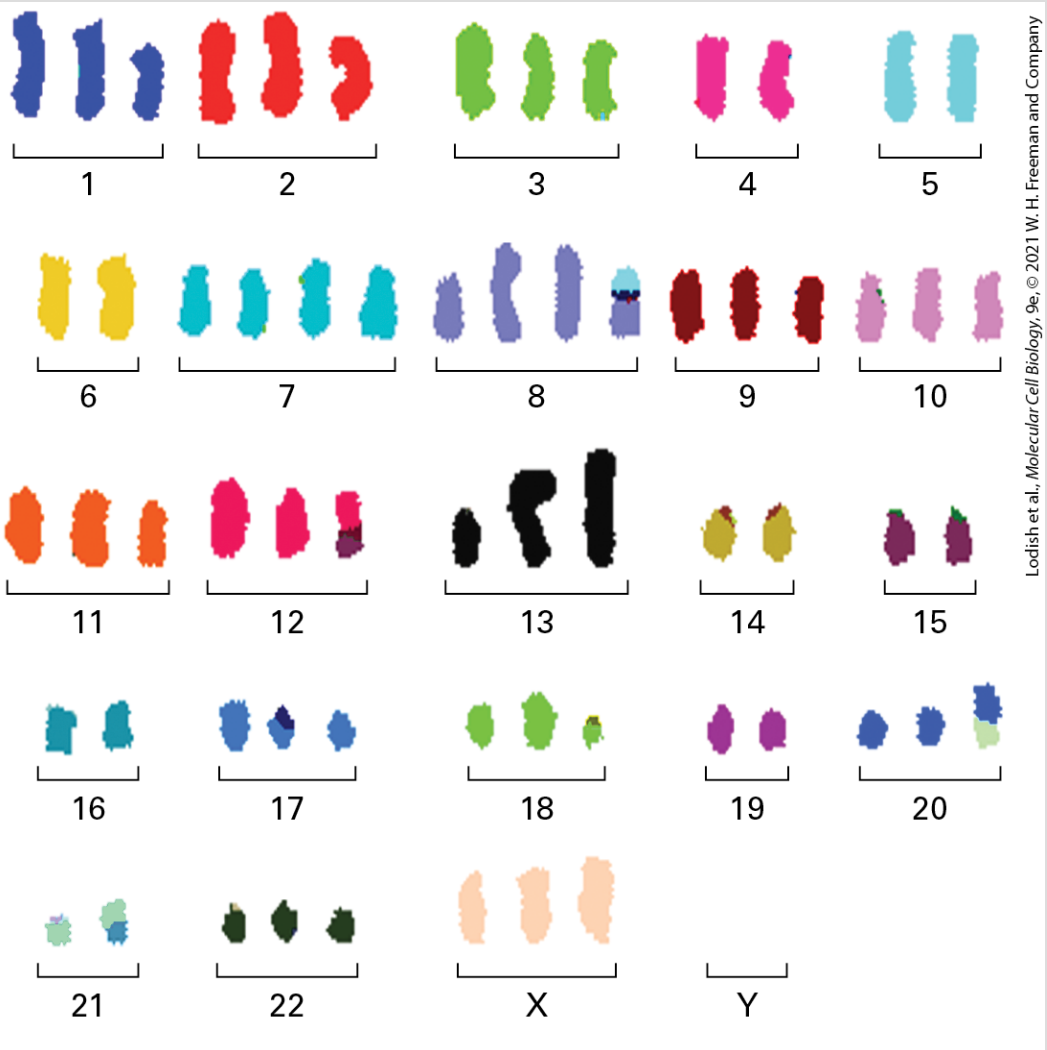
Before examining the genetic basis of cancer in detail, let's consider the general properties of tumor cells that distinguish them from normal cells. The change from a normal cell to a cancer cell commonly involves multiple steps, each one adding properties that make cells more likely to grow into a tumor. The genetic changes that underlie oncogenesis alter several fundamental properties of cells, allowing those cells to evade normal growth controls, modulating their tissue microenvironment, and ultimately conferring the full cancer phenotype (see Figure 25-1). Cancer cells acquire a drive to proliferate that does not require an external inducing signal. They fail to sense signals that restrict cell division, and they survive when they should die. They often change their attachment to surrounding cells or to the extracellular matrix, breaking loose to move away from their tissue of origin. Solid tumors characteristically outgrow their initial blood supply and become *hypoxic* (oxygen starved), so to grow to more than a small size, tumors must obtain an additional source of blood flow. They often do so by inducing the growth of blood vessels into the tumor. As cancer progresses, further adaptations are acquired by selection allowing growth and invasion of surrounding tissues, often spreading to distant sites in the body.

In this section, we describe the characteristics of cancer cells. We first discuss the changes in the cancer cell's genetic makeup that affect virtually all cellular functions, allowing the cancer cell to escape proliferation regulation and acquire the ability to divide indefinitely. We then see how the genetic changes in a tumor cell and its interactions with its environment facilitate its escape from the constraints of the tissue it was once a part of and allow it to invade neighboring tissues and colonize distant sites in the body.

The Genetic Makeup of Most Cancer Cells Is Dramatically Altered

At the turn of the twentieth century, David von Hansemann and Theodor Boveri first documented what we now know to be an almost universal feature of cancer cells: their entire genetic makeup differs dramatically from that of normal cells. Chromosomes from cancer cells stained to be visible in the light microscope often show a highly altered karyotype and large chromosomal amplifications and deletions, translocations, and aberrant numbers of chromosomes — generally too many, a condition known as *aneuploidy*. Typical cancer cells exhibit whole chromosome or chromosome arm gains and losses involving a quarter of their genome (Figure 25-2). The advent of efficient next generation DNA-sequencing methods has allowed complete cancer genome sequences to be obtained from thousands of different tumors. In

addition to gross chromosomal abnormalities, sequencing reveals single-base mutations and small, local amplifications and deletions often affecting about 10 percent of the cancer cell's genome (Figure 25-3). Perhaps the most surprising result revealed by the sequencing of cancers is the high degree of variation in mutation rates across different cancers. Mutations are rare in pediatric cancers, with substitution rates as low as a few hundred base changes per genome, but those rates may be as high as 500,000 base changes per genome in mutagen-induced cancers such as certain lung cancers and melanomas. As we will see in Section 25.2, the frequency and type of genomic mutations in a given tumor can often be traced to the presence of mutations in DNA repair pathways. These genetic changes affect virtually all aspects of cellular homeostasis, proliferation, tissue organization, and migratory properties as well as survival and proliferation at foreign sites in the body.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 25-2 Cancers often have highly abnormal karyotypes. Image of chromosomes obtained from an SW403 colorectal adenocarcinoma cell line. Chromosomes are arranged and color coded according to labeling with probes specific for each of the 24 chromosomes. Two characteristics of the cancer cells are evident. First, a normal karyoptype would contain exactly two copies of each chromosome and the cancer cells contain one or two extra copies of many of the chromosomes. Second, many chromosomes are composites of pieces from different chromosomes, which result from translocation events.

[From W. M. Abdel-Rahman et al., 2001, "Spectral Karyotyping Suggests Additional Subsets of Colorectal Cancers Characterized by Pattern of Chromosome Rearrangement." *Proc. Nat'l Acad. Sci. USA* **98**(5):2538–2543, Fig. 3c. Copyright (2001) National Academy of Sciences, U.S.A.]

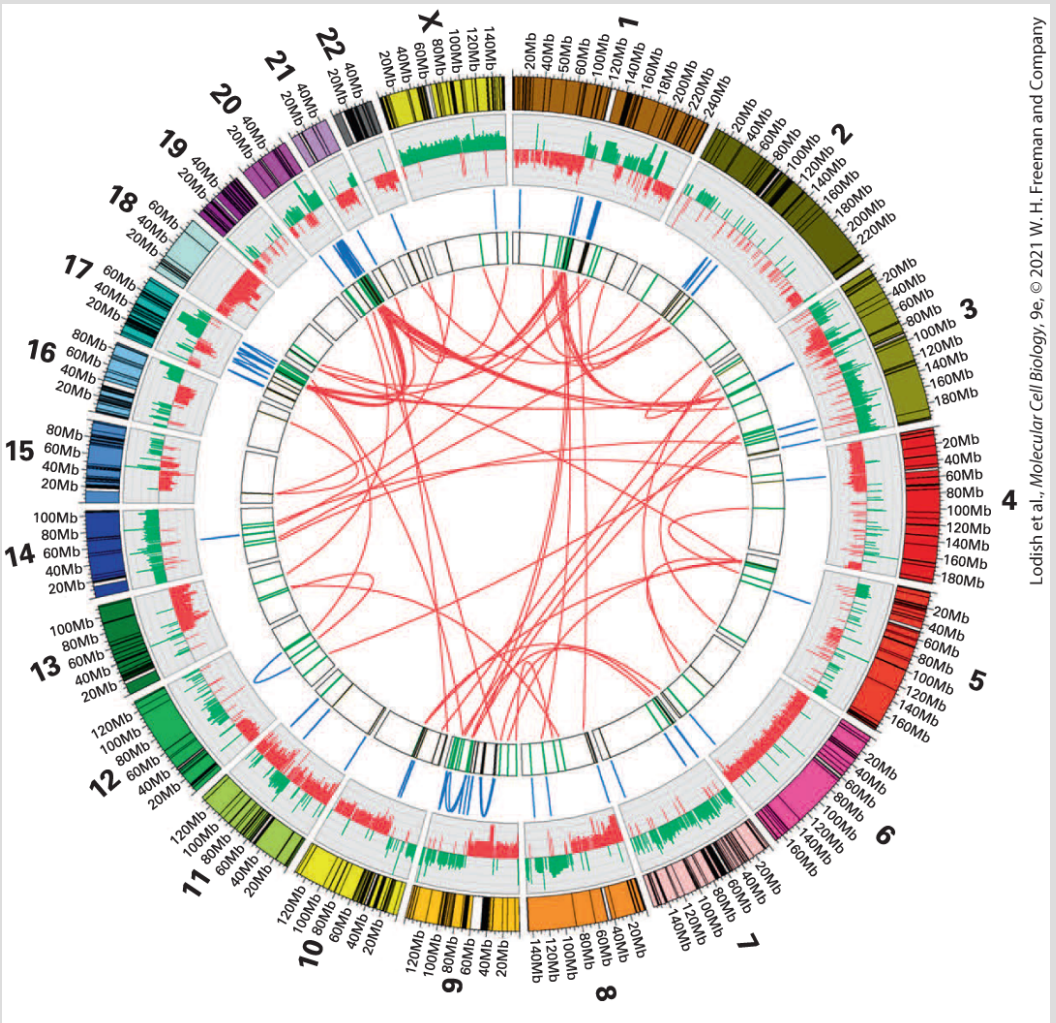


FIGURE 25-3 Cancer genome sequencing reveals thousands of sequence alterations. The complete genome sequence of the MCF-7 adenocarcinoma breast cancer cell line displayed as a Circos plot. The 23 chromosomes are arranged around the outside ring. In the next ring, DNA copy number is represented by increases (green) or decreases (red) from the normal 2n (diploid) number of copies, showing that most of the MCF-7 genome has an abnormal copy number. Intrachromosomal and interchromosomal rearrangements are represented respectively by blue and red arcs. A total of 157 chromosomal breakpoints were found.

[Oliver A. Hampton et al., 2009, *Genome Res.* **19**:167–177; Copyright © 2009 by Cold Spring Harbor Laboratory Press.]

Uncontrolled Proliferation Is a Universal Trait of Cancer

In normal tissues, cell proliferation is a highly controlled process. Growth-promoting factors are released in a highly controlled fashion to ensure that target cells proliferate only as needed to replenish a tissue. A universal feature of cancer cells is that they have acquired oncogenic mutations that allow them to escape these tight controls and to proliferate continuously and indefinitely. Here we will briefly introduce three of the most common *oncogenic mutations* as illustrative examples, and in Sections 25.3 and 25.4, we will develop a more comprehensive picture of the ways that different mutations can enable cancer cell proliferation.

The **Ras protein** described in Chapter 16 acts a GTPase switch in many of the signaling pathways that receive growth stimulatory signals via a receptor tyrosine kinase. Ras activates the MAP kinase signal transduction pathway that in turn activates many transcription factors regulating aspects of cell proliferation. The first oncogenic mutation to be described is a point mutation in the *Ras* gene that causes constitutive activation of the MAP kinase pathway. This pathway normally is only activated when a growth factor is present in the cell's environment. Thus precancerous cells that carry an oncogenic Ras mutation proliferate as if they are continually receiving a growth signal.

Most human tumors have oncogenic mutations in **p53** or in proteins that regulate p53 activity. Normally, p53 plays a major role in maintaining the integrity of the genome. As described in Chapter 19, when DNA damage is first detected in a cell, p53 is responsible for causing cell cycle arrest before M phase, giving repair enzymes sufficient time to correct the damage before the cell cycle continues. If the DNA damage cannot be repaired, p53 will eventually activate apoptosis to ensure that cells with a compromised genome are eliminated from the body. Precancerous cells that have insufficient p53 activity continue to divide even when their DNA is badly damaged. Such unchecked cell divisions destabilize the genome and greatly accelerate the acquisition of additional oncogenic mutations.

A third example of oncogenic mutations that allows uncontrolled cell proliferation commonly found in tumors are DNA rearrangements that cause increased expression of **telomerase**, the enzyme responsible for adding telomeres, short tandem DNA sequences, to the ends of linear chromosomes. As discussed in Chapter 7, telomeres become shorter every time a chromosome is replicated. Most human somatic cells produce only a small amount of telomerase and uncontrolled proliferation would eventually lead to extensive shortening of telomeres, which is recognized by the cell as a double-strand break that consequently triggers cell cycle arrest and apoptosis. Tumor cells overcome this fate by up-regulating telomerase expression to produce a state of replicative immortality.

Cellular Housekeeping Functions Are Fundamentally Altered in Cancer Cells

Cancer cells can often be distinguished from normal cells by microscopic examination. They are usually less well differentiated than normal cells. Cancer cells frequently exhibit the characteristics of non-cancerous rapidly growing cells: a high nucleus-to-cytoplasm ratio, prominent nucleoli, an increased frequency of mitotic cells, and relatively little specialized structure.

Tumor cells differ from normal cells not only in their appearance, but in their entire protein composition. Losses and gains of whole chromosomes or chromosome parts are characteristic of cancer cells and have a profound effect on the protein composition of the cell and hence on many cellular functions. Imbalances in the composition of multiprotein particles cause many proteins to remain only partially folded; this in turn induces a stress response, similar to that caused by raising the cell to an elevated temperature (see Chapter 21), aimed at offsetting these protein imbalances. As a result, cancer cells rely heavily on protein-folding chaperones and degradation mechanisms, including proteasomes, for their survival as a direct result of their dramatically altered chromosome composition.

Another hallmark of cancer cells is their dependence on an energy-generating pathway that resting cells only use if they are deprived of

oxygen. Most normal differentiated cells rely on relatively efficient mitochondrial oxidative phosphorylation to satisfy their energy needs. Cells metabolize glucose to carbon dioxide by oxidation of pyruvate through the tricarboxylic acid (TCA) cycle in the mitochondria (see Chapter 12). Only under anaerobic conditions do normal resting cells undergo anaerobic glycolysis and produce large amounts of lactate. Most cancer cells, however, rely on glycolysis for energy production irrespective of whether oxygen levels are high or low (Figure 25-4). The use of glycolysis to produce energy even in the presence of oxygen, called *aerobic* glycolysis, was first discovered in cancer cells by the biochemist Otto Warburg and is therefore called the **Warburg effect**. Aerobic glycolysis generates only 2 ATP molecules per molecule of glucose and appears to be much less efficient than oxidative phosphorylation, which generates up to 36 molecules of ATP per molecule of glucose. However, by not using the TCA cycle for generating energy, proliferating cells can instead utilize the molecules in the TCA cycle and their precursors to synthesize nucleotides, amino acids, and lipids. This rewiring of glucose metabolism diverts the flow of metabolic intermediates away from ATP synthesis and toward the production of cellular building blocks, a balance more favorable for rapid cell growth.

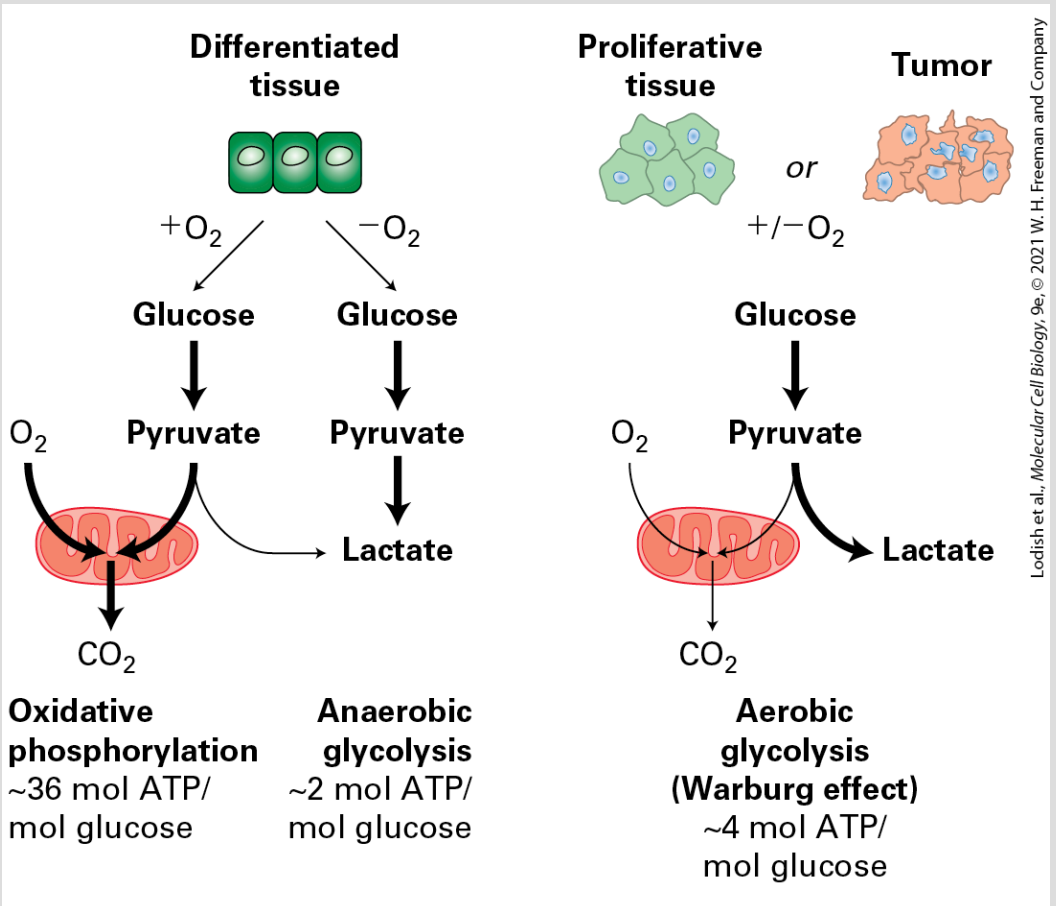


FIGURE 25-4 Energy production in cancer cells by aerobic glycolysis. In the presence of oxygen, nonproliferating (differentiated) cells metabolize glucose into pyruvate via glycolysis. Pyruvate is then transported into mitochondria, where it is fed into the TCA cycle. Oxygen is required as the final electron acceptor during oxidative phosphorylation. Thus when oxygen is limiting, cells metabolize pyruvate into lactate, allowing glycolysis to continue by cycling NADH back to NAD^+. Cancer cells and proliferating cells convert most glucose to lactate regardless of whether oxygen is present or not. The production of lactate in the presence of oxygen is called *aerobic* glycolysis.

[Data from M. G. Vander Heiden, L. C. Cantley, and C. B. Thompson, 2009, *Science* **324**:1029.]

Not only do cancer cells rewire their metabolic pathways, but they also produce increased levels of some metabolites and metabolic by-

products. These so-called *onco-metabolites* directly or indirectly alter patterns of gene expression, including genes involved in blood vessel formation.

Cancer Cells Exhibit Altered Cell-Cell Interactions to Form Heterogeneous Organs

Many human cell types grow in well-ordered sheets. Cells do not continue to divide and grow outside the plane of the sheet because of a phenomenon of *contact inhibition*, which causes arrest in the G_0 phase of the cell cycle when a cell is completely surrounded by other cells. Precancerous cells have lost contact inhibition, are less adherent and more rounded than normal cells, and will continue to divide as a three-dimensional cluster of cells (a *focus*) that can be recognized under a microscope. Adhesion molecules such as E-cadherin, cell polarity factors, actin cytoskeleton regulators, and the Hippo pathway (see Chapter 21) all function in mediating cell cycle arrest when cell-cell contacts are established. However, the exact mechanisms whereby this occurs, and how these pathways are disrupted in cancer, remain to be worked out.

Not all tumors are made up of uniform cells, even if they originated from a single initiating cell. In some types of tumors, for example, only certain tumor cells, called **cancer stem cells**, are capable of

seeding a new tumor. Within these tumors, some cells cease dividing, while others can continue cancerous growth. The latter, of course, are the most dangerous and the most important to destroy with anticancer treatments. Cancer stem cells are thought to give rise to some cells with high replicative capabilities and others with more limited replicative potential. The origins of these cancer stem cells are not clear. In some cancers, a normal tissue stem cell may give rise to the cancer stem cells. In others, dedifferentiation of terminally differentiated cells to form progenitor cells may give rise to cancer stem cells. Irrespective of their origin, cancer stem cells share gene expression signatures with normal tissue stem cells, leading to their designation as stem cell-like cells.

The immediate environment of a tumor — the **tumor microenvironment** — contributes to the heterogeneity of cells within the tumor, influencing the behavior of the cancer stem cells and the tumor cells in general. Some neighboring cells may be more conducive to tumor growth than others. The importance of the tumor microenvironment extends to one of the most common environmental influences on a tumor cell: inflammatory cells. It is now widely accepted that cells of the immune system interact with the tumor. As we detail in Section 25.4, $CD8^+$ cytotoxic T lymphocytes and natural killer cells surround and often migrate into the tumor, where they can inhibit tumor formation. Mice deficient in these and other components of the immune system are more prone to carcinogen-induced tumors than normal mice. These

findings lead to the idea that the immune system eliminates cancer cells. Current information about how cancer cells escape this immune surveillance is discussed in Section 25.4.

More and more evidence is mounting that immune-system cells can also have tumorigenic properties. It has been known for a long time that cancers frequently arise at sites of injury or chronic infection. It is estimated that up to 20 percent of cancers are linked to chronic infection. For example, persistent infection of the stomach by the bacterium *Helicobacter pylori* is associated with gastric cancer. Crohn's disease, an inflammatory autoimmune disease that affects the intestines, is associated with colon cancer. Infection with hepatitis B or C viruses increases the risk of a form of liver cancer, hepatocellular carcinoma. Immune-system cells migrate to sites of injury or infection and produce growth factors, thereby stimulating tumor cell proliferation. They also produce factors to induce the growth of blood vessels, which — as we will discuss next — is an essential aspect of tumor growth and dissemination to distant sites.

Tumor Growth Requires Formation of New Blood Vessels

Tumors must recruit new blood vessels in order to grow to a large size. In the absence of a blood supply, a tumor can grow into a mass of about 10^6 cells, roughly a sphere 2 mm in diameter. At this point,

division of cells on the outside of the tumor mass is balanced by death of cells in the center from an inadequate supply of nutrients. Such growth-limited tumors, unless they secrete hormones, cause few problems. However, most tumors induce the formation of new blood vessels that invade the tumor and nourish it, a process called **angiogenesis**. This complex process requires several discrete steps: degradation of the basement membrane that surrounds a nearby capillary, migration of endothelial cells lining the capillary into the tumor, division of these endothelial cells, and formation of a new basement membrane around the newly elongated capillary.

Many tumors produce growth factors that stimulate angiogenesis; other tumors induce surrounding normal cells to synthesize and secrete such factors. Basic fibroblast growth factor (β-FGF), transforming growth factor α (TGF-α), and vascular endothelial growth factor (VEGF), which are secreted by many tumors, all have angiogenic properties. New blood vessels allow the tumor to increase in size and thus increase the probability that additional harmful mutations will occur. The presence of an adjacent blood vessel also facilitates the process of metastasis.

The VEGF receptors, which are tyrosine kinases, regulate several aspects of blood vessel growth, such as endothelial cell survival and growth, endothelial cell migration, and vessel wall permeability. VEGF expression can be induced by oncogenes and by hypoxia. As detailed in Chapter 21, the hypoxia signal is mediated by hypoxia-

inducible factor 1α (HIF-1α), a transcription factor that is activated in low-oxygen conditions and that binds to and induces transcription of the VEGF gene and about 30 other genes, including many enzymes in the glycolytic pathway. An enhancement of glycolysis can in turn stimulate the growth of many cancer cells. HIF-1α activity is controlled by an oxygen sensor composed of a prolyl hydroxylase that is active at normal O_2 levels but inactive when deprived of O_2. Hydroxylation of HIF-1α causes ubiquitinylation and degradation of the transcription factor, a process that is blocked when O_2 is low. Mutations in genes encoding a subunit of the ubiquitin ligase that degrades HIF-1α at ambient oxygen levels trigger certain kidney tumors. Compounds that inhibit angiogenesis have excited much interest as potential therapeutic agents, but their success in the clinic has thus far been limited.

Invasion and Metastasis Are Late Stages of Tumorigenesis

Tumors arise with great frequency, especially in older individuals, but most pose little risk to their host because they are small and localized. We call such tumors **benign**; an example is a wart, a benign skin tumor. The cells composing benign tumors closely resemble and may function like normal cells. The cell-adhesion molecules that hold tissues together keep benign tumor cells, like normal cells, localized to the tissues where they originate. A fibrous

capsule usually delineates the extent of a benign tumor and makes it an easy target for a surgeon. Benign tumors become serious medical problems only if their sheer bulk interferes with normal functions or if they secrete excess amounts of biologically active substances, such as hormones. For example, acromegaly, the overgrowth of head, hands, and feet, can occur when a benign pituitary tumor causes overproduction of growth hormone.

In contrast to benign tumor cells, **malignant** tumor cells are able to invade nearby tissue, spreading and seeding additional tumors while the cells continue to proliferate (Figure 25-5). This ability is a major characteristic that differentiates malignant tumors from benign ones. Some malignant tumors, such as those in the ovary or breast, remain localized and encapsulated, at least for a time. When these tumors progress, however, the cells invade surrounding tissues and form metastases (Figure 25-6a).

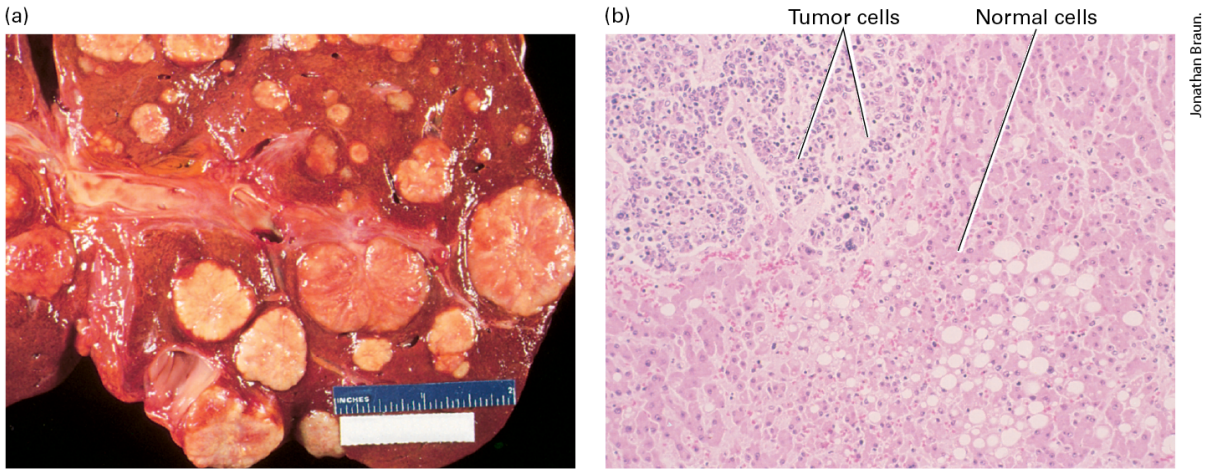
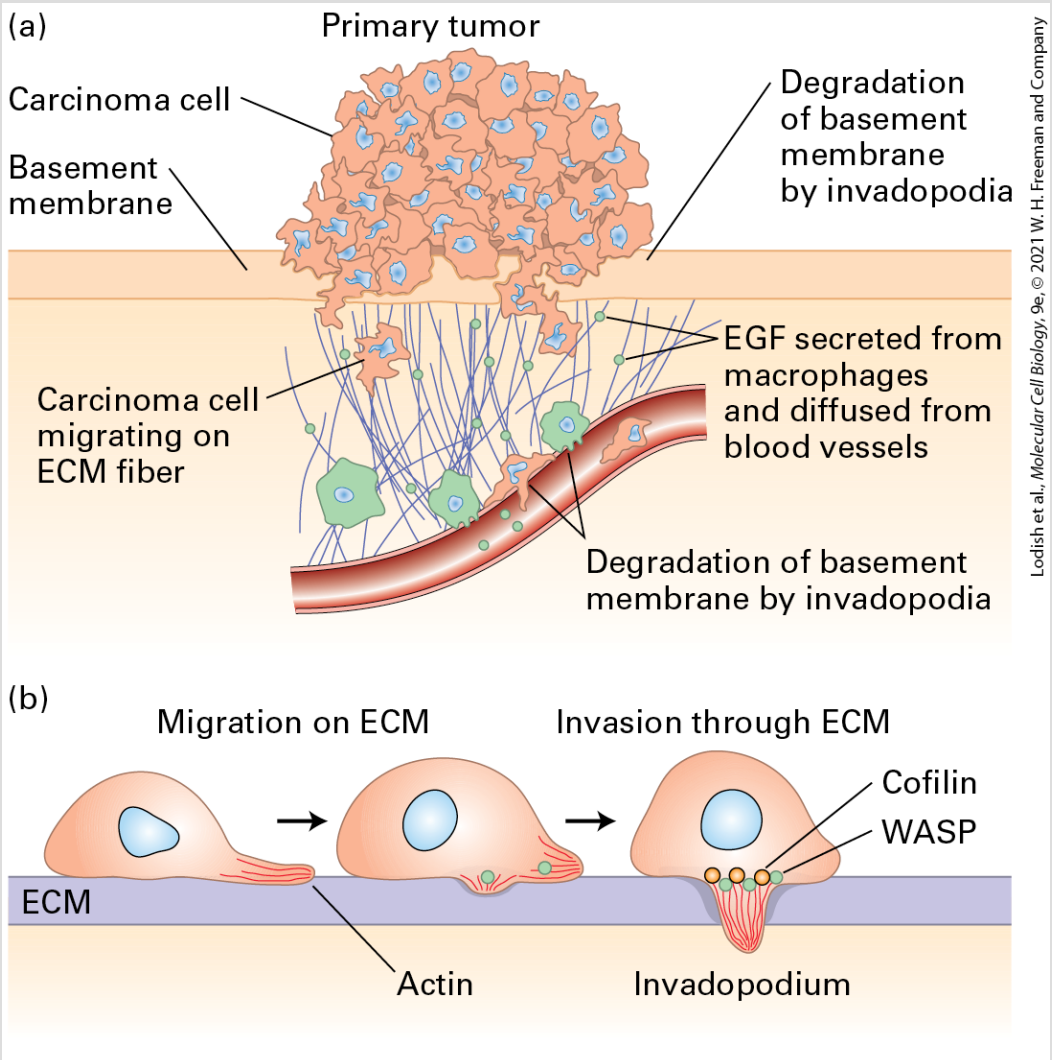


FIGURE 25-5 Gross and microscopic views of a tumor invading normal liver tissue. (a) The gross morphology of a human liver in which a metastatic lung tumor is growing. The white

protrusions on the surface of the liver are the tumor masses. (b) A light micrograph of a section of the tumor in (a), showing areas of small, dark-staining tumor cells invading a region of larger, light-staining, normal liver cells.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 25-6 Metastasis. (a) First steps in metastasis, using breast carcinoma cells as an example. Cancer cells leave the main tumor and attack the basement membrane, using extracellular matrix (ECM) fibers to reach the blood vessels. The cancer cells can be attracted by signals such as epidermal growth factor (EGF), which can be secreted by macrophages. At the blood vessels, they penetrate the layer of endothelial cells that forms the vessel walls and enter the bloodstream. (b) Carcinoma cells penetrate the extracellular matrix and blood vessel wall by extending invadopodia (actin-rich protrusions of the plasma

membrane), which produce matrix metalloproteases and other proteases to open up a path.

[Data from H. Yamaguchi, J. Wyckoff, and J. Condeelis, 2005, *Curr. Opin. Cell Biol.* **17**:559.]

Normal cells are restricted to their place in an organ or tissue by cell-cell adhesion and by physical barriers such as the *basement membrane*, which underlies layers of epithelial cells and also surrounds the endothelial cells of blood vessels (see Chapter 20 and Figure 1-26). In contrast, cancer cells have acquired the ability to penetrate basement membranes using a cell protrusion called an *invadopodium* and to migrate to distant sites in the body (Figure 25-6b). A developmental process known as the **epithelial-to-mesenchymal transition** (**EMT**) is thought to play a crucial role during the process of metastasis in some cancers. During normal development, the conversion of epithelial cells into mesenchymal cells is a step in the formation of some organs and tissues. An EMT requires distinct changes in patterns of gene expression and results in fundamental changes in cell morphology, such as loss of cell-cell adhesion, loss of cell polarity, and the acquisition of migratory and invasive properties. During metastasis, the EMT regulatory pathways are thought to be activated at the invasive front of a tumor, producing single migratory cells.

As the basement membrane disintegrates, some tumor cells enter the bloodstream, but fewer than 1 in 10,000 cells that escape the primary tumor survive to colonize another tissue and form a

secondary, metastatic tumor. Much of preventative medicine is currently focused on developing methods to identify the rare tumor cells that circulate in the bloodstream. The ability to capture these *circulating tumor cells* would not only provide a powerful and noninvasive tool for the early detection of cancer, but also their analysis could provide insights into the nature of the disease and inform treatment.

In order to produce metastases, tumor cells must not only enter the bloodstream, but also adhere to the lining of the blood vessel in a new location and migrate through it into the underlying tissue, in a process called *extravasation* (see Chapter 20). In order to seed a metastasis at a distant site, the tumor cells must not only disseminate, but also adapt to a foreign tissue environment. At least initially, metastatic tumor cells may not be well adapted to their new environment, but they are thought to evolve to survive and thrive in a foreign context. Little is known about the molecular pathways that facilitate this adaptation, but mounting evidence suggests that some environments are more conductive to cancer cell colonization than others.

Because metastasis is the most common reason for morbidity associated with cancer, much effort is being put into understanding which tumors will become metastatic and how metastasis occurs. Traditionally, the properties of tumor and normal cells have been assessed using microscopic tools, and the prognosis for many

tumors could be determined, within certain limits, from their histology. However, the appearance of cells alone has limited information content, and better ways to discern the properties of cells are desirable, both to understand tumorigenesis and to arrive at meaningful and accurate decisions about prognosis and therapy. The advent of methods to assess a tumor's patterns of RNA, protein, lipid, and metabolite production is allowing for a more detailed examination of tumor properties. Not surprisingly, primary tumors are often distinguishable from metastatic tumors by the RNAs and proteins that they produce. Analyses of global patterns of gene expression (described in Chapter 6) are now routinely used to predict patient outcomes and to determine the best course of treatment for many types of cancers. They will soon become the standard in determining treatment options. ■

KEY CONCEPTS OF SECTION 25.1

How Tumor Cells Differ from Normal Cells

- The genomes of most cancer cells acquire thousands of mutations and other genetic alterations that include point mutations, deletions, amplifications, chromosomal rearrangements and whole chromosome gains and losses.
- The changes in genetic makeup of cancer cells include oncogenic mutations that affect a wide variety of cellular functions.
- Oncogenic mutations in growth control pathways, cell cycle regulation, and in telomere maintenance are examples of how cancer cells acquire the ability to proliferate indefinitely.
- Cancer cells often exhibit altered metabolism, such as a switch of glucose metabolism to anaerobic glycolysis.
- Cancer cells lose contact inhibition, which allows them to grow in a mass outside the confines of a normal cell layer.

- Tumors are complex organs composed of different cell types that interact with their environment to obtain a maximal growth advantage.
- Tumors require angiogenesis, the formation of new blood vessels, in order to grow to a large mass.
- Cancer cells sometimes invade surrounding tissues, often breaking through the basement membranes that define the boundaries of tissues and spreading through the body to establish secondary areas of growth, a process called *metastasis*.
- Metastatic tumor cells acquire migratory properties in a process called the *epithelial-to-mesenchymal transition*.

25.2 Genetic and Genomic Basis of Cancer

Cancer can be considered to be a genetic disease in two different ways. Considering the mechanism of cancer formation, we know that tumors arise from somatic cells that have acquired a collection of mutations that give them the ability to proliferate when their neighbors cannot. In this sense, somatic mutations are the cause of cancer. When we consider the tumors themselves, we can see that they typically have acquired thousands of point mutations and genomic rearrangements; in this sense, mutations are a consequence of cancer. These contrasting relationships can be reconciled by viewing cancer progression as a form of *Darwinian evolution* acting on the growth of clones of cancer cells with cycles of mutation and selection. In this view, early in the process of cancer formation, the genome of a precancerous cell becomes highly mutable and the cells and its descendants begin to acquire random mutations. Although the vast majority of these mutations are not in genes and have little or no phenotypic effect on the cells that harbor them, a small fraction will have functional consequences. Those mutations that enable the precancerous cells to grow more rapidly, to avoid being eliminated, or to spread through the body will be selected for and will become part of the lineage that forms a tumor. Oncogenic mutations are those that are selected because they are causal drivers of cancer progression. Most modern cancer biology

research aims at identifying oncogenic mutations and understanding their underlying mechanism.

In this section, we first consider the major causes of somatic mutations in cancer — these are environmental DNA-damaging agents known as **carcinogens** and either inherited or somatic mutations that prevent DNA damage from being repaired with high fidelity. We then consider the spectrum of mutations found in cancer genomes which include a small fraction of oncogenic driver mutations. Next we consider the major types of oncogenic mutations that include virus-borne onncogenes, and gain-of-function mutations and loss-of-function mutations in the genome. Finally, we consider oncogenic mutations that exert their effect through general changes in gene expression via chromatin structure or microRNA expression.

Carcinogens Induce Cancer by Damaging DNA

The ability of chemical carcinogens to induce somatic-cell mutations results from the DNA damage they cause as well as the errors introduced into DNA that result from the mechanisms to repair that damage. The strongest evidence that carcinogens act through mutagenesis comes from the observation that there is a strong correlation between the ability of a compound to cause

mutations in DNA of cultured cells and to transform cells and induce cancer in animal models.

Although substances identified as chemical carcinogens have a broad range of chemical structures, they can be classified into two general categories. *Direct-acting carcinogens*, of which there are only a few, are highly reactive compounds that can react with and modify nucleotide bases in DNA so as to distort the normal pattern of base pairing. If the modified nucleotides are not repaired, they allow an incorrect nucleotide to be incorporated during replication. This class of carcinogens includes ethylmethane sulfonate (EMS), dimethyl sulfate (DMS), and nitrogen mustards.

In contrast, *indirect-acting carcinogens* are relatively unreactive compounds that become modified to more reactive forms by cellular *cytochrome P-450 enzymes*. P-450 enzymes normally function to add electrophilic centers, such as OH groups, to nonpolar foreign chemicals in order to solubilize them so that they can be excreted from the body. For some compounds, modification by cytochrome P-450 increases the reactivity of the compound with DNA converting it into a mutagen and a carcinogen.

Some Carcinogens Have Been Linked to Specific Cancers

Since the earliest awareness of cancer as a disease intrinsic to the cells of the body, it became clear that at least some cancers resulted from the action of environmental poisons. As early as the eighteenth century, chimney sweeps' exposure to soot was associated with a high incidence of scrotal cancer, and the use of tobacco as snuff was associated with nasal cancer. Environmental chemicals were originally associated with cancer through experimental studies in animals. The classic test for a carcinogen is to repeatedly paint a test substance on the back of a mouse and look for development of local or systemic tumors in the animal. Such assays led to the purification of a pure chemical carcinogen, benzo(*a*)pyrene, from coal tar in 1933.

Although chemical carcinogens are believed to be risk factors for many human cancers, a direct link to specific cancers has been established in only a few cases, the most important being lung cancer and the other cancers that are associated with smoking (of the larynx, pharynx, stomach, liver, pancreas, bladder, cervix, and more). Epidemiological studies first indicated that cigarette smoking was the major cause of lung cancer. The chemical *benzo(*a*)pyrene*, found in cigarette smoke as well as in coal tar, undergoes metabolic activation by cytochrome P-450 in the lungs (Figure 25-7) to form a potent mutagen that mainly causes conversion of cytosine (C) to adenine (A) bases, a transversion mutation. When applied to cultured bronchial epithelial cells, activated benzo(*a*)pyrene induces transversion mutations at many loci. *Benzo(*a*)pyrene* thus leaves a

specific fingerprint in the DNA of cancer cells from heavy smokers. Cigarette smoke is known to contain more than 60 different carcinogens, and the overall profile of mutations in lung cancer appear as a complex pattern of overlapping signatures.

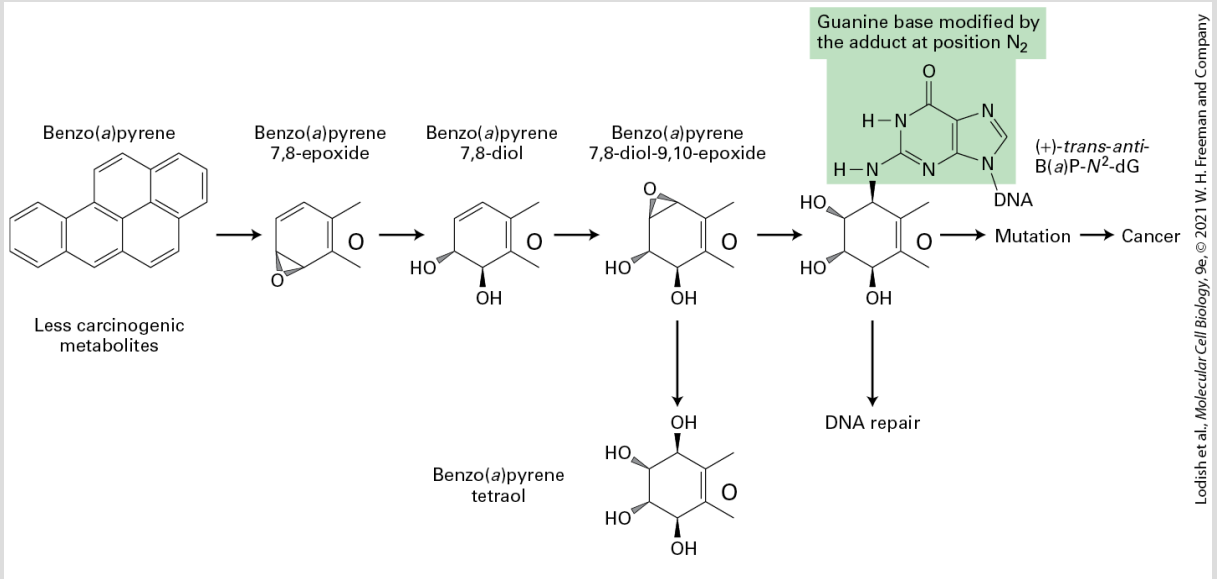


FIGURE 25-7 Enzymatic processing of benzo*(a)*pyrene to a more potent mutagen and carcinogen. Liver enzymes, particularly P-450 enzymes, modify benzo(*a*)pyrene in a series of reactions, producing 7,8-diol-9,10-epoxide, a highly potent mutagenic species that reacts with DNA primarily at the N_2 atom of a guanine (G) base. The resulting adduct, $(+)\text{-}trans\text{-}anti\text{-}\mathrm{B}(a)\mathrm{P}\text{-}N^2\text{-}\mathrm{dG}$, causes polymerase to insert an A rather than a C opposite the modified G base. The next time the DNA is replicated, a T will be inserted opposite the A, and the mutation will be complete. Horizontal arrows indicate alterations toward greater potency, while vertical arrows indicate changes in the direction of reduced toxicity. The large "O" symbol represents the rest of the multi-ring structure shown in the complete benzo(*a*)pyrene molecule at the left.

In addition to chemical mutagens, electromagnetic radiation of sufficient energy to damage DNA is also carcinogenic. X-rays and gamma rays are of a category of radiation known as ionizing

radiation because they have sufficient energy to strip electrons from atoms, generating ions. When ionizing radiation hits DNA it has sufficient energy to break a strand of DNA. X-rays have been known to cause mutations in *Drosophila* since the classic experiments of Hermann J. Muller were published in 1927. A connection between ionizing radiation and cancer was first indicated by the high incidence of cancer in the mouths of "radium girls" who worked in watch factories and ingested radium from licking the paintbrushes used to paint luminous dials. A causal relationship was solidified from a large-scale study from atomic bomb survivors who were shown to have an increased incidence of cancer, principally leukemia, in proportion to their exposure to radiation from the bomb blast. Because the repair of DNA breaks is often error prone, ionizing radiation causes a very broad spectrum of mutations.

UV light does not have the energy to break DNA strands but can interact with adjacent pyrimidine residues to create chemically altered DNA that is repaired by the nucleotide excision repair pathway (see Figure 5-18). The predominant signature of melanoma, a cancer caused by exposure to sunlight, is a preponderance of cytosine (C) to thymine (T) bases conversions, consistent with the known mutagenic effect of UV light.

Familial Syndromes That Cause Loss of DNA Repair Can Lead to Cancer

Even without exposure to any external carcinogens or mutagens, normal biological processes generate a large amount of DNA damage. That damage is due to depurination reactions, to alkylation reactions, and to the generation of reactive species such as oxygen radicals, all of which alter DNA. It has been estimated that in every cell, more than 20,000 alterations to the DNA occur each day from reactive oxygen species and depurination. The vast majority of this damage is correctly repaired by the high-fidelity DNA-repair systems that are described in Chapter 5. As we will see, the large number of mutations found in most cancers results from defects in one or more of these repair systems.

One well-recognized way for DNA repair to fail is by familial inheritance of a defect in one of the DNA-repair pathways. Table 25-1 shows the best understood familial syndromes that result from defects in DNA-repair pathways and their corresponding effects on an increase in the propensity to develop cancer. For example, people who inherit the disease xeroderma pigmentosum (XP) have a 1000-fold increase in the propensity to develop skin cancer. Seven of the eight known XP genes have been identified as components of the nucleotide excision-repair machinery, and the connection to skin cancer is clear since repair of DNA damage caused by UV light primarily requires nucleotide excision repair. Hereditary nonpolyposis colorectal cancer (HNPCC, also known as *Lynch syndrome*) causes a greatly increased propensity to develop colon cancers. HNPCC genes encode components of the mismatch-repair

system; in the absence of mismatch repair, colon cancer progresses from benign polyps to full-fledged tumors much more rapidly than usual, presumably because the precancerous cells acquire oncogenic mutations at an increased rate. Defects in mismatch repair should in theory increase the probability of oncogenic mutations in all cells in the body. The specificity of HNPCC for colon cancer is not well understood.

TABLE 25-1 • Some Human Hereditary Diseases and Cancers Associated with DNA-Repair Defects

Disease	DNA-Repair System Affected	Sensitivity	Cancer Susceptibility	Symptoms
Hereditary nonpolyposis colorectal cancer	DNA mismatch repair	UV irradiation, chemical mutagens	Colon, ovary	Early development of tumors
Xeroderma pigmentosum	Nucleotide excision repair	UV irradiation, point mutations	Skin carcinomas, melanomas	Skin and eye photosensitivity, keratoses
Bloom's syndrome	Repair of double-strand breaks by homologous recombination	Mild alkylating agents	Carcinomas, leukemias, lymphomas	Photosensitivity, facial telangiectases, chromosome alterations
Fanconi anemia	Repair of double-strand	DNA cross-linking	Acute myeloid leukemia,	Developmental abnormalities

	breaks by homologous recombination	agents, reactive oxidant chemicals	squamous-cell carcinomas	including infertility and deformities of the skeleton, anemia
Hereditary breast cancer, BRCA1 and BRCA2 deficiency	Repair of double-strand breaks by homologous recombination		Breast and ovarian cancer	Breast and ovarian cancer

SOURCE: Modified from A. Kornberg and T. Baker, 1992, *DNA Replication*, 2d ed., W. H. Freeman and Company, p. 788; J. H. J. Hoeijmakers, 2001, *Nature* **411**:366; and L. H. Thompson and D. Schild, 2002, *Mutat. Res.* **509**:49.

Double-strand breaks are particularly severe lesions because incorrect rejoining of double strands of DNA can lead to gross chromosomal rearrangements and translocations, such as those that produce a hybrid gene or bring a growth regulatory gene under the control of a different promoter or enhancer. Often the repair of such damage depends on using the homologous chromosome as a template (see Figure 5-21). The B and T cells of the immune system are particularly susceptible to DNA rearrangements caused by double-strand breaks created during rearrangement of their immunoglobulin or T-cell receptor genes, which explains the frequent involvement of these loci in leukemias and lymphomas. *BRCA1* and *BRCA2*, genes implicated in human breast and ovarian cancers, encode important components of DNA-break repair systems. As we saw in Chapter 5, broken DNA ends can normally be

repaired by recombination with the homologous chromosome, which is usually a high-fidelity process, returning the broken chromosome to the wild-type sequence. However, if homologous recombination fails, the broken ends can be repaired by nonhomologous end joining (NHEJ), which is highly mutagenic. Cells lacking either of the BRCA functions are unable to repair broken DNA strands by homologous recombination and they have no alternative but to repair by NHEJ, which introduces an insertion, a deletion, or a chromosome rearrangement at the site of the break.

Somatic Mutations in the DNA Damage Response Pathway Are Oncogenic

Inherited predispositions for a defect in DNA repair account for only a small fraction of tumors, yet the vast majority of cancer cells exhibit a much greater propensity for genome instability and a higher rate of mutagenesis than normal somatic cells. This propensity of cancer cells to acquire new mutations at a high rate is the result of somatic mutations acquired during cancer progression that compromise efficient DNA repair. The majority of these repair-defective somatic mutations do not directly affect the repair pathways themselves but alter the ways that cells respond to DNA damage.

We saw in Chapter 19 that a general response to DNA damage of all dividing eukaryotic cells is to arrest the cell cycle to give the repair systems time to repair the damage before entry into mitosis. In brief outline, serine kinases ATM or ATR are recruited to sites of DNA damage and become activated (Figure 19-34). The activated kinases then signal the presence of DNA damage by phosphorylating p53 on a serine residue in the N-terminus of the protein. This phosphorylation causes the protein to evade ubiquitin-mediated degradation, leading to a marked stabilization of p53 that activates transcription of the gene encoding p21, which in turn binds to and inhibits mammalian CDK2, CDK1, and CDK4/6 complexes (Figure 19-34). The net outcome of this signaling process is for cells with damaged DNA to activate cell cycle checkpoints causing an arrest in the cell cycle, allowing time for DNA repair to be completed before chromosome segregation in mitosis. Once DNA repair has been completed, p53 is degraded and the cell cycle will recommence to complete mitosis. If for some reason the DNA damage cannot be repaired, the cell will be arrested permanently and become senescent.

This pathway that enables DNA-repair enzymes to operate effectively appears to be the aspect of DNA repair that is most vulnerable to somatic mutation. Most, if not all, human tumors have mutations either in p53 itself or in other proteins that regulate p53 activity. When the p53, G_1, and S phase checkpoint control does not operate properly, damaged DNA can replicate, generating mutations

and DNA rearrangements that are passed on to daughter cells and make their transformation into metastatic cells more likely. For example, loss of p53 function leads to a hundredfold or greater increase in the frequency of gene amplification. At the same time, loss of p53 function limits the duration of G_2 arrest, allowing cells with damaged DNA to enter mitosis prematurely. The activity of p53 is not limited to inducing cell cycle arrest, and in Section 25.4 we will describe the role of p53 in inducing apoptosis.

The fact that cancer cells usually contain one or more defective DNA-repair systems provides an Achilles heel that allows tumors to be killed by chemicals that intentionally damage DNA. For example, traditional anticancer chemotherapeutic agents like cis-platinum cause DNA cross-links that tumor cells cannot efficiently repair, while drugs like doxorubicin and etoposide, which inhibit DNA topoisomerases, cause large numbers of DNA breaks during S phase, blocking entry into mitosis and resulting in tumor cell death. This loss of normal DNA repair, combined with the rapid proliferation of tumor cells, explains why commonly used chemotherapy drugs preferentially kill cancer cells, but do not kill the normal nonproliferating cells in the body. An exception to this are the rapidly proliferating normal cells in the bone marrow and immune system, which are also damaged to some extent by these drugs. This explains why cancer patients develop anemia and are susceptible to infections during chemotherapy treatment. ■

Cancer Genome Sequencing Reveals an Enormous Diversity of Somatic Mutations

The advent of next-generation DNA-sequencing technology has enabled rapid sequencing of thousands of cancer genomes and compilation of the vast array of somatic mutations in cancer cells. The large collection of cancer genome sequences shown in Figure 25-8 shows the complexity and variation in the types and numbers of different somatic mutations found in tumors. Normal somatic cells acquire an average of about one mutation per cell division. Given the total accumulation of $>10^5$ mutations in cancer cells, the mutation rate during cancer progression is at least 1000 times greater than normal. Despite the enormous diversity of somatic mutations in cancer, a number of useful generalizations can be made.

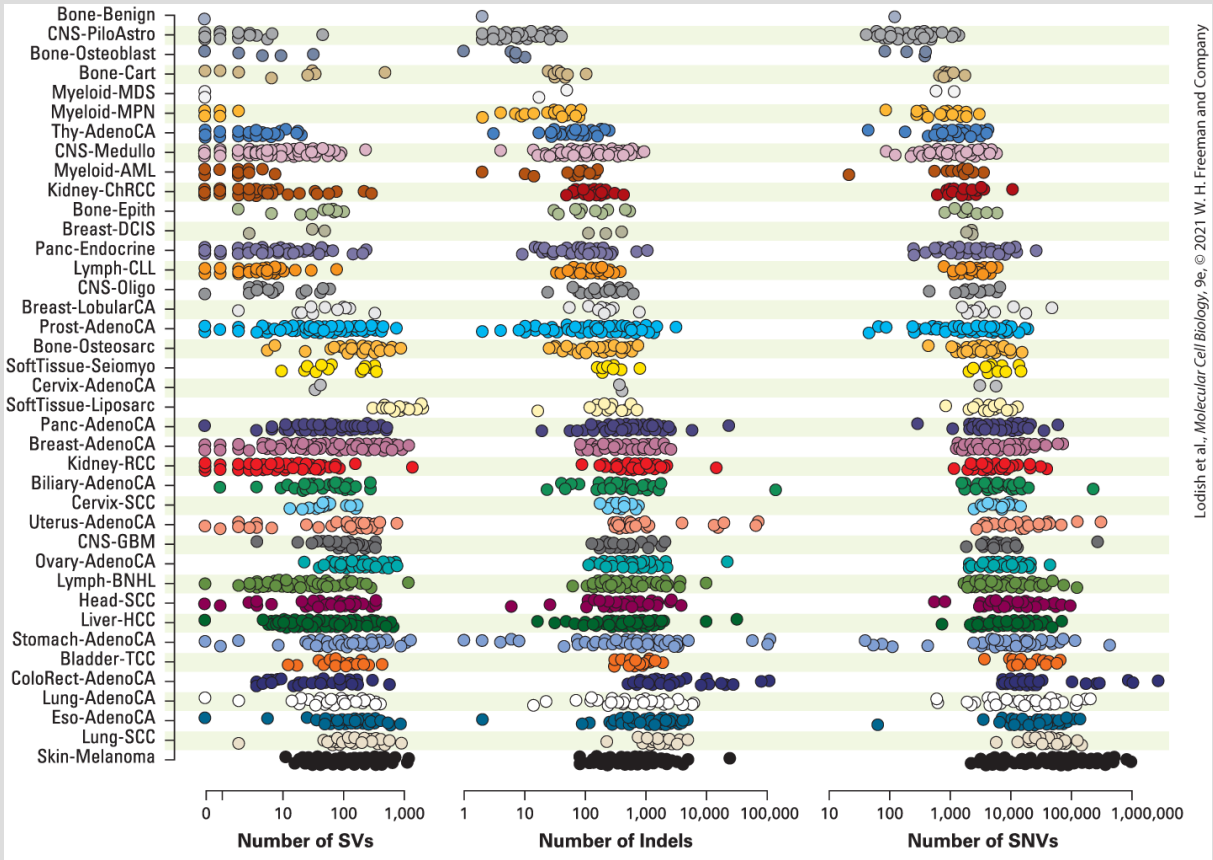


FIGURE 25-8 Sequenced cancer genomes carry a large number and wide variety of somatic mutations. A compendium of 2658 cancer genome sequences compared to matching normal somatic cells. The sequences are grouped into 38 tumor types and each dot specifies the number of mutations in a single tumor categorized as single nucleotide variants (SNVs), short insertions or deletions (Indels), or structural variants (SVs).

[P. J. Campbell et al., 2020, "Pan-Cancer Analysis of Whole Genomes," *Nature* **578**:82–93, https://doi.org/10.1038/s41586-020-1969-6, (Creative Commons Attribution 4.0 International License).]

1. Somatic mutations can be grouped into three general types: *s*ingle *n*ucleotide *v*ariants (SNVs, or point mutations); *in*sertions or *del*etions (Indels); or *s*tructural *v*ariants (SVs), which include chromosomal rearrangements, duplications, and copy number variations.

2. Although a typical tumor might contain 5000 SNVs, 500 Indels, and 50 SVs, there is enormous variation in number from one tumor to another even for the same type of cancer.
3. The cancers that have the greatest number and variation in mutations arise from tissues that have been exposed to mutagens: melanoma from sunlight and lung tumors from heavy smokers.
4. The types of mutations can reflect the underlying mechanism for increased mutagenesis. For example, breast tumors from women carrying a BRCA1 mutation tend to have a greater proportion of Indels and SVs than sporadic breast tumors. This is consistent with the expectation that a BRCA1 defect would lead to error-prone repair of DNA breaks, leading to Indels and SVs.

Oncogenes Were Discovered by Their Association with Tumor Viruses

Now that we have seen some of the ways that cancer cells acquire a large number of random somatic mutations, we turn to the problem of identifying the small fraction of these mutations that have functional consequences and have contributed to cancer progression. With the modern capability of analyzing entire genomes, we now know that a typical tumor may have approximately five known *driver mutations* that contribute to the uncontrolled cell proliferation, avoidance of programmed cell

death, and unstable genome that are hallmarks of cancer cells. Finding a few driver mutations among tens of thousands of somatic mutations has only been possible recently; the first known genetic drivers of oncogenesis were revealed in cases in which a single powerful driver was activated through an unusual genetic event. The genes that caused the oncogenic **transformation** are known as **oncogenes**.

Pioneering studies by Peyton Rous beginning in 1911 led to the initial recognition that a virus could cause cancer when injected into a suitable host animal. Many years later, molecular biologists showed that his Rous sarcoma virus (RSV) is a **retrovirus** whose RNA genome is reverse-transcribed into DNA, which is then incorporated into the host-cell genome (see Figure 5-44). In addition to the normal genes present in all retroviruses, oncogenic transforming viruses such as RSV contain an oncogene — in the case of RSV, the v-*src* gene. Subsequent studies with mutant forms of RSV demonstrated that only the v-*src* gene, not the other viral genes, was required for cancer induction.

In the late 1970s, scientists were surprised to find that normal cells from chickens and other species contain a gene that is closely related to the RSV v-*src* gene. This normal cellular gene, a **proto-oncogene**, is commonly distinguished from the viral gene by the prefix *c* for "cellular" (c-*SRC*). The product of this gene, c-Src, is a cytosolic protein tyrosine kinase which participates in many signal

transduction pathways. RSV and other oncogenic transforming viruses are thought to have arisen by incorporating a normal host cellular proto-oncogene into their genome. Subsequent mutations in the incorporated gene then converted it into a dominantly acting oncogene, encoding a constitutively active kinase that is able to transform host cells even in the presence of the normal c-*SRC* proto-oncogene. When this phenomenon was first discovered, it was startling to find that these dangerous viruses were turning the hosts' own genes against them.

In other cases, retroviruses can cause cancer not by carrying an oncogene but by integrating into the host-cell DNA near a cellular proto-oncogene and activating its expression. For example, in the cells from tumors caused by avian leukosis virus (ALV), the retroviral DNA is inserted near the *MYC* gene. These cells overproduce MYC protein, which causes abnormally rapid proliferation of cells, initiating the cancer progression. Additional subsequent mutations have to occur before a full-fledged tumor becomes evident.

Single Oncogenic Drivers Can Be Activated by Chromosome Rearrangements

In the 1960s, researchers first realized that some cancers harbor characteristic chromosome rearrangements that could be detected by light microscopy. Chronic myelogenous leukemia (CML), a common leukemia in humans, was found to be associated with the **Philadelphia chromosome** (Figure 25-9a), which is generated by a translocation between chromosomes 22 and 9. The two chromosomes exchange their terminal regions, which leads to a characteristic alteration in the size of chromosome 22. Subsequent cloning and analysis of the DNA at the breakpoint of this translocation showed that a new fusion protein, called BCR-ABL, is generated. This hybrid protein kinase phosphorylates proteins that the wild-type ABL kinase normally does not phosphorylate, thereby inappropriately activating many intracellular signal-transducing proteins. If this translocation occurs in a hematopoietic cell in the bone marrow, the activity of the chimeric *BCR-ABL* oncogene results in the initial phase of CML, characterized by an expansion in the number of white blood cells.

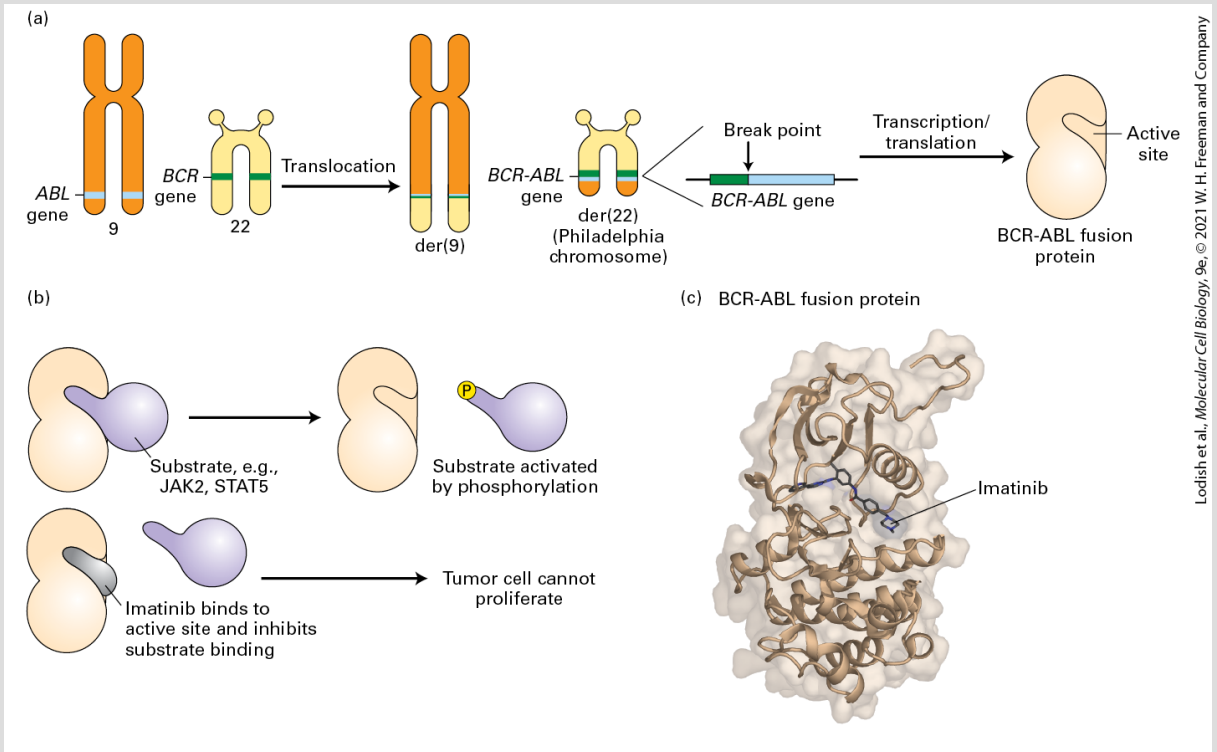


FIGURE 25-9 BCR-ABL protein kinase. (a) Origin of the Philadelphia chromosome from a translocation of the tips of chromosomes 9 and 22 and the oncogenic fusion protein formed by that translocation. (b) The BCR-ABL fusion protein is a constitutively active kinase that phosphorylates multiple signal-transducing proteins. Imatinib binds near the active site of BCR-ABL and stabilizes the inactive form that lacks kinase activity. (c) Imatinib bound to the BCR-ABL active site.

[Data from B. Nagar et al., 2002, *Cancer Res.* **62**:4236, PDB ID 1iep.]

The discovery of the Philadelphia chromosome and the critical oncogene it creates, BCR-ABL, combined with the discovery of the molecular action of the ABL protein, together have led to a powerful new therapy for CML. After a painstaking screening of compounds that can inhibit the enzymatic activity of ABL kinase, a highly specific inhibitor named *imatinib* (Gleevec) was identified as a possible treatment for CML. Imatinib, which binds directly to the

active site of the ABL kinase and holds it in an inactive conformation to inhibit its kinase activity, is lethal to CML cells while sparing normal cells (see Figure 25-9b, c). After clinical trials showing that imatinib is remarkably effective in treating CML despite some side effects, it was approved by the FDA in 2001 as the first cancer drug targeted to a signal-transducing protein unique to tumor cells. Imatinib inhibits several other tyrosine kinases that are implicated in different cancers and has been successful in trials for treating those diseases, including forms of gastrointestinal tumors, as well. There are 90 functional protein tyrosine kinases encoded in the human genome, and drugs that inhibit several of these are proving useful in treating other types of cancers.

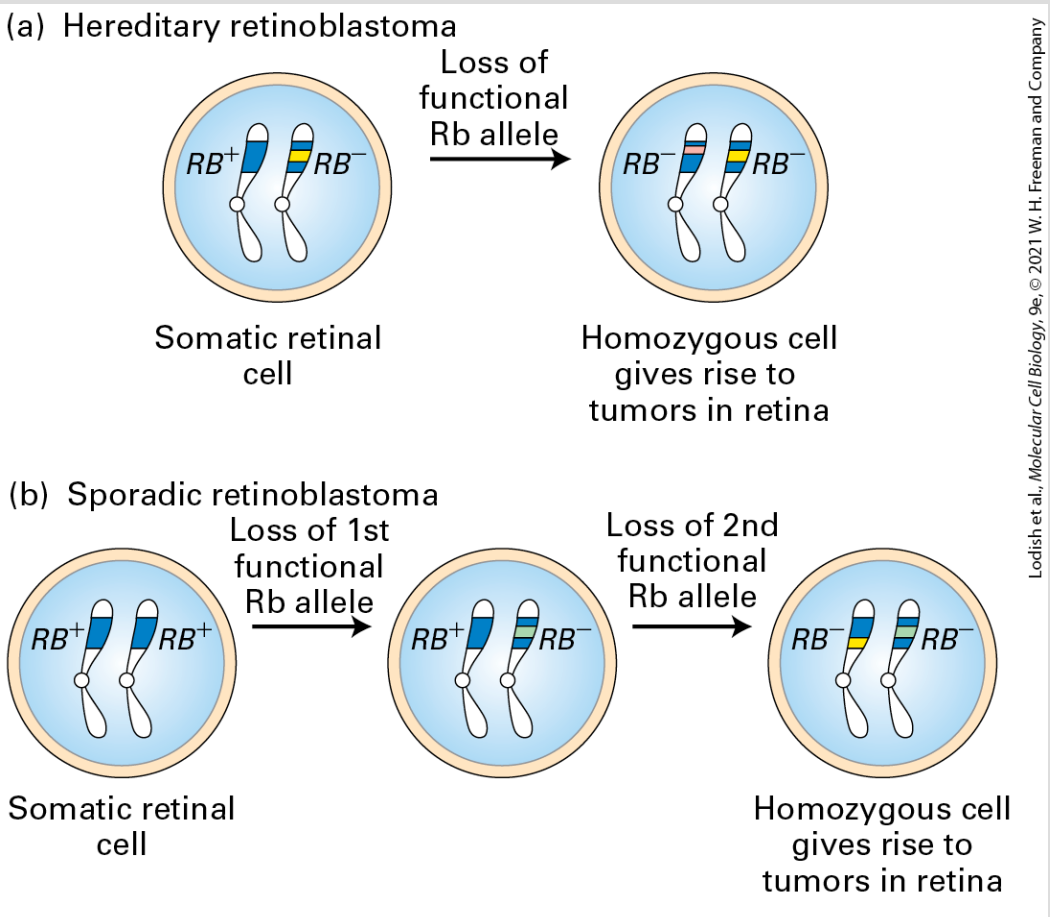
The development of imatinib represents one of the great successes of an approach to cancer therapy known as *targeted therapy*. In its general outline, the idea behind targeted therapy is to find the oncogenic drivers for a particular type of cancer and then to design a drug that reverses the effect of the oncogenic mutation. The case of targeting BCR-ABL that is the oncogenic driver of CML demonstrates both the power of this method and also reveals its limitations. CML is unusual in that it forms as the result of a single oncogenic driver whose effect can be reversed to cure the disease. As we will see, the majority of cancers result from multiple driver mutations; reversing the effect of any one of them would not cure the disease. In addition, we have seen that most cancers have a high mutation rate and thus would rapidly acquire mutations that confer

resistance to a specific enzyme inhibitor. Indeed, CML tumors eventually acquire mutations in the gene encoding the BCR-ABL fusion protein that prevent binding of imatinib. Researchers have identified molecules that specifically inhibit some of these mutant BCR-ABL kinases; these are used as a second-line therapy for CML. The twin challenges of multiple pathways affected and generation of drug resistance through mutation are features of the underlying genetic complexity of most cancers and remain the most significant obstacles to development of targeted therapies.

Inherited Predisposition for Cancer Enabled Identification of Some Oncogenic Drivers

Another way that individual genes responsible for driving progression were identified was through genetic analysis of syndromes that have a hereditary predisposition to certain cancers. For most of these syndromes, individuals inherit from a parent a loss of function mutation in one allele of the gene; then a somatic mutation of the second allele of the gene initiates cancer progression. A classic inherited cancer syndrome is retinoblastoma, which is caused by loss of function of *RB*. As discussed in Chapter 19, the protein encoded by *RB* regulates cell cycle entry.

Children with hereditary retinoblastoma inherit one defective copy of the *RB* gene, sometimes seen as a small deletion on one of the two copies of chromosome 13. These children develop multiple retinal tumors early in life and generally in both eyes. The loss or inactivation of the normal *RB* gene on the other chromosome is an essential step in tumor formation, giving rise to a cell that produces no functional **Rb protein** (Figure 25-10a). Individuals with sporadic retinoblastoma, in contrast, inherit two normal *RB* alleles, each of which has undergone a loss-of-function somatic mutation in a single retinal cell (Figure 25-10b). Because losing both copies of the *RB* gene in a somatic cell is far less likely than losing one, sporadic retinoblastoma in individuals with two normal copies of the *RB* gene is very rare and usually affects only one eye.



Lodish et al., *Molecular Cell Biology*, 9e, © 2021 W. H. Freeman and Company

FIGURE 25-10 Role of spontaneous somatic mutation in retinoblastoma. This disease is marked by retinal tumors that arise from cells carrying two mutant RB^- alleles. (a) In hereditary (familial) retinoblastoma, a child inherits a normal RB^+ allele from one parent and a mutant RB^- allele from the other parent. When the second normal allele is lost in a heterozygous somatic retinal cell, a cell is generated that lacks any *RB* gene function. (b) In sporadic retinoblastoma, a child inherits two normal RB^+ alleles. Two separate *RB* loss events must occur in a particular retinal cell to produce a cell lacking all *RB* function.

If retinal tumors are removed before they become malignant, children with hereditary retinoblastoma often survive until adulthood and produce children, but they are at an increased risk of

developing other types of tumors later in life. Because their germ cells contain one normal and one mutant *RB* allele, these individuals will, on average, pass on the mutant allele to half their children and the normal allele to the other half. Children who inherit the normal allele are normal if their other parent has two normal *RB* alleles. However, those who inherit the mutant allele have the same enhanced predisposition to develop retinal tumors as their affected parent, even though they inherit a normal *RB* allele from their other, normal parent. Thus the tendency to develop retinoblastoma is inherited as a dominant trait: one mutant copy is sufficient to predispose a person to develop the cancer.

As we will see, many human tumors (not just retinal tumors) contain mutant *RB* alleles or mutations affecting other components of the *RB* pathway; most of these tumors arise as the result of somatic mutations. Although hereditary retinoblastoma cases number about 100 per year in the United States, about 100,000 other cancer cases each year involve *RB* mutations acquired in somatic cells.

Estimates vary, but hereditary cancers (cancers that arise due in part to an inherited version of a gene) are thought to constitute about 10 percent of human cancers. Further work tracing the contributions of human genes seems likely to increase the percentage. It is important to remember, however, that the inherited germ-line mutation alone is not sufficient to cause tumor development. Not only must the inherited normal allele be lost or inactivated, but also

mutations affecting other genes must occur for cancer to develop. Thus a person with a recessive tumor suppressor gene mutation can be exceptionally susceptible to environmental mutagens, such as radiation.

Mutation in only one copy of a tumor suppressor gene itself typically does not cause cancer because the remaining normal allele prevents aberrant growth. However, the subsequent loss or inactivation of the remaining normal allele in a somatic cell, referred to as **loss of heterozygosity (LOH)** causes cancer to develop. Three mechanisms exist that can cause the loss of the normal allele. First, the normal allele can become inactive due to a de novo inactivating mutation or deletion. Second, chromosome missegregation can cause loss of the chromosome carrying the normal allele. Neither mechanism is particularly common. By far the most frequent mechanism for LOH is mitotic recombination between a chromatid bearing the normal allele and a homologous chromatid bearing a mutant allele. Subsequent chromosome segregation can generate a daughter cell that is homozygous for the mutant tumor-suppressor allele.

Oncogenic Driver Mutations Have Been Identified in Many Genes

After discovery of the first oncogenes carried by tumor viruses, in the ensuing 40 years dozens of different genes that can carry oncogenic driver mutations have been identified. These genes were typically identified by studying the unique genetic markers of a particular tumor sample or transformed cell line. The greatest advances have come from studying tumor types, like pediatric cancers and rare types of adult cancer, where the total number of genetic changes is small. Most oncogenic drivers were first identified by one of the following six approaches (we have already seen examples of the first five):

1. Identification of an oncogene associated with a tumor-causing virus.
2. Identification of a gene activated to cause cancer by insertion of a retrovirus.
3. Identification of a gene at the breakpoint in a chromosome rearrangement associated with cancer.
4. Mapping of a gene that causes a familial predisposition to cancer.
5. Identification of a gene that, when activated or up-regulated, can cause cell transformation.
6. Identification of a gene specifically up-regulated in cancer cells.

Oncogenic Driver Mutations Can Be Identified by Comparing Cancer Genomes

All of the oncogenic driver mutations in a given tumor should be discoverable by comparing the genomic sequence of the tumor with that of the sequence of a normal tissue-matched somatic cell. The difficulty with this approach is that for most tumors, out of tens of thousands of somatic mutations only a few will be oncogenic driver mutations. However, when a large number of different cancer genome sequences are compared, patterns emerge that allow driver mutations to be identified among the much larger number of adventitious somatic mutations in tumor cells. The most obvious pattern would be to look for mutations in the same gene appearing across multiple tumors of the same type. This method is based on the concept that cancer cells of the same type have similar morphology, metabolism, and regulatory and gene expression networks and therefore are likely to be the result of similar sets of oncogenic driver mutations. More sophisticated analysis of the sequence information will identify common mutations of the same type across tumors. For example, mutations in *RB* tend to be deletions that cause a loss of gene function, whereas mutations in *Ras* are usually point mutations that have a gain of function.

This type of analysis, applied to 2500 different tumor sequences from different types of cancer, yields a set of common oncogenic drivers shown in Figure 25-11. The average number of oncogenic driver mutations found in these tumors was five and the most common driver gene mutated is p53 (*TP53*), which was found in 77 percent of the tumors. Although at first inspection the number and

variety of oncogenic drivers can be bewildering, almost all of these genes had previously been associated with cancer progression and their basic cellular functions are known. Moreover, many of the oncogenic driver genes operate in a relatively small number of regulatory pathways and cellular processes, as shown in Table 25-2. In Section 25.3, we describe in greater detail the function of oncogenic drivers that participate in cell growth control pathways; in Section 25.4, we consider oncogenic drivers that prevent programmed cell death and allow evasion of immune surveillance.

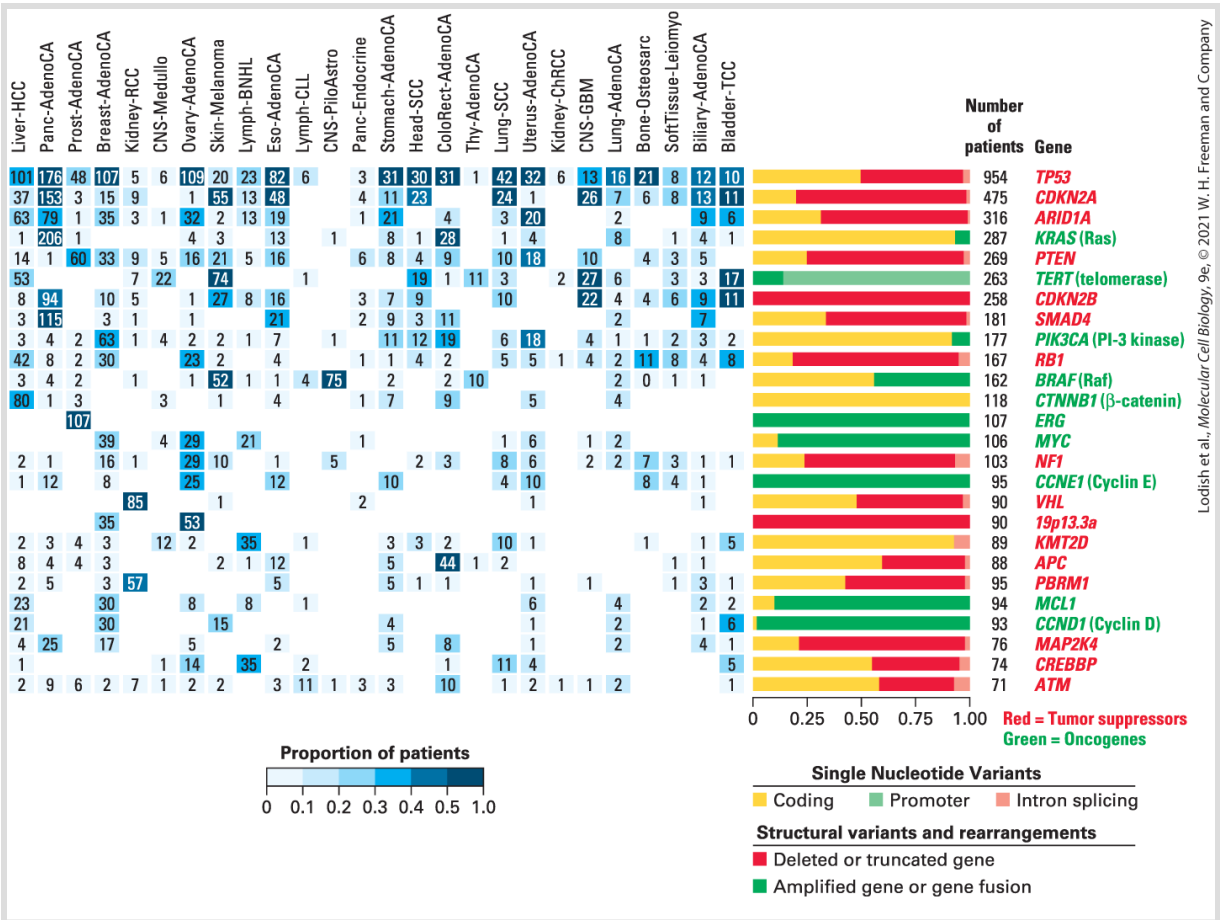


FIGURE 25-11 Oncogenic driver mutations identified in sequenced cancer genomes. A compendium of 2583 cancer genome sequences from 26 different tumor types were searched for shared oncogenic driver mutations. The columns of the table represent

different types of cancer and the rows represent the most prevalent driver mutations in order of frequency. The numbers in the table represent the number of tumors of a given type found with a given driver mutation. The colored bars represent the proportion of driver mutations of a given type. The most common types of mutations are single nucleotide variants in the coding sequence (yellow), which can be either loss-of-function mutations as in the case of TP53 or gain-of-function mutations as in the case of KRAS. Common loss-of-function mutations are structural variants which lead to a deleted or truncated gene (red) and less common intron splicing mutants (light red). Common gain-of-functions mutations are structural variants that lead to an amplified gene or a gene fusion (green) and less commonly promoter mutations (light green). Tumor suppressor genes (labeled red) predominantly carry loss-of function oncogenic driver mutations, whereas oncogenes (labeled green) carry gain-of function oncogenic driver mutations.

[Data from ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020, *Nature* **258**:82.]

Oncogenic Drivers Can Be Gain-of-Function or Loss-of-Function Mutations

In the simplest terms, gene function can be altered by a mutation in one of two ways. A gene can be partially or completely inactivated by a mutation to cause a loss of function, or a gene can possibly acquire from a mutation increased activity or a new activity that could be considered a gain of function. It is usually possible to deduce from a cancer genome sequence, whether an oncogenic mutation causes a loss of function or a gain of function and thus to obtain important information about the mechanism by which the mutation drives cancer progression.

Typical loss-of-function mutations:

1. A *deletion* of all or part of a gene.
2. A *point mutation* in the coding sequence that introduce a stop codon or a frameshift.

Typical gain-of-function mutations:

1. *Duplication or amplification* of the gene leading to overproduction of the encoded protein.
2. A *point mutation* that results in a hyperactive or constitutively active protein product.
3. A *promoter mutation* that results in increased expression of the protein product.
4. A DNA rearrangement that places the gene under the control of an enhancer or promoter that results in increased or inappropriate expression of the protein product.

Genes for which the oncogenic forms arise because of a loss-of-function mutation are usually known as **tumor suppressor genes**. *RB* is the prototypical tumor suppressor gene and we have already seen that retinoblastoma arises because of loss-of-function mutations in both copies of the *RB* gene. We could also deduce that *RB* was a tumor suppressor gene from the cancer genome sequences in Figure 25-11, since most of the oncogenic mutations in *RB* are deletions. Similarly, it is evident that p53 is a tumor

suppressor gene since about half of the oncogenic mutations in the *TP53* gene are deletions.

Genes for which the oncogenic forms arise because of a gain-of-function mutation are usually known as oncogenes. For example, the oncogenic mutation in the *RAS* gene is due to a single base change that permanently activates the Ras protein in the active GTP-bound state and inappropriately activates the MAP kinase pathway, generating an excessive or uncontrolled proliferation-promoting signal. Almost all of the *RAS* mutations identified in tumors are of this type. In contrast, the *MYC* oncogene typically acquires a gain of function by amplification of the *MYC* gene, which increases the abundance of the protein product.

Tumor Suppressor Genes and Oncogenes Often Operate in the Same Pathway

Most of the common oncogenic mutations shown in Figure 25-11 affect pathways that regulate growth control or cell cycle progression. These pathways contain some elements that act positively, but they usually contain negatively acting elements as well (Figure 25-12). For pathways that promote cell growth and cancer progression, gain-of-function oncogenic mutations can occur in pathway elements that act positively, whereas loss-of-

function mutations may occur in pathway elements that act negatively. Thus both oncogenes and tumor suppressor genes can be found that function in the same pathway. However, most tumors usually have only one oncogenic mutation that activates a given pathway. For example, many tumors have oncogenic mutations in *Ras* or *Raf* (see Figure 25-11), but very few have mutations in both genes since these proteins operate in the same signal transduction pathway (see Figure 16-13 and Table 25-2).

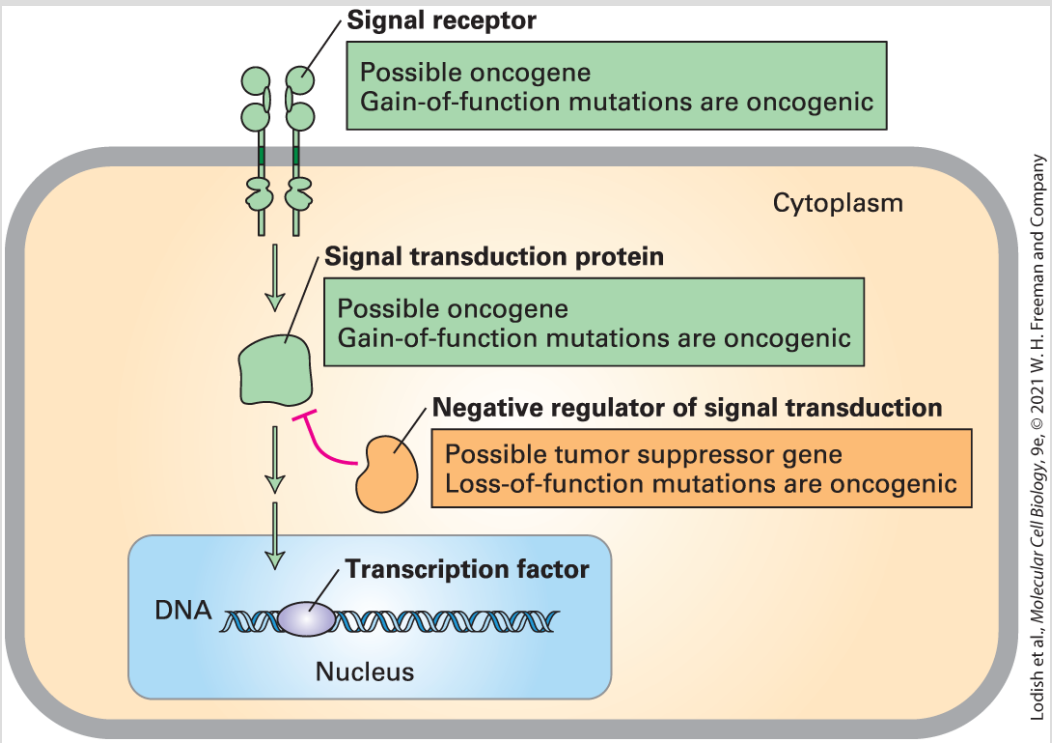


FIGURE 25-12 A typical cell growth pathway showing how gain-of-function or loss-of-function mutations in the same pathway can drive cancer progression. A typical pathway is shown that activates cell proliferation by gene transcription and initiation of cell cycle progression in response to an extracellular mitogenic signal. If proteins that act positively in this pathway, such as the receptor protein or an intracellular signaling protein, can become oncogenes by acquiring a gain-of-function mutation. In contrast, a negative

regulator of the pathway may activate the pathway because of a loss-of-function mutation. In such a case the corresponding gene for the negative regulator would be a tumor suppressor gene.

TABLE 25-2 • Common Types of Mutations That Drive Cancer Progression

Cellular Process	Protein (Human Gene)	Type of Mutation (Gain of Function or Loss of Function)
Activation of a mitogenic growth pathway (e.g., the MAP kinase pathway)	Ras (*KRAS*)	Activating coding (GoF)
	Raf (*BRAF*)	Activating coding and promoter (GoF)
	Nf1 (*NF1*)	Mostly deletion (LoF)
Inactivation of a growth inhibitory pathway (e.g., TGFβ pathway)	Smad4 (*SMAD4*)	Mostly deletion (LoF)
Activation of cell cycle initiation	p16 (*CDKN2A*)	Mostly deletion (LoF)
	Rb (*RB*)	Mostly deletion (LoF)
	Cyclin D (*CCND1*)	Gene amplification (GoF)
Sustaining chromosome replication by elongation of telomeres	Telomerase (*TERT*)	Promoter (GoF)
Loss of DNA damage checkpoint	p53 (*TP53*)	Deletion and coding (LoF)
	ATM (*ATM*)	Deletion and coding (LoF)

Avoidance of apoptosis	p53 (*TP53*)	Deletion and coding (LoF)
	Bcl2 (*MCL1*)	Gene amplification (GoF)
	PI-3 kinase (*PIK3CA*)	Activating coding (GoF)
	PTEN (*PTEN*)	Mostly deletion (LoF)
Change in chromatin state by SWI/SNF	ARID1 (*ARID1A*)	Mostly deletion (LoF)
Activation of nuclear transcription factors	MYC (*MYC*)	Gene amplification (GoF)
	β-catenin (*CTNNB1*)	Activating coding (GoF)

MicroRNAs Can Promote and Inhibit Tumorigenesis

In the last two decades, a new class of oncogenic factors has emerged. Noncoding RNAs (RNAs that do not encode proteins), especially microRNAs (miRNAs), play a critical role in tumorigenesis. As discussed in Chapter 9, generation of miRNAs typically involves the transcription of a precursor RNA that, through a number of processing steps, is trimmed down to a mature miRNA that is 20–22 nucleotides long. The mature miRNA usually base-pairs with the 3′ untranslated region (UTR) of its target RNA and inhibits

its translation or sometimes causes its degradation (Chapter 9). To date, more than 1500 miRNAs have been identified in humans and have been implicated in the regulation of as many as 30 percent of the cell's mRNAs, with fundamental roles in cell proliferation, differentiation, and apoptosis. A number of miRNAs have also been shown to function as tumor suppressor genes or oncogenes.

The first known role for miRNAs in tumorigenesis was revealed by the analysis of chromosomal region 13q14.3. This genomic region is found deleted in most cases of chronic lymphocytic leukemia (CLL), prostate cancer, and pituitary adenomas. Characterization of the disease-causing deletion showed the absence of two distinct miRNAs encoded in the deleted region. Deletions of the homologous miRNAs in a mouse strain caused the mice to develop CLL, showing that these miRNAs act as tumor suppressors. The two miRNAs appear to repress translation of certain cell proliferation proteins; in the absence, proliferation of B cells is increased. Similarly, the *let-7* family of miRNAs has been implicated in lung, colon, breast, and ovarian cancer. Let-7 miRNAs repress the translation of Ras mRNA. Thus in the absence of the miRNAs, Ras is constitutively overproduced, contributing to tumorigenesis. Let-7 miRNAs have other targets as well, such as the oncogenic transcription factor MYC, which we will discuss in detail in the next section. A general theme that emerges in the study of miRNAs in cancer is that each miRNA has multiple targets and therefore ample opportunities to contribute to tumorigenesis.

Other miRNAs have also been found to be overexpressed in cancer. Of particular interest is miR-21, which is overexpressed in most solid tumors, including glioblastomas and breast, lung, pancreatic, and colon tumors. This miRNA targets several tumor suppressor genes, among them the gene encoding the PTEN phosphatase, which normally inhibits the PI-3 kinase signal transduction pathway (Chapter 16). Much more needs to be learned about how miRNAs contribute to tumorigenesis, but it is clear that through their ability to regulate many different genes, they can influence disease progression in more than one way.

Epigenetic Changes Can Contribute to Tumorigenesis

We have just seen how mutations can be oncogenic by causing a loss of function of a tumor suppressor gene or a gain of function in an oncogene. As we saw in Chapter 9, gene expression can either decrease or increase according to the chromatin state. Genes with global effects on chromatin state through changes in DNA methylation, as well as changes in the activity of *histone-modifying enzymes* or *chromatin-remodeling complexes*, are now recognized as major drivers of tumorigenesis.

DNA methylation occurs at cytosines of CpG islands, which are found largely in promoters of genes. Methylation of these Cs leads to

repression of the promoters. A large fraction of colorectal cancers are characterized by DNA hypermethylation. DNA hypomethylation is also a hallmark of cancer. The promoters of many genes involved in cancer are hypomethylated, and expression of the genes under their control is therefore increased. For example, 25 percent of acute myeloid leukemias are characterized by DNA hypomethylation that is due to inactivating mutations in an enzyme that catalyzes the methylation of CpG dinucleotides.

Genes encoding chromatin modifiers and regulators have also emerged as drivers of tumorigenesis. Systematic, whole genome sequencing of many tumor types has revealed highly recurrent alterations in approximately 40 genes encoding epigenetic regulators. Recurrent mutations were found in genes encoding enzymes that modify histones or proteins that bind to these post-translational modifications of histones. Genes encoding histone methyl transferases, histone demethylases, and histone acetyl transferases have all been found mutated in a wide variety of tumors. Interestingly, tumors typically harbor only a single mutated allele of a gene encoding a chromatin-modifying enzyme, indicating that these mutations are haplo-insufficient. Presumably, losing both alleles would kill the cell, but having only one functional allele alters the extent of histone modification and thus the expression of target genes sufficiently to promote tumorigenesis.

Central among the chromatin-remodeling factors implicated in cancer are the SWI/SNF complexes. These large and diverse

multiprotein complexes, which have an ATP-dependent helicase at their core, often control histone modification and chromatin remodeling (see Chapter 8). For example, they can cause changes in the positions or structures of nucleosomes, making genes accessible or inaccessible to DNA-binding proteins that control transcription. If a target gene is normally activated or repressed by SWI/SNF-mediated chromatin changes, mutations in the genes encoding SWI or SNF proteins will cause changes in the expression of that gene. Studies with transgenic mice suggest that SWI/SNF normally plays a role in repressing the expression of *E2F* genes, thereby inhibiting progression through the cell cycle. Thus loss of SWI/SNF function can lead to overgrowth and perhaps cancer.

Recent evidence from humans and mice has strongly implicated the *SNF5* gene in cancer. The SNF5 protein is a core member of the SWI/SNF complex. In humans, inactivating somatic *SNF5* mutations cause rhabdoid tumors, which most commonly form in the kidney, and an inherited (familial) disposition to form brain and other tumors. Subsequent studies have found genes encoding various BAF proteins, which are also subunits of the SWI/SNF complex, to be mutated in 40 percent of renal cancers, 50 percent of ovarian cancers, and a high fraction of liver and bladder cancers. In summary, epigenetic misregulation has emerged as a major contributor to tumorigenesis. In hindsight, this notion is probably not surprising, given that epigenetic regulation offers the

opportunity to change the expression of many factors and regulatory pathways simultaneously.

KEY CONCEPTS OF SECTION 25.2

Genetic and Genomic Basis of Cancer

- Oncogenic mutations can result from DNA-copying errors or through the effect of environmental mutagens such as DNA-alkylating agents, UV light, and ionizing radiation.
- Mutagens that can cause cancer are known as carcinogens. Some carcinogens such as benzo(*a*)pyrene must first be activated by cytochrome P-450 enzymes.
- A major driver of mutagenesis in cancer cells is through the loss of normal DNA-repair processes. Inherited defects in DNA-repair processes, such as mutations in the *BRCA1* gene, are associated with an increased susceptibility for certain cancers.
- The most common somatic mutations that increase mutagenesis in cancer cells cause loss of the DNA damage checkpoint. Most such mutations are in the tumor suppressor gene *p53*.
- Sequencing of cancer genomes typically reveals thousands of mutations, which include single nucleotide variants (SNVs), insertions or deletions (Indels), or structural variants (SVs) that include chromosomal rearrangements, duplications, and copy-number variations.
- The first cancer-causing oncogenes were identified by their association with tumor viruses.
- Cellular oncogenic driver mutations have been identified at the break points of chromosome rearrangements, by mapping genes responsible for familial cancer syndromes, and by assays for transformation of tissue-culture cells.
- The first tumor suppressor gene to be identified, *RB*, is mutated in retinoblastoma and many other tumors. Inheritance of a single mutant allele of *RB* greatly increases the probability that a cancer will develop, because a somatic cell readily lose the functional copy of *RB* by loss of heterozygosity (LOH) by mutation or deletion of the normal allele, chromosome mis-segregation, or mitotic recombination.
- Comparison of large numbers of genome sequences from the same tumor type allows oncogenic driver mutations to be distinguished from the many adventitious mutations present in cancer genomes.
- Many oncogenic driver genes cause cell proliferation by activating growth-promoting pathways or by inhibiting growth-inhibiting and apoptotic pathways.

- A gene that becomes an oncogenic driver as a result of a gain-of-function mutation is known as an *oncogene*.
- A gene that becomes an oncogenic driver as a result of a loss-of-function mutation is known as a *tumor suppressor gene*.
- Micro-RNAs (miRNAs) can promote or inhibit tumorigenesis by affecting the expression of multiple oncoproteins or tumor suppressor proteins.
- Mutations affecting epigenetic regulators such as histone-modifying enzymes or chromatin remodelers are associated with a variety of tumors.

25.3 Dysregulation of Cell Growth and Developmental Pathways Initiates Tumorigenesis

Uncontrolled cellular proliferation is the central hallmark of cancer. We have already seen that many of the known oncogenic drivers of cancer are in genes involved in the pathways that keep growth of normal tissues in check by a balance of growth-promoting and growth-inhibiting signals. In this section, we examine these pathways in greater detail, with an emphasis on the mechanism by which oncogenic mutations exert their effects.

Receptor Mutations Can Cause Proliferation in the Absence of External Growth Factors

Oncogenes encoding cell-surface receptors that transduce growth-promoting signals have been associated with several types of cancer. A major class of these receptors, known as *receptor tyrosine kinases (RTKs)*, have intrinsic protein tyrosine kinase activity in their cytosolic domains, which is quiescent until activated. As detailed in Chapter 16, ligand binding to the external domains of RTKs leads to their dimerization and activation of their kinase activity, initiating

an intracellular signaling pathway that ultimately promotes proliferation.

In some cases, a point mutation changes a wild-type RTK into one that dimerizes and is constitutively active even in the absence of ligand. For instance, a single point mutation converts the normal human EGF receptor 2 (HER2) into the NEU oncoprotein ("NEU" for its first known role, in neuroblastoma) (Figure 25-13, *left*). Similarly, human tumors called *multiple endocrine neoplasia type 2* produce a constitutively active dimeric glia-derived neurotrophic factor (GDNF) receptor that results from a point mutation in the extracellular domain. The GDNF receptor and the HER2 receptor are both protein tyrosine kinases, so by the signal transduction pathways detailed in Chapter 16, the constitutively active forms excessively activate their downstream target proteins, which ultimately promote cell cycle progression, cell survival and proliferation. In other cases, deletion of much of the extracellular ligand-binding domain produces a constitutively active oncogenic receptor. For example, deletion of the extracellular domain of the normal EGF receptor (Figure 25-13, *right*) converts it to the dimeric ErbB oncoprotein (from *er*ythro*b*lastosis virus, in which a viral version of the altered gene was first identified). Mutations leading to overproduction of a normal RTK can also be oncogenic. For instance, many human breast cancers overproduce a normal HER2 receptor because of amplification of its encoding gene. As a result, the cells are stimulated to proliferate in the presence of very low

concentrations of EGF and related hormones, concentrations too low to stimulate proliferation of normal cells (see Chapter 16).

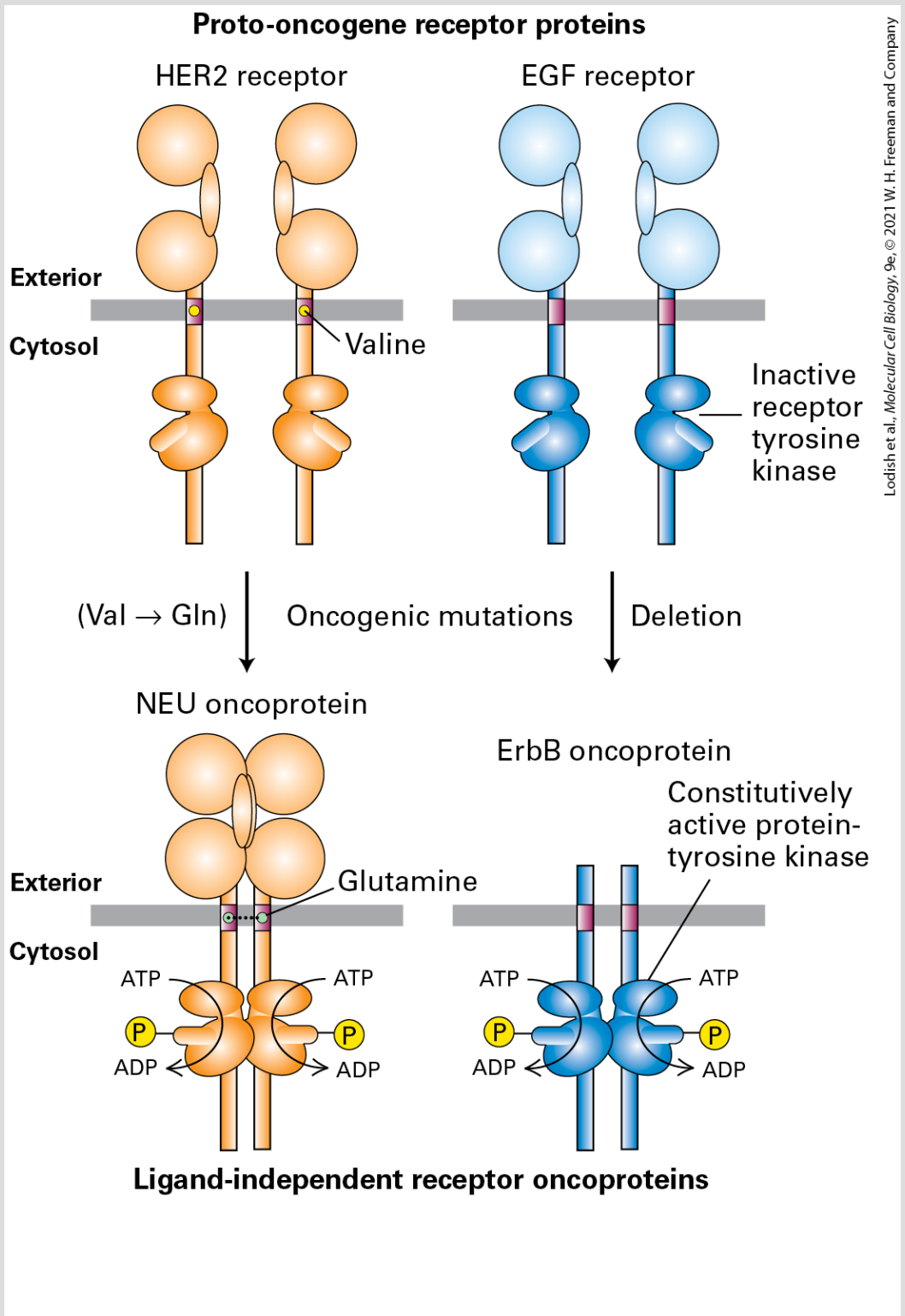


FIGURE 25-13 Effects of oncogenic mutations in proto-oncogenes that encode cell-surface receptors. *Left:* A mutation that alters a single amino acid (valine to glutamine) in the transmembrane region of the HER2 receptor causes dimerization of the receptor, even in the absence of the normal EGF-related ligand, transforming it into the oncoprotein NEU, a constitutively active protein tyrosine kinase. *Right:* A deletion that causes loss of the extracellular ligand-binding domain in the EGF receptor leads, for unknown reasons, to constitutive activation of the kinase activity of the resulting oncoprotein, ErbB.

Many Oncogenic Mutations Constitutively Activate Signal-Transducing Proteins

A large number of oncogenes are derived from proto-oncogenes whose encoded proteins are components or regulators of signal transduction pathways — most prominent among them the Ras pathway. As we saw in Chapter 16, Ras is a key component in the transduction of signals from activated receptors to a cascade of protein kinases. In the first part of this pathway, a signal from an activated RTK is carried via two adapter proteins to Ras, converting it to the active GTP-bound form (see Figure 16-10). In the second part of the pathway, activated Ras transmits the signal via two intermediate protein kinases to MAP kinase. The activated MAP kinase then phosphorylates a number of transcription factors that induce synthesis of important growth and proliferation proteins (see Figures 16-13 and 16-14). Virtually every component of this

RTK/Ras/MAP kinase signaling cascade has been identified as an oncogene or tumor suppressor gene (Figure 25-14).

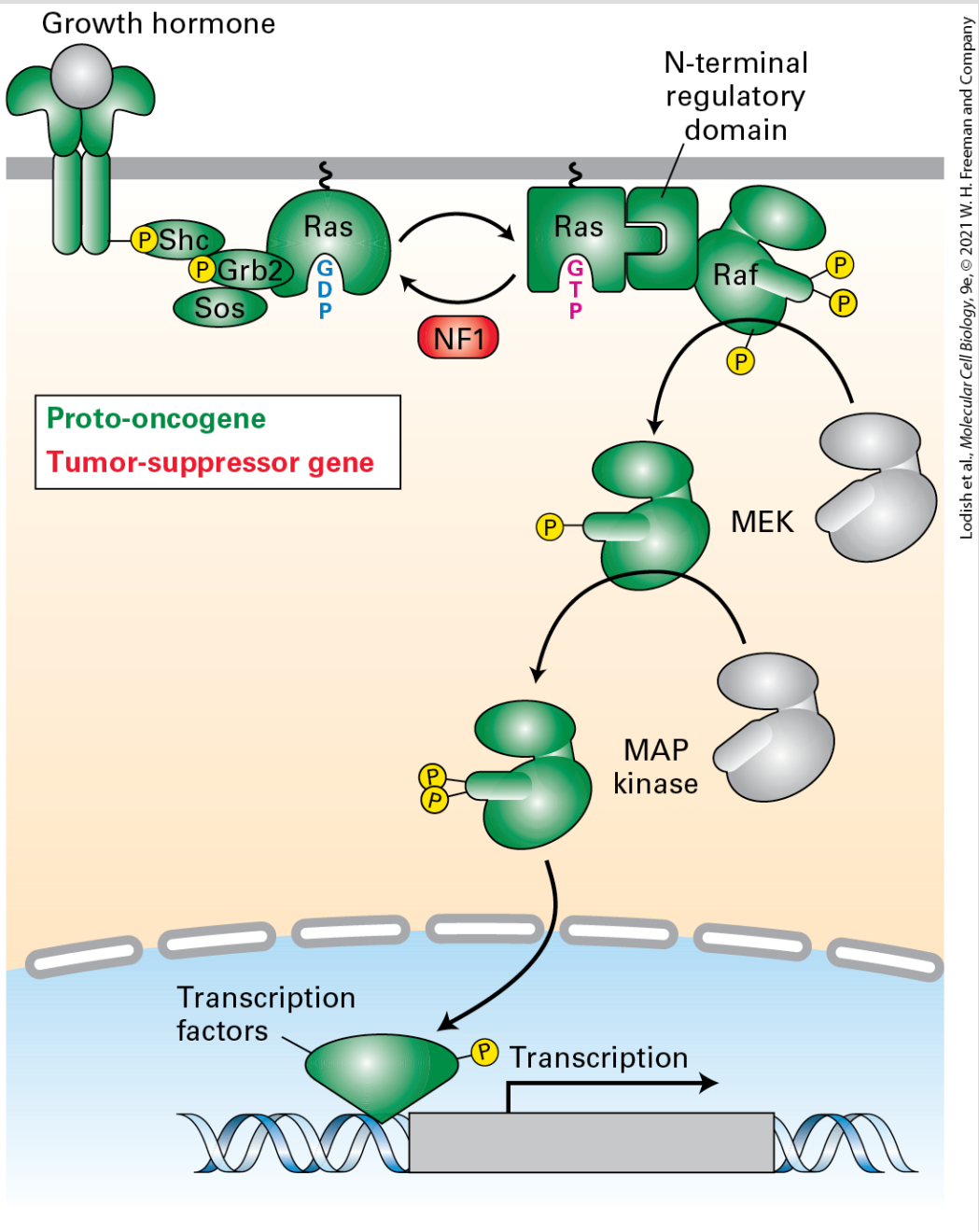


FIGURE 25-14 RTK/Ras/MAP kinase pathway components are frequently mutated in cancer. Oncogenic mutations that activate the RTK/Ras/MAP kinase pathway have been identified in many human cancers. Most components,

highlighted in green, activate the pathway and oncogenic mutations cause a gain of function. In contrast, NF1, highlighted in red, normally inactivates the pathway and the oncogenic mutations cause a loss of function of both copies of the gene.

Among the best studied oncogenes are the RAS^D genes themselves, which were the first nonviral oncogenes to be recognized. Any one of a number of changes in Ras can lead to its uncontrolled and therefore dominant activity. In particular, if a point mutation substitutes any other amino acid for the glycine at position 12 in the Ras sequence, the normal protein is converted into a constitutively active oncoprotein (see Chapter 16). This single mutation reduces the protein's GTPase activity, thus maintaining Ras in the active GTP-bound state. Activating Ras mutations short-circuit the first part of the RTK pathway, making upstream activation triggered by ligand binding to the receptor unnecessary. Constitutively active Ras oncoproteins are produced by many types of human tumors, including bladder, colon, mammary, skin, and lung carcinomas, neuroblastomas, and leukemias.

Constitutive Ras activation can also arise from loss-of-function mutations in a GTPase-activating protein (GAP). The normal function of a GAP is to accelerate hydrolysis of GTP and thus the conversion of active GTP-bound Ras to inactive GDP-bound Ras (see Figure 3-35). Loss-of-function mutations in the Ras-GAP protein, the product of the *NF1* gene, leads to sustained activation of downstream signal-transducing proteins. The relationship between

RAS, an oncogene, and *NF1*, a tumor suppressor gene acting in the same pathway, is a good example of the regulatory circuit shown in Figure 25-12. *NF1* was first discovered as the underlying cause of the familial cancer syndrome neurofibromatosis. Individuals who have inherited a single mutant *NF1* allele then develop neurofibromas, a benign tumor of the sheath cells that surround nerves, caused by loss of both alleles through LOH.

Oncogenes encoding other altered components of the RTK/Ras/MAP kinase pathway have also been identified (see Figure 25-14). For example, constitutively active forms of Raf have been identified in approximately 50 percent of melanomas. As in the case of constitutively active forms of Ras, these mutant Raf forms no longer require regulatory signals coming from the cell surface and signal continuously for cell growth and proliferation.

Growth Control Pathways Ultimately Regulate Initiation of the Cell Cycle

Growth stimulatory pathways such as the RTK/Ras/MAP kinase pathway ultimately have two outputs: transcription of a set of genes required for cell growth and activation of the cell cycle to initiate a new round of cell division. During the cell cycle, once a cell progresses past a certain point in G_1, called the **restriction point**, it becomes irreversibly committed to entering S phase and replicating

its DNA. Cyclin Ds, cyclin-dependent kinases (CDKs), and the Rb protein are all elements of the control system that regulates passage through the restriction point. Many of these proteins that regulate cell cycle initiation are targets for oncogenic mutations.

The pathway that controls entry into the cell cycle is estimated to be misregulated in approximately 80 percent of human cancers. At the heart of this pathway are cyclin D–CDK4/6 complexes and the transcriptional repressor Rb (see Figure 19-16). The expression of cyclin D genes is induced by many extracellular growth factors, or *mitogens*. These cyclins assemble with a partner, CDK4 or CDK6, to generate catalytically active cyclin-CDK complexes whose kinase activity promotes progression through G_1. Mitogen withdrawal prior to passage through the restriction point leads to accumulation of two CDK inhibitors. As described in Chapter 19, these two proteins, p15 and p16, bind to cyclin D–CDK4/6 complexes and inhibit their activity, thereby causing G_1 arrest.

Most tumors contain an oncogenic mutation that causes the overproduction or loss of one of the components of the pathway that controls entry into S phase, so that the cells are propelled into S phase in the absence of the proper extracellular growth signals. For example, elevated levels of cyclin D1, one of the three cyclin Ds, are found in many human cancers. One mechanism that results in overproduction of cyclin D is translocation. In certain tumors of antibody-producing B lymphocytes, the *cyclin D1* gene is

translocated such that its transcription is under the control of an antibody-gene enhancer, causing elevated cyclin D1 production throughout the cell cycle irrespective of extracellular signals. That cyclin D1 can function as an oncoprotein was shown by studies with transgenic mice in which the *cyclin D1* gene was placed under the control of an enhancer specific for mammary duct cells. Initially, the duct cells underwent hyperproliferation, and eventually breast tumors developed in these transgenic mice. A second mechanism that can lead to overproduction of cyclin D is gene amplification. Amplification of the *cyclin D1* gene and concomitant overproduction of the cyclin D1 protein is common in human breast cancers; the extra cyclin D1 helps to drive cells through the cell cycle.

We have already seen that inactivating mutations in both *RB* alleles lead to childhood retinoblastoma, a relatively rare type of cancer. However, loss-of-function mutations in the *RB* gene are also found in the more common cancers that arise later in life (see Figure 25-11). These tissues, unlike retinal tissue, probably produce other proteins (e.g., p107 and p130, both structurally related to *RB*) whose function is redundant with that of *RB*, and thus *RB* is not so critical for preventing cancer in these tissues. In the retina, however, regulation of cell cycle entry appears to rely exclusively on the Rb protein, which is why patients heterozygous for the *RB* gene first develop tumors in this tissue. Rb function can be eliminated not only by inactivating mutations, but also by the binding of an inhibitory protein, designated *E7*, that is encoded by human

papillomavirus (HPV), another nasty viral trick to create virus-producing tissue. At present, this binding is known to occur only in cervical and oropharyngeal cancers.

The proteins that function as cyclin-CDK inhibitors are also targets for oncogenic mutations. In particular, loss-of-function mutations in p16 (*CDKN2A*) that prevent it from inhibiting cyclin D–CDK4/6 kinase activity are among the most common oncogenic drivers in several cancers (see Figure 25-11 and Table 25-2). Loss of p16 mimics overproduction of cyclin Ds and thus p16 normally acts as a tumor suppressor. Although the *p16* tumor suppressor gene is deleted in some human cancers, the *p16* sequence is normal in others. In some of these latter cancers (e.g., lung cancer), the *p16* gene, or genes encoding other functionally related proteins, is inactivated by hypermethylation of its promoter region, which prevents its transcription. What promotes this change in the methylation of *p16* is not known, but it prevents production of this important cell cycle control protein.

The locus encoding p16 is highly unusual in that it encodes no less than three tumor suppressor genes, which makes it the most vulnerable locus in the human genome to oncogenic changes. In addition to harboring the p16-encoding gene, *CDKN2A*, it has the *CDKN2B* locus immediately upstream, which encodes p15, another cyclin D–CDK4/6 inhibitor (Figure 25-15). The locus also encodes a key activator of the tumor suppressor p53. This protein, p14ARF

(p19ARF in the mouse), is encoded by an exon upstream of the first *CDKN2A* exon and shares its exon 2 and exon 3 with *CDKN2A*. As we will see in Section 25.4, this protein controls the stability of p53. Thus mutations in this locus can simultaneously affect the two major tumor-suppressor pathways in the cell, the Rb and p53 pathways.

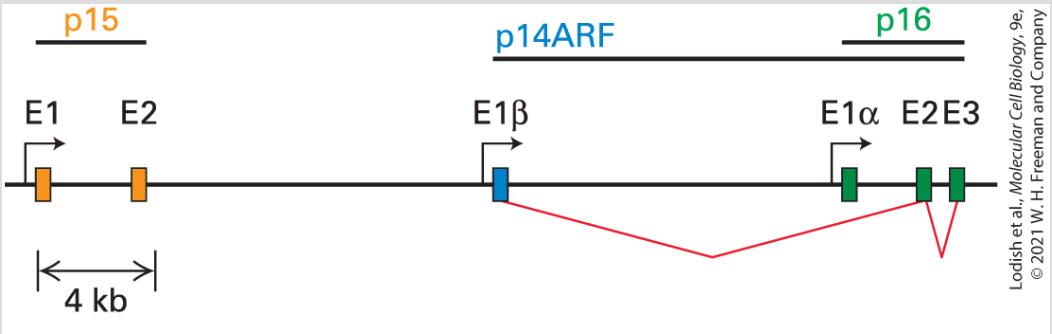


FIGURE 25-15 The p15-ARF-p16 locus encodes three tumor suppressor genes. Exons are designated as *E*. The two *p15* exons (orange) are located upstream of the *ARF* locus. *ARF* (blue) is encoded by a unique E1β exon but shares exons E2 and E3 with *p16* (green). *ARF* encodes a p53 activator.

[Data from C. J. Sherr, 2006, *Nat. Rev. Cancer* **6**:663–673.]

Inappropriate Production of Nuclear Transcription Factors Can Induce Transformation

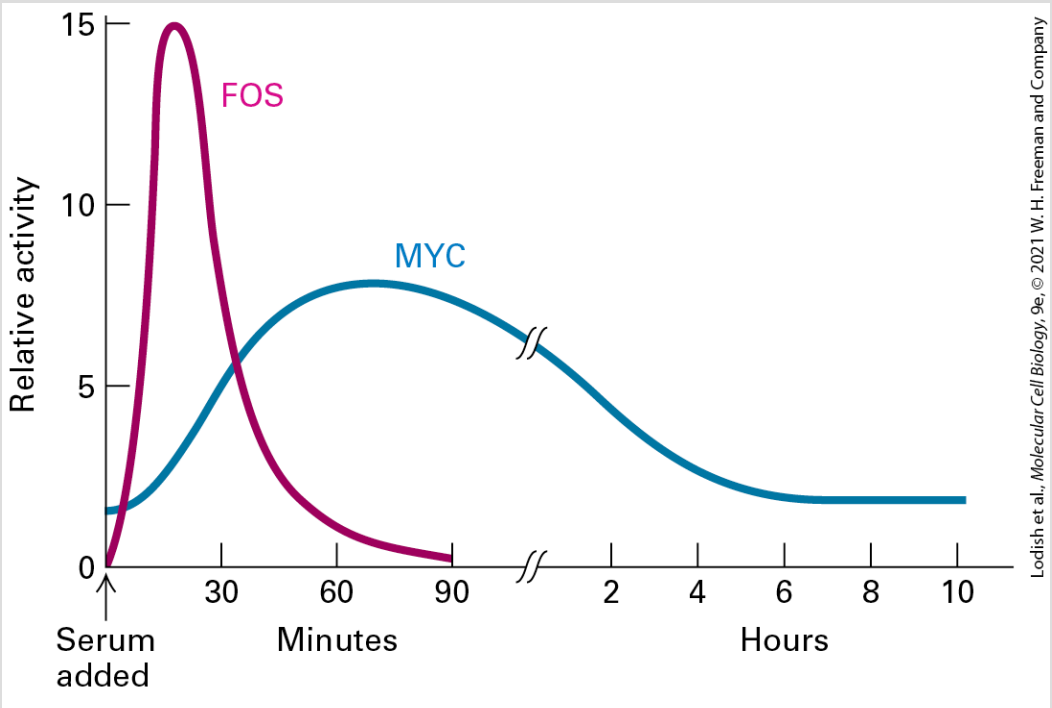
Mutations that create oncogenes or inactivate tumor suppressor genes eventually cause broad changes in gene expression. These changes can be measured by comparing the amounts of different

mRNAs produced in normal cells and in tumor cells. Since the most direct effect on gene expression is exerted by transcription factors, it is not surprising that many oncogenes encode transcription factors. Two such transcription factors with a clear role in tumorigenesis are the FOS and MYC proteins, which stimulate transcription of genes encoding proteins that promote progression through the G_1 phase of the cell cycle and the G_1-to-S transition. We discuss the deregulation of these proteins below.

JUN and FOS were initially identified in transforming retroviruses and later found to be overexpressed in some human tumors. The *JUN* and *FOS* proto-oncogenes encode proteins that sometimes associate to form a heterodimeric transcription factor, called *AP1,* that binds to a sequence found in promoters and enhancers of many genes (see Figure 8-26a and Chapter 16). These proteins function as oncoproteins by activating the transcription of key genes that encode growth-promoting proteins and by inhibiting the transcription of growth-repressing genes.

Many nuclear proto-oncogene proteins are produced when normal cells are stimulated to grow, indicating their direct role in growth control. For example, platelet-derived growth factor (PDGF) treatment of quiescent mouse 3T3 cells induces an approximately 50-fold increase in the production of the transcription factors FOS as well as MYC, the products of the *FOS* and *MYC* proto-oncogenes. Initially, there is a transient rise of FOS and later a more prolonged

rise of MYC (Figure 25-16). The levels of both proteins decline within a few hours, a regulatory effect that may, in normal cells, help to avoid cancer.



EXPERIMENTAL FIGURE 25-16 Addition of serum to quiescent 3T3 cells yields a marked increase in the activity of two proto-oncogene products, FOS and MYC. Serum contains factors such as platelet-derived growth factor (PDGF) that stimulate the growth of quiescent cells. One of the earliest effects of growth factors is to induce expression of *FOS* and *MYC*, whose encoded proteins are transcription factors.

[Data from M. E. Greenberg and E. B. Ziff, 1984, *Nature* **311**:433.]

The oncogenic forms of *FOS* and *MYC* are due to gain-of-function mutations. In normal cells, the mRNAs for these genes and the proteins they encode are intrinsically unstable and degrade rapidly after the genes are transcribed. Some of the genetic changes that

turn *FOS* from a normal gene into an oncogene involve deletions of the sequences that normally make the *FOS* mRNA and protein short-lived. Conversion of the *MYC* proto-oncogene into an oncogene can occur by different mechanisms. In cells of the human tumor known as **Burkitt's lymphoma**, the *MYC* gene is translocated to a site near the heavy-chain antibody genes, which are normally active in antibody-producing white blood cells (Figure 25-17). The *MYC* translocation is a rare aberration of the normal DNA rearrangements that occur during maturation of antibody-producing cells. The translocated *MYC* gene, now regulated by the antibody-gene enhancer, is continually highly expressed, causing the cell to become cancerous. Localized amplification of a segment of DNA containing the *MYC* gene, which occurs in several human tumors, also causes inappropriately high production of the otherwise normal MYC protein. This mechanism of oncogenic activation is similar to the formation of BCR-ABL oncogene by a translocation that produces the Philadelphia chromosome.

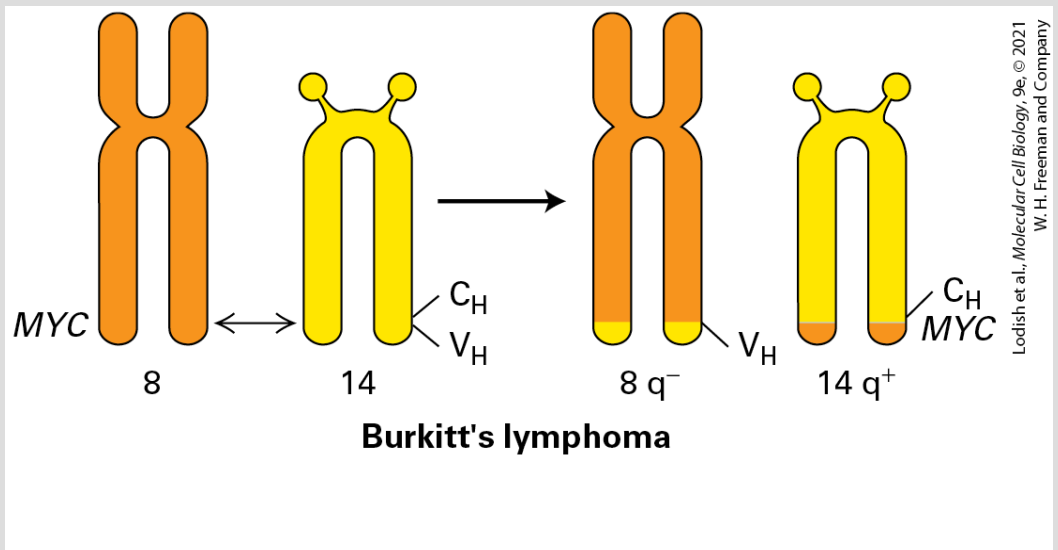


FIGURE 25-17 Chromosomal translocation in Burkitt's lymphoma. As a result of a translocation between chromosomes 8 and 14, the *MYC* gene is placed adjacent to the gene for part of the antibody heavy chain (C_H), leading to overproduction of the MYC transcription factor in lymphocytes and hence their growth into a lymphoma.

The *MYC* gene encodes a basic helix-loop-helix protein that acts as part of a set of interacting proteins that can dimerize in various combinations, bind to DNA, and coordinately regulate the transcription of target genes. Other members of this protein set include MAD, MAX, and MNT. MAX can heterodimerize with MYC, MAD, and MNT. MYC-MAX dimers regulate genes that control proliferation, such as cyclins. MAD proteins inhibit MYC proteins, which has led to an interest in using MAD proteins, or drugs that stimulate MAD proteins, to rein in excessive MYC activity that contributes to tumor formation. MYC protein complexes affect transcription by recruiting chromatin-modifying complexes containing histone acetyl transferases (which usually stimulate transcription; see Chapter 8) to MYC target genes. MAD and MNT work with the SIN3 co-repressor protein to bring in histone deacetylases that help to block transcription. Together, all these proteins form a regulatory network that employs protein-protein association, variations in DNA binding, and transcriptional regulation to control cell proliferation. Overproduction of MYC protein tips the scales in favor of cell growth and division.

Aberrations in Signaling Pathways That Control Development Are Associated with Many Cancers

During normal development, secreted signals such as Hedgehog (Hh), Wnt, and TGF-β are used to direct cells to particular developmental fates, which may include the property of rapid cell cycling. The effects of such signals must be regulated so that growth is limited to the right time and place. Among the mechanisms available for reining in the effects of these powerful developmental signals are inducible intracellular antagonists, receptor blockers, and competing signals. Mutations that prevent such restraining mechanisms from operating are likely to be oncogenic, causing inappropriate or cancerous growth.

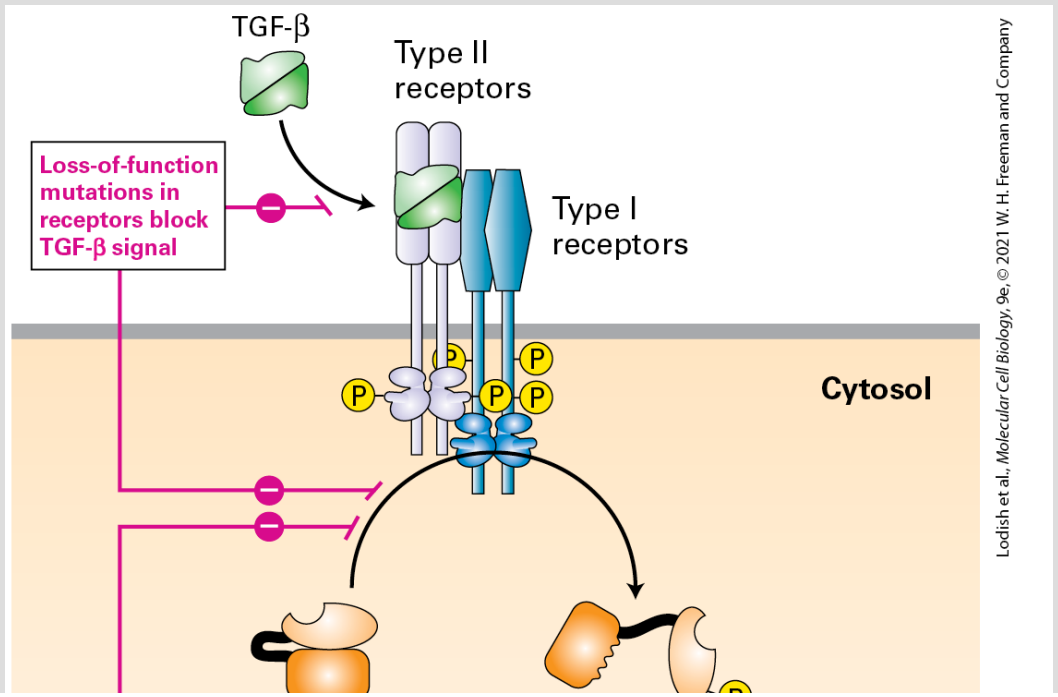
The Hedgehog signaling pathway, which is used repeatedly during development to control differentiation, is a good example of a signaling pathway implicated in cancer induction. In the skin and cerebellum, one of the human Hh proteins, Sonic Hedgehog, stimulates cell division by binding to and inactivating a membrane protein called *Patched1 (PTC1)* (see Figure 16-29). Loss-of-function mutations in *PTC1* permit cell proliferation in the absence of an Hh signal; thus *PTC1* is a tumor suppressor gene. People who inherit a defective copy of *PTC1* have a propensity to develop skin and brain cancer; either can occur when the remaining *PTC1* allele is lost through the mechanism of LOH that we saw for RB and NF1 cancer

syndromes. Spontaneous mutations in both copies of this gene have also been observed in sporadic cases of these cancers. Mutations in other genes in the Hh signaling pathway are also associated with cancer. Some such mutations create oncogenes that turn on Hh target genes inappropriately; others are recessive mutations that affect negative regulators such as PTC1.

Many of the signaling pathways described in Chapters 16 and 20 also play roles in controlling embryonic development and cell proliferation in adult tissues. In recent years, mutations affecting components of most of these signaling pathways have been linked to cancer. Indeed, once one gene in a developmental pathway has been linked to a type of human cancer, knowledge of that pathway gleaned from model organisms such as worms, flies, or mice allows focused investigations of the possible involvement of additional pathway genes in other cases of the cancer. For example, *APC*, a gene that is mutated early on in colon cancer, is now known to be part of the Wnt signaling pathway (see Chapter 16). That knowledge, in turn, led to the discovery of β-catenin mutations in colon cancer.

Mutations in tumor-suppressor developmental genes promote tumor formation in tissues where the affected gene normally acts to restrain growth. For example, transforming growth factor β (TGF-β), despite its name, primarily acts to inhibit proliferation of many cell types, including most epithelial and immune-system cells. Binding of TGF-β to its receptor activates cytosolic Smad transcription

factors (see Figure 16-24). After translocating to the nucleus, Smads can promote expression of the gene encoding p15, an inhibitor of cyclin-dependent kinase 4 (CDK4), which causes cells to arrest in G_1. TGF-β signaling also promotes expression of genes encoding extracellular matrix proteins and plasminogen activator inhibitor 1 (PAI-1), which reduces the plasmin-catalyzed degradation of the matrix. Loss-of-function mutations in either TGF-β receptors, as noted above, or in Smads thus promote cell proliferation and probably contribute to the invasiveness and metastasis of tumor cells (Figure 25-18). Such mutations have in fact been found in a variety of human cancers. For example, deletion of the *Smad4* gene occurs in many human pancreatic cancers; retinoblastoma and colon cancer cells lack functional TGF-β receptors and therefore are unresponsive to TGF-β growth inhibition. Originally Smad4 was called *DPC (deleted in pancreatic carcinoma)*.



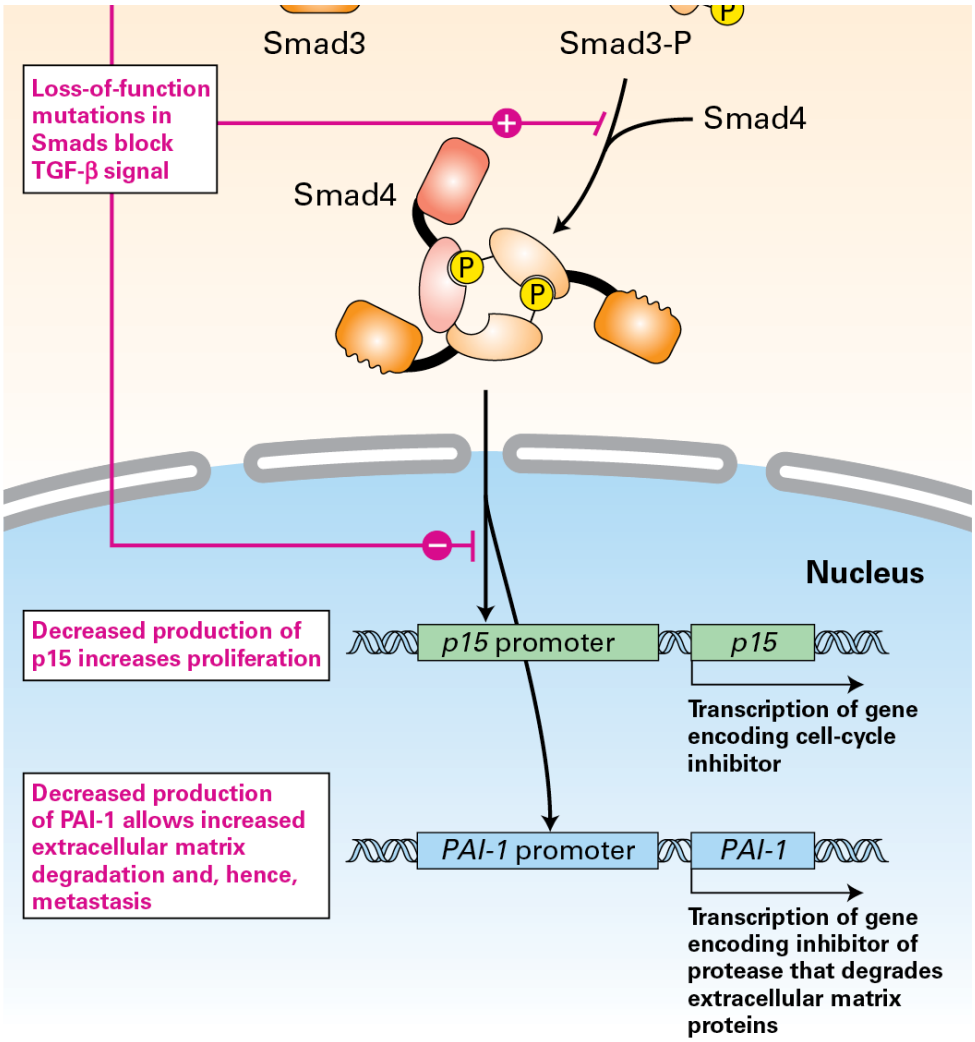


FIGURE 25-18 Effect of loss of TGF-β signaling. Binding of TGF-β, an anti-growth factor, to its receptor causes activation of Smad transcription factors. In the absence of effective TGF-β signaling owing to either a receptor mutation or a Smad mutation, cell proliferation and invasion of the surrounding extracellular matrix increase. See X. Hua et al., 1998, *Genes Dev.* **12**:3084.

Experimental Reconstruction of the Multi-Hit Model for Cancer

As we have seen from the sequencing of cancer genomes, a typical tumor may have acquired thousands of mutations, five of which have been selected for because they are drivers of oncogenesis; many of which are genes we just discussed that regulate cell growth and division. As the data in Figure 25-19 show, the incidence of many types of human cancer increases drastically with age, indicating that it can take decades for the required multiple mutations to occur. This model for the slow progression of cancer raises fundamental questions about the sequence of events needed to lead to a tumor. In what sequential order do the driver mutations occur? Do different driver mutations act synergistically? Answers to these questions are necessary to complete the picture of how cancer forms and will be important for strategies for early diagnosis and cancer prevention.

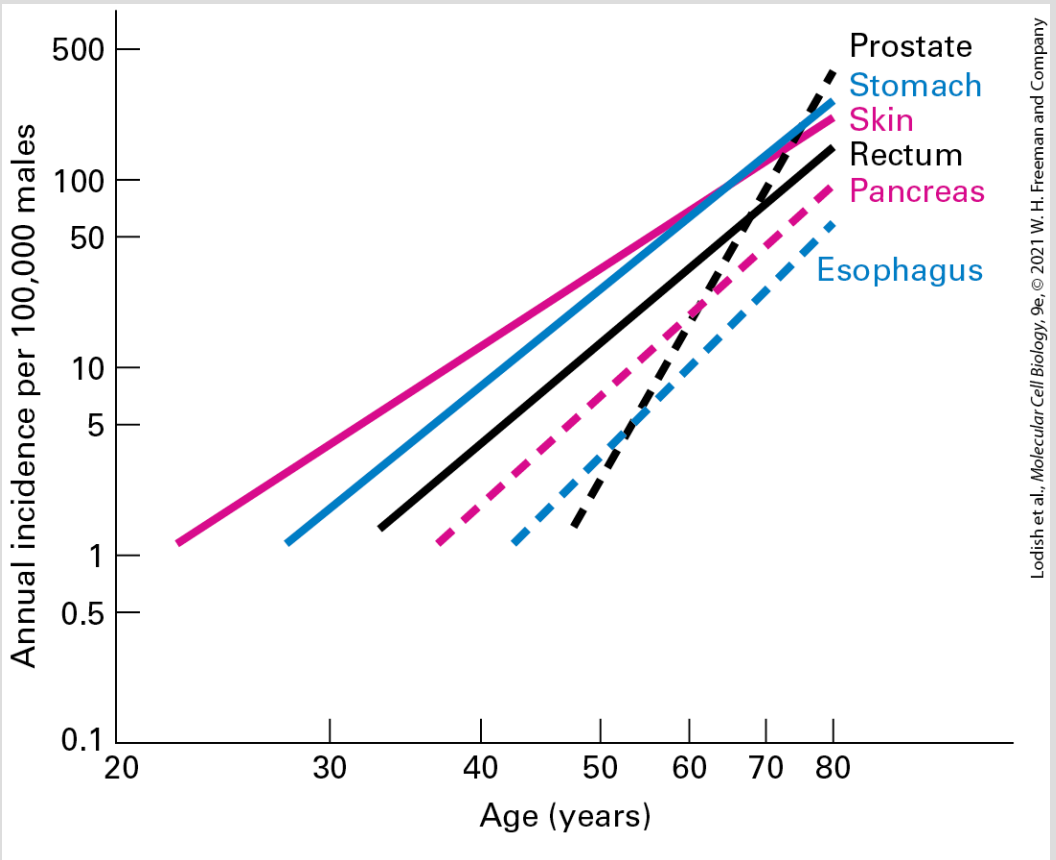


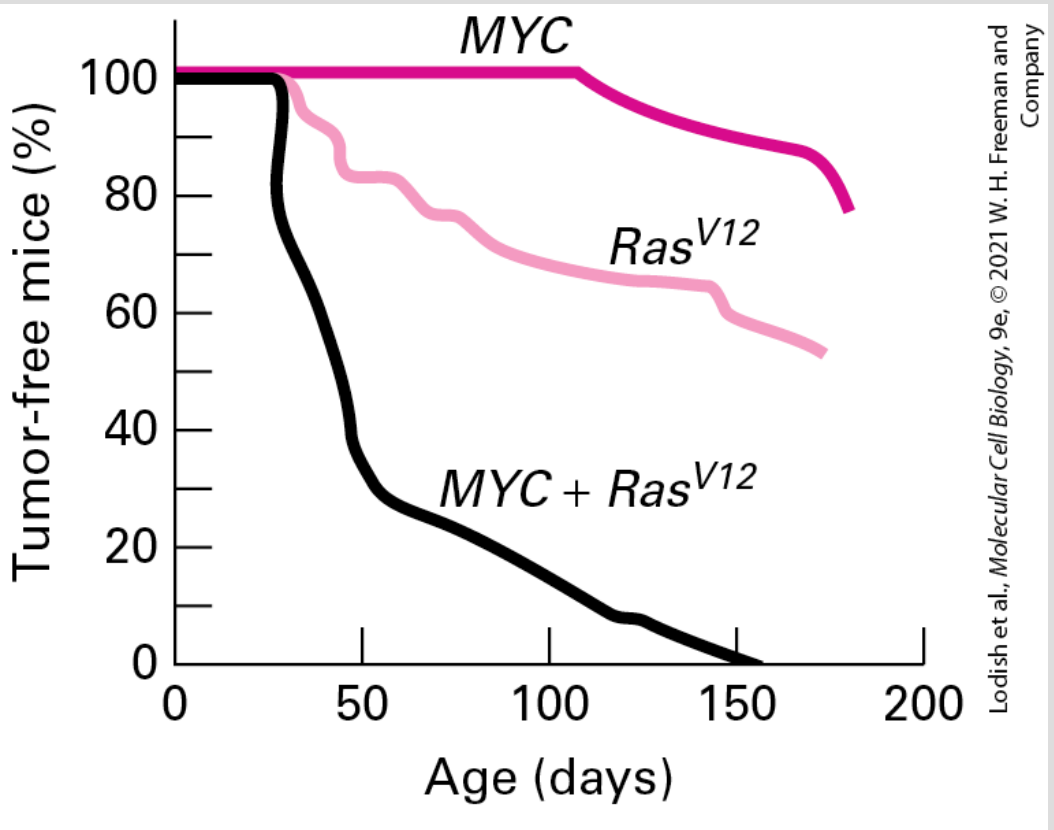
FIGURE 25-19 The incidence of human cancers increases as a function of age. The marked increase in the incidence of cancer with age is consistent with the multi-hit model of cancer induction. Note that the logarithm of annual incidence is plotted versus the logarithm of age.

[Data from B. Vogelstein and K. W. Kinzler, 1993, *Trends Genet*. **9**:138–141.]

Experimental reconstruction of a multi-hit progression has been carried out with transgenic mice, which have shown that a variety of combinations of oncogenes can cooperate in causing cancer. For example, mice have been made that carry either the mutant RAS^{V12} dominant oncogene or the *MYC* proto-oncogene, in each case under the control of a mammary-cell-specific promoter/enhancer. This promoter is induced by endogenous hormone levels and tissue-

specific regulators, leading to overexpression of *MYC* or RAS^{V12} in breast tissue.

As noted, the MYC protein is a transcription factor that induces expression of many genes required for the transition from the G_1 to the S phase of the cell cycle. Heightened transcription of *MYC* in these mice mimics previously identified oncogenic mutations that increase *MYC* transcription, converting the proto-oncogene into an oncogene. By itself, the *MYC* transgene causes tumors only after 100 days, and then in only a few mice; clearly only a minute fraction of the mammary cells that overproduce the MYC protein actually become malignant. Production of the mutant Ras^{V12} protein alone causes tumors earlier, but still slowly and with about 50 percent efficiency over 150 days. When the *MYC* and Ras^{V12} overexpressing transgenic mice are crossed, however, all mammary cells in their offspring overproduce both MYC and Ras^{V12}. In these mice, tumors arise much more rapidly, and all animals succumb to cancer (Figure 25-20). Such experiments emphasize the synergistic effects of multiple oncogenes. They also suggest that the long latency of tumor formation, seen even in the double-transgenic mice, is due to the need to acquire still more mutations.

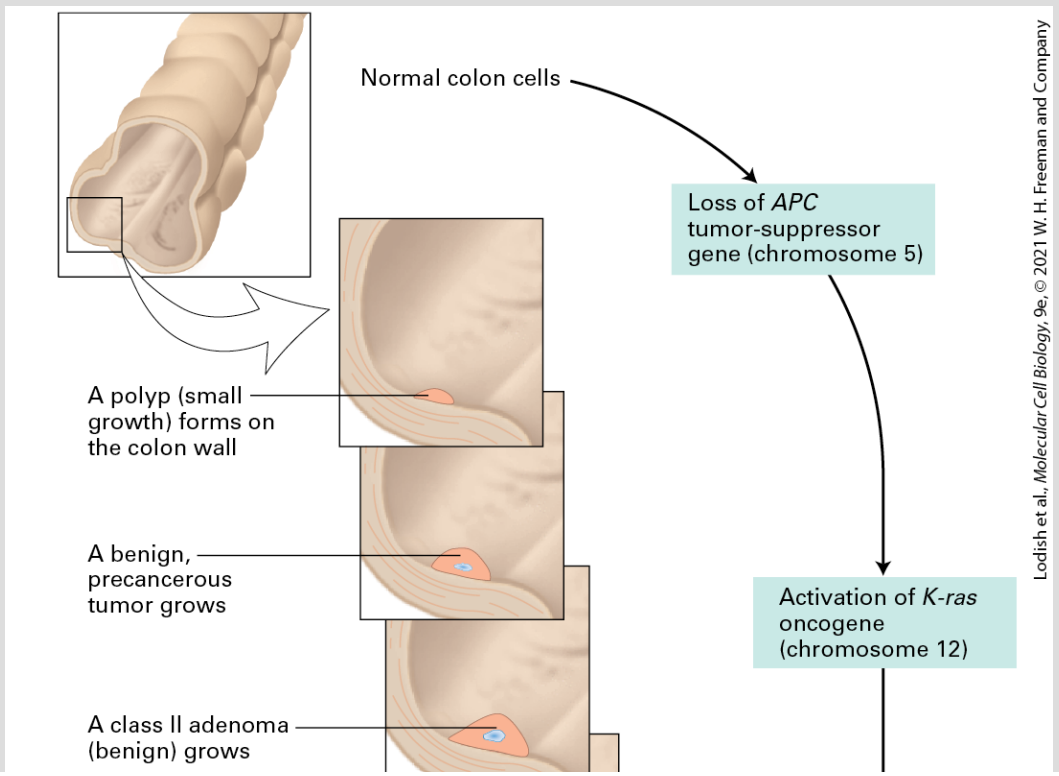


EXPERIMENTAL FIGURE 25-20 The kinetics of tumor appearance in female mice carrying either one or two oncogenic transgenes shows the cooperative nature of multiple mutations in cancer induction. Each of the transgenes was driven by the mouse mammary tumor virus (MMTV) breast-specific promoter. The hormonal stimulation associated with pregnancy activates the MMTV promoter and hence the overexpression of the transgenes in mammary tissue. The graph shows the time course of tumorigenesis in mice carrying either *MYC* or Ras^{V12} transgenes as well as in the progeny of a cross of *MYC* carriers with Ras^{V12} carriers, which contain both transgenes. The results clearly demonstrate the cooperative effects of multiple mutations in cancer induction. See E. Sinn et al., 1987, *Cell* **49**:465.

The Succession of Oncogenic Mutations Can Be Traced in Colon

Cancers

A direct way to reconstruct the progression of a tumor is to take sequential samples from a slowly developing tumor. Surgeons can obtain fairly pure samples of many human cancers, but since the tumor is observed at only one time, its exact stage of progression cannot be easily determined. An exception is colon cancer, which evolves through distinct, well-characterized morphological stages. Its intermediate stages — polyps, benign adenomas, and carcinomas — can be isolated by a surgeon during colonoscopies, allowing mutations that occur in each of these stages to be identified. Numerous studies have shown that colon cancer arises from a series of mutations that commonly occur in a well-defined order, providing strong support for the **multi-hit model** (Figure 25-21).



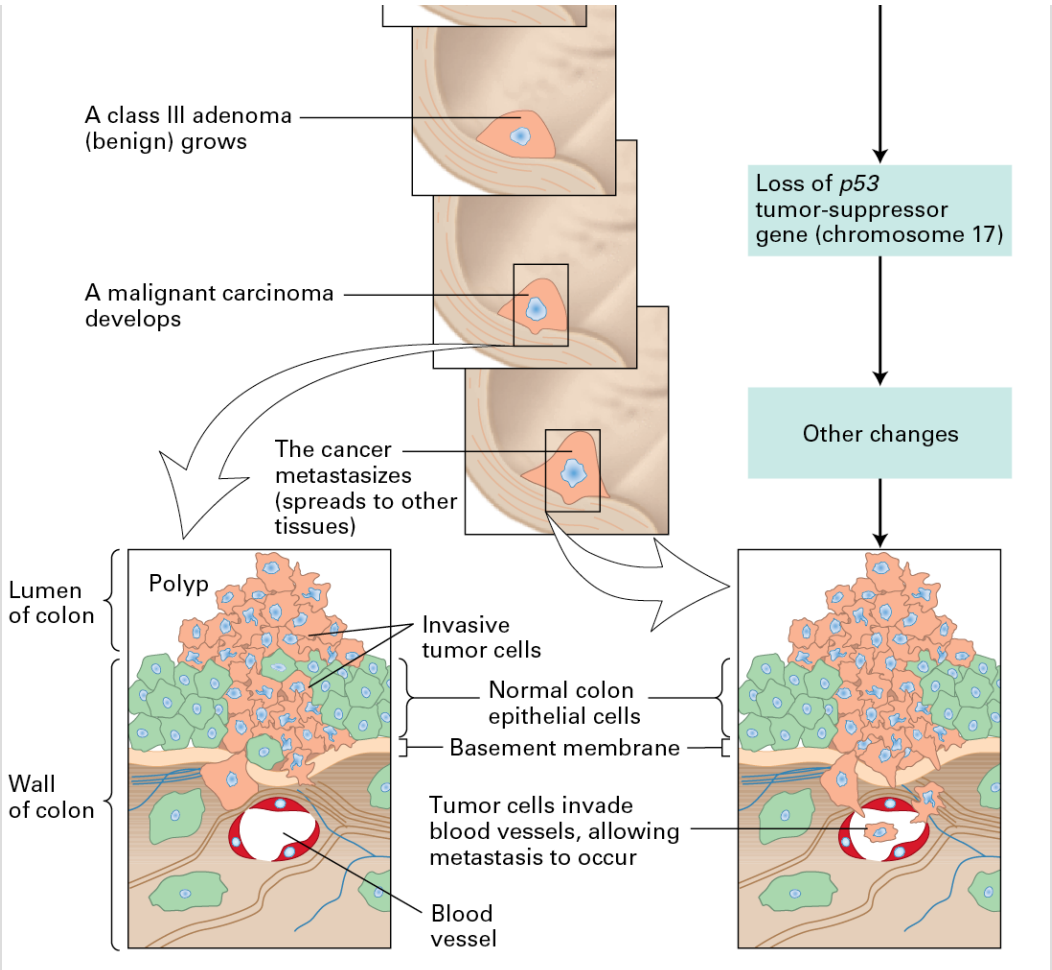


FIGURE 25-21 The development and metastasis of human colorectal cancer and its genetic basis. A mutation in the *APC* tumor suppressor gene in a single epithelial cell causes the cell to divide (although surrounding epithelial cells do not), forming a mass of localized benign tumor cells, called a *polyp*. Subsequent mutations lead to expression of a constitutively active Ras protein and loss of the tumor suppressor gene *p53.* These mutations, together with additional genetic changes yet to be identified, generate a malignant cell. The cell continues to divide, and its progeny invade the basement membrane that surrounds the tissue, but do not penetrate the basement membrane of capillaries (*bottom left*). Some tumor cells spread into blood vessels that will distribute them to other sites in the body (*bottom right*). Additional mutations permit the tumor cells to exit from the blood vessels and proliferate at distant sites. See B. Vogelstein and K. W. Kinzler, 1993, *Trends Genet*. **9**:138–141.

Insight into the progression of colon cancer first came from the study of inherited predispositions to colon cancer such as familial adenomatous polyposis (FAP). Mutations in the Wnt signaling pathway have been identified in many of these syndromes, and it is now believed that deregulation of Wnt signaling results in formation of polyps (precancerous growths) on the inside of the colon wall — not only in people with inherited polyposis syndromes, but also in people afflicted with sporadic (noninherited) forms of colon cancer. The APC (*a*denomatous *p*olyposis *c*oli) protein is a negative regulator of Wnt signaling (see Figure 16-26), which promotes cell cycle entry by activating expression of the *MYC* gene. The absence of functional APC protein thus leads to inappropriate production of MYC, and cells homozygous for *APC* mutations proliferate at a rate higher than normal and form polyps. Loss-of-function mutations in the *APC* gene are the most frequent mutations found in early stages of colon cancer. Most of the cells in a polyp contain the same one or two mutations in the *APC* gene that result in its loss or inactivation, indicating that they are clones of the cell in which the original mutation occurred. Thus *APC* is a tumor suppressor gene, and both alleles of the *APC* gene must carry an inactivating mutation for polyps to form because cells with one wild-type *APC* allele express enough APC protein to function normally.

If one of the cells in a polyp undergoes another mutation, this time an activating mutation of the *ras* gene, its progeny divide in an even more uncontrolled fashion, forming a larger adenoma. Inactivation

of the *p53* gene follows and results in the gradual loss of normal regulation and the consequent formation of a malignant carcinoma (see Figure 25-21). The p53 protein is a tumor suppressor that halts progression through the cell cycle in response to DNA damage and other stresses and can induce apoptosis (see Section 25.4). While the three "hits" listed here are certainly crucial parts of the picture, there are likely to be additional contributing genetic events. Not every colon cancer, however, acquires all the later mutations or acquires them in the order depicted in Figure 25-21. Thus different combinations of mutations may result in the same phenotype.

DNA from different human colon carcinomas generally contains mutations in all three genes mentioned here — loss-of-function mutations in the tumor suppressors *APC* and *p53* and an activating (gain-of-function) mutation in the oncogene *K-RAS* (one of the *RAS* family of genes) — establishing that multiple mutations in the same cell are needed for the cancer to form. Some of these mutations appear to confer growth advantages at an early stage of tumor development, whereas other mutations promote the later stages, including invasion and metastasis, which are required for the malignant phenotype.

For the majority of tumors that are only biopsied or removed at a late stage in cancer development there are now sequencing-based methods that allow some of the events in cancer progression to be reconstructed. The idea is to gain information from multiple

individual cells from the tumor to create a lineage map for the tumor as a whole. To see how this analysis works, imagine a simple case for a tumor that arose as the result of two oncogenic driver mutations. At the time that the tumor is biopsied, all of the cells should have the first driver mutation that occurred, whereas only some of the tumor cells would have the second. This method will allow the order of appearance to be determined for any two oncogenic driver mutations as long as at least a few cells in the tumor can be found that lack one of the mutations. This method for reconstructing cell lineages within a tumor is now feasible because of advances in next generation sequencing that enable genomic sequence information to be obtained from a single tumor cell.

Cancer Development Can Be Studied in Animal Models

Genetically engineered mice have also provided tremendous insights into the steps of tumor initiation and progression. Using mouse models to study cancer is not always straightforward, however. Many tumor suppressor genes serve essential functions during normal mouse development, so mice lacking both copies of these genes are not viable. The essential functions of these genes during early embryogenesis preclude the study of their role in tumor progression. To circumvent this problem, researchers have begun to employ conditional *knock-in* and *knockout* strategies that

allow for the targeted activation or inactivation of a gene in a certain tissue or at a certain stage of development.

In the conditional mouse model, an allele of a particular oncogene or tumor suppressor gene is wild type until activated or inactivated with exogenous chemicals or viruses in a tissue- or time-specific manner. At the heart of these conditional systems are the Cre and FLP recombinases. These recombinases facilitate homologous recombination between *loxP* and *FRT* sites, respectively (Figure 25-22; see also Figure 6-40). When the recombinases are under the control of a tissue-specific promoter, recombination occurs only in the tissue that produces the recombinase. The recombinase method can be used in two ways. First, the recombinase target sites may flank an exon. Upon induction of the recombinase, that exon is lost and the gene is inactivated (Figure 25-22a). This method is especially useful for inactivating tumor suppressor genes in a tissue-specific manner. Second, expression of an oncogene can be controlled by introducing into the oncogene an additional exon that contains a stop codon, which makes the gene nonfunctional. However, if the additional exon is flanked by recombinase target sites, the oncogene will be expressed upon induction of the recombinase (Figure 25-22b). Using this system, researchers have examined the role of oncogenic forms of Ras in the mouse and have, using a conditional oncogenic *ras* allele, created a mouse model of human lung cancer.

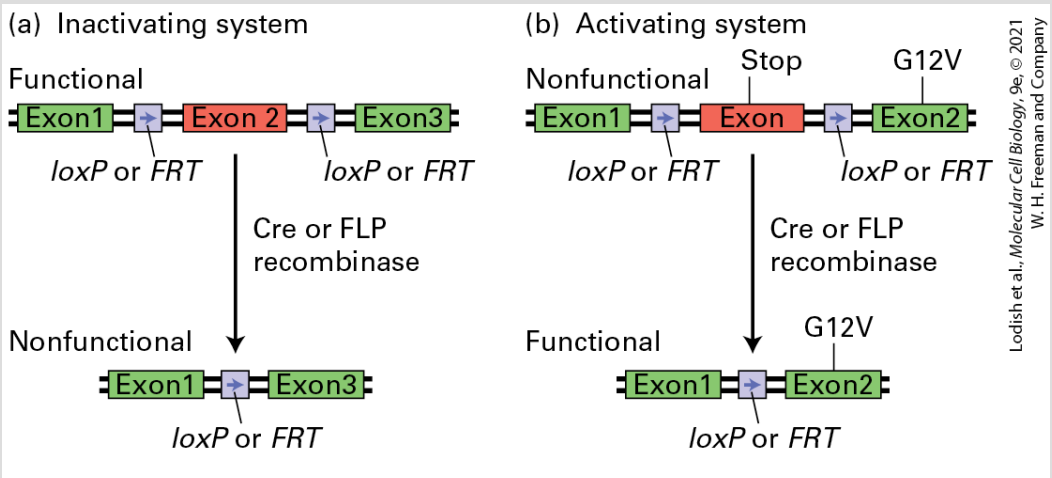


FIGURE 25-22 Conditional mouse models of cancer. In the inactivating system (a), an exon of interest is flanked by two *loxP* or *FRT* sites as shown. Expression of the Cre or FLP recombinase leads to homologous recombination between the two *loxP* and *FRT* sites, respectively. This recombination leads to excision of the exon, rendering the gene nonfunctional. In the activating system (b), an additional exon with a stop codon is introduced into the gene of interest, making the gene nonfunctional. This exon is flanked by *loxP* or *FRT* sites. When Cre or FLP recombinase is induced, the stop codon–containing exon is recombined out, and the gene of interest is expressed.

Molecular Cell Biology Is Changing How Cancer Is Diagnosed and Treated

The identification of oncogenic driver mutations, such as mutations in genes that regulate cell growth and development pathways, has not only provided us with a molecular understanding of how cancer arises and progresses, but has also revolutionized the way cancers are diagnosed and treated. Historically, tumors were characterized and treated based on the organ in which they arose and the

histologic characteristics of the tumor cells when they were viewed under the microscope. We now know that similarly appearing tumors may behave very differently and can be better separated into specific tumor subsets by examining their genomic changes. This type of characterization, which is extensively used in breast and brain cancer diagnosis, has significant implications for treatment. Each difference between the cancer cells and normal cells provides a new opportunity to identify a specific drug or treatment that kills only the cancer cells or at least stops their uncontrolled growth. Thus knowledge of the molecular cell biology of a tumor is critical information that can be exploited by researchers to develop anticancer treatments that more precisely target cancer cells.

Breast cancer provides a good example of how molecular cell biology techniques have affected treatments, both curative and palliative. Breast cancer is the second most frequent cause of women's cancer deaths. Breast cancers are often diagnosed during routine mammogram (x-ray) examinations. Typically, a very small biopsy is taken with a needle to confirm the diagnosis and is tested with antibodies to determine whether a high level of estrogen or progesterone receptors is present. These steroid receptors are capable of stimulating tumor growth and are sometimes expressed at high levels in breast cancer cells. If either receptor is present, it is exploited in the treatment. A drug called *tamoxifen*, which inhibits the estrogen receptor (see Figure 8-26), can be used to deprive the tumor cells of this growth-stimulating hormone. The biopsy is also

tested for amplification of the proto-oncogene *HER2/NEU*, which, as we saw in Chapter 16, encodes the human EGF receptor HER2. A monoclonal antibody specific for HER2 has been a strikingly successful new treatment for the subset of breast cancers that overproduce HER2. HER2 antibody injected into the blood recognizes HER2 and causes it to be internalized, selectively killing the cancer cells without any apparent effect on normal breast (and other) cells that produce only moderate amounts of HER2. More recently, covalent conjugates of HER2 monoclonal antibodies with a potent cytotoxic drug became available to target these agents into HER2-overproducing tumor cells, enhancing the killing effects of these monoclonal antibodies. These advanced antibody conjugates are used as a second line of defense for tumors that have resisted more conventional treatments.

Breast cancer is still initially treated with a combination of surgery, radiation therapy, and chemotherapy. The first step is surgical resection (removal) of the tumor and examination of lymph nodes for evidence of metastatic disease. The subsequent treatment typically involves 8 weeks of chemotherapy with several different types of agents and 6 weeks of radiation. These harsh treatments are designed to kill the dividing cancer cells; however, they also cause a variety of side effects, including suppression of blood cell production, hair loss, nausea, and neuropathy. These effects can reduce the strength of the immune system, risking infection, and cause weakness due to poor oxygen supply. To help offset these

effects, patients are given the cytokine G-CSF to promote the formation of neutrophils (a type of white blood cell that fights bacterial and fungal infections) and erythropoietin (Epo) to stimulate red blood cell formation (see Chapter 16). Even with this aggressive treatment, about a third of women will nevertheless succumb to cancer over the next 10 years. This risk can be reduced by 15 percent by hormone-blocking treatment such as tamoxifen and by another 10 percent by treatment with antibodies against the HER2/NEU oncoprotein. In aggregate, advanced molecular methods can reduce morbidity from breast cancer by as much as 25 percent.

KEY CONCEPTS OF SECTION 25.3

Dysregulation of Cell Growth and Developmental Pathways Initiates Tumorigenesis

- Mutations that permit receptors for growth factors to dimerize in the absence of their normal ligands lead to constitutive receptor activity. Overproduction of growth-factor receptors can have the same effect and lead to abnormal cell proliferation.
- Most tumor cells produce constitutively active forms of one or more intracellular signal-transducing proteins, causing growth-promoting signaling in the absence of normal growth factors.
- A feature of the proliferation of most cancers is unregulated passage through the restriction point of the cell cycle. This can be caused by gain-of-function mutations in cyclin D1 or loss-of-function mutations in p16 or RB.
- Inappropriate production of nuclear transcription factors such as FOS, JUN, and MYC can induce transformation. In Burkitt's lymphoma cells, *MYC* is translocated close to an antibody gene, leading to overproduction of MYC.
- Loss of signaling by TGF-β, a negative growth regulator, promotes cell proliferation and development of malignancy.

- A typical tumor carries about five oncogenic mutations in genes of cell growth and development pathways that are acquired over many years. A major unsolved problem in understanding how cancer progresses will depend on reconstruction of the sequential order of these mutational events.
- Colon cancer develops through distinct, identifiable morphological stages that can be traced to specific oncogenic mutations.
- Genome sequencing of multiple individual cells in a tumor can enable the lineages of cells with specific oncogenic mutations to be reconstructed.

25.4 Evasion of Programmed Cell Death and Immune Surveillance Processes

So far in this chapter, we have seen that early stages of the progression of cancer involve two types of oncogenic mutations: those discussed in Section 25.2 that compromise DNA repair and lead to destabilization of the genome, and those discussed in Section 25.3 that activate growth-promoting pathways and inactivate growth-inhibiting pathways. In addition to these mutations that drive oncogenesis, many cells that have reached this precancerous stage also carry a significant burden of adventitious mutations, chromosomal rearrangements, and other defects in chromosomal DNA. Cells that carry significant chromosomal abnormalities or irreparable damage to their DNA will enter the apoptotic pathway and be targeted for programmed cell death. Cells that carry many point mutations in protein-coding genes will express mutant proteins that may cause them to be recognized and targeted for destruction by cells of the immune system. Although we have little direct visibility into what happens to small numbers of precancerous cells when they do arise, it is likely that the vast majority of them are efficiently eliminated by one or both of these mechanisms. For this reason, the tumor cells that we eventually do observe owe their existence to having acquired mutations that allowed them to evade both apoptosis and immune surveillance.

The acquisition by most cancer cells of mutations that block the apoptotic pathway and of other mutations that suppress an immune response to them is a powerful example of how cancer progression results from an evolutionary process of mutation and selection. We will see in this section some of the kinds of oncogenic mutations that allow tumor cells that carry a high mutational load to survive. At the end of the section, we will highlight some of the promising new therapies that are based on reactivating the immune system to target cancer cells.

Oncogenic Driver Mutations Enable Cancer Cells to Evade Apoptosis

As we saw in Chapter 22, programmed cell death, or apoptosis, plays an important role in development. For example, during mammalian limb formation apoptosis is responsible for sculpting the digits out of a paddle-shaped limb bud. In addition, apoptosis serves an editing function to maintain the organization of mature tissues by removing abnormal excesses of cells not needed for development of a working organ and by removing cells that have been compromised by errors in mitosis or DNA damage (see Chapter 22). For some cells, apoptosis appears to be the outcome by default, and signals from the PI 3-kinase/AKT pathway are required to ensure cell survival (see Figure 22-42).

A major class of oncogenic mutations blocks apoptosis, allowing tumor cells to continue to proliferate inappropriately. For example, chronic lymphocytic leukemia (CLL) occurs because cells survive when they should not. The cells accumulate slowly, and most are not actively dividing, but they do not die. CLL cells have chromosomal translocations that activate the *BCL2* gene, a critical blocker of apoptosis (see Figure 22-42). The resultant inappropriate overproduction of the BCL2 protein prevents normal apoptosis and allows survival of these tumor cells. CLL tumors are therefore attributable to a failure of cell death. Another dozen or so proto-oncogenes that are normally involved in negatively regulating apoptosis have been found to be mutated to become oncogenes. Overproduction of their encoded proteins prevents apoptosis even when death is needed to stop cancer cells from growing.

p53 Can Activate Either the DNA Damage Checkpoint or Apoptosis in Response to DNA Damage

We have already seen that the p53 protein is a central player in tumorigenesis, and it is thought that most, if not all, human tumors have mutations either in p53 itself or in proteins that regulate p53 activity. In Section 25.2, we saw that p53 has a major part in causing an arrest in the cell cycle in response to DNA damage. Because cells carrying oncogenic p53 mutations do not arrest to give enough time

for the DNA damage to be repaired, these mutants have unstable genomes and accumulate somatic mutations. p53 mutant cells also do not respond to extensive DNA damage by entering the apoptotic pathway and thus can continue to proliferate even after they have sustained damage that would cause a normal cell to be removed by programmed cell death.

Although the oncogenic form of p53 is due to a loss of function, it is nevertheless possible for certain missense mutations in only one of the two *p53* alleles in a cell to abrogate almost all p53 DNA activity. This behavior comes from the fact that the active form of p53 is tetramer of four identical subunits and mixed complexes that contain at least one defective subunit have reduced ability to activate transcription. Precancerous cells with one missense mutant of p53 are probably selected for because of a proliferative growth advantage. However, the loss of function is incomplete and in order to proliferate more rapidly tumor cells often lose the remaining functional allele. This unusual feature of positive selection for a first and then a second mutation in the p53 gene may explain why tumor cells so frequently carry oncogenic driver mutations in p53.

We do not yet understand why p53 causes cell cycle arrest under some circumstances and apoptosis under others, but both responses can occur as a result of p53 being stabilized in response to DNA damage. Unlike other cell cycle proteins, p53 is present at very low levels in normal cells because it is extremely unstable and rapidly

degraded. The activity of p53 is normally kept low by a protein called *MDM2*. When MDM2 is bound to p53, it inhibits the transcription-activating ability of p53 and at the same time, because it has E3 ubiquitin ligase activity, catalyzes the ubiquitinylation of p53, thus targeting it for proteasomal degradation. After DNA damage, ATM or ATR are recruited to the sites of damage and their serine kinase activity is activated. Phosphorylation of p53 on a serine residue in the N-terminus of the protein by ATM or DNA-PK displaces bound MDM2 from p53, thereby stabilizing it (Figure 25-23). The stabilized p53 activates transcription of the gene encoding p21, which binds to and inhibits mammalian CDK2, CDK1, and CDK4/6 complexes (see Figure 19-34).

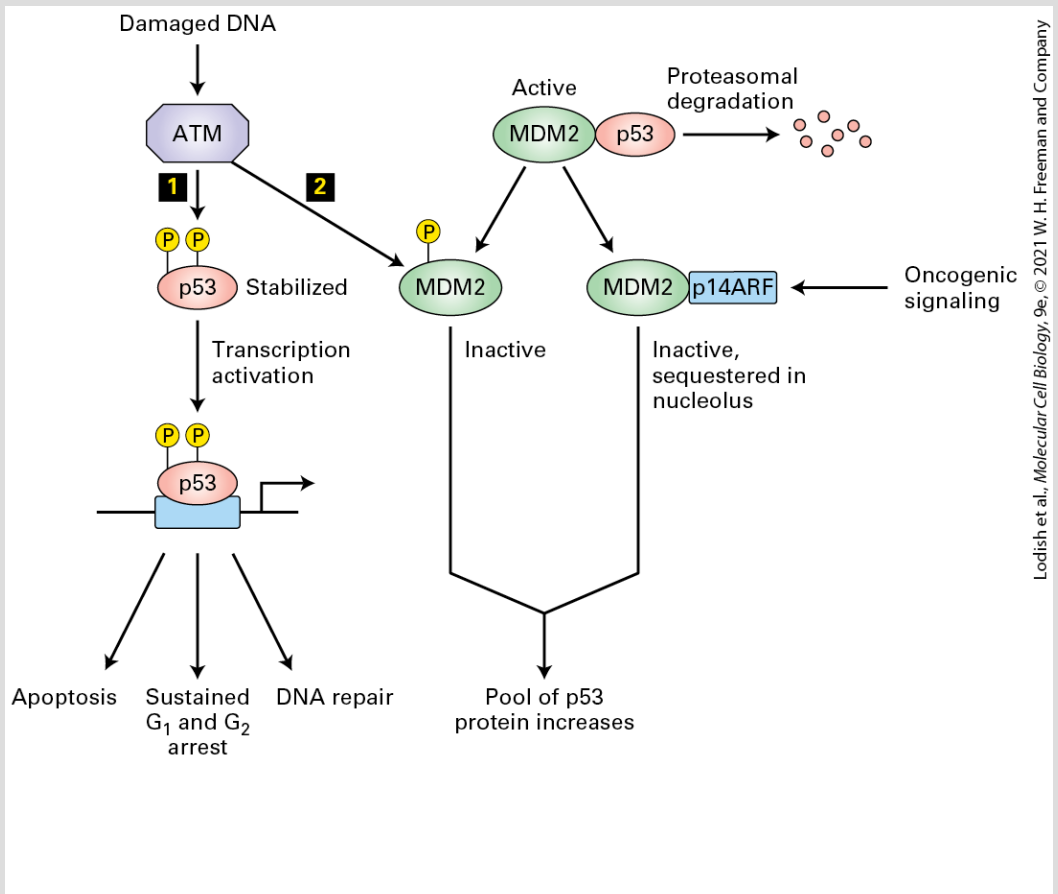


FIGURE 25-23 Arrest in G_1 in response to DNA damage. The kinase activity of ATM, ATR, and DNA-PK is activated in response to DNA damage due to various stresses (e.g., UV irradiation, heat). Activated ATM and ATR then trigger two principal pathways leading to arrest in G_1. Step **1**: Phosphorylation of p53 stabilizes it, permitting p53-activated expression of genes encoding proteins that cause arrest in G_1, promote apoptosis, or participate in DNA repair. Step **2**: The second pathway is another way of controlling the pool of p53. The MDM2 protein in its active form can form a complex with p53, inhibiting its transcription-factor activity and causing p53 ubiquitinylation and subsequent proteasomal degradation. ATM and DNA-PK phosphorylate MDM2 to inactivate it, causing increased stabilization of p53. In addition, MDM2 levels in the cytoplasm are controlled by ATR and p14ARF (p19ARF in the mouse), sequesters it in the nucleolus, where it cannot access p53. The p14ARF gene is induced by high levels of mitogenic signaling, which are frequently observed in cells carrying oncogenic mutations in growth factor signaling pathways. The human *MDM2* gene is frequently amplified in sarcomas, which presumably causes excessive inactivation of p53. Similarly, p14ARF is also found mutated in some cancers.

Stabilized p53 also can induce expression of proteins such as Bax and PUMA that bind to and inhibit Bcl2 and thus induce apoptosis (see Figure 22-42). When cells suffer extensive DNA damage, the p53-induced expression of pro-apoptotic proteins leads to their quick elimination. While apoptosis may seem like a drastic response to DNA damage, it prevents proliferation of cells that are likely to accumulate multiple mutations. When p53 function is lost, apoptosis cannot be induced, and the accumulation of additional mutations required for cancer to develop and progress becomes more likely.

Apoptosis can be triggered in response to a number of intrinsic and extrinsic signals. For example, we will see in the next section that cytotoxic T cells can kill a target by inducing apoptosis in the targeted cell.

The Immune System Is a Second Line of Defense Against Cancer Formation

We have seen that tumor cells carry thousands of mutations, including point mutations, Indels, and chromosomal structural variants (see Figure 25-8). For a typical tumor with 5000 mutations, on average about 50 of these would lie in protein-coding sequences and thus may give rise to a protein product that was altered in some way. For tumors that arise from exposure to mutagens, such as melanomas (UV light) and lung carcinomas (tobacco smoke) the number of mutations may be hundredfold greater than this. These altered proteins are known as *neo-antigens* and would be expected to cause the tumor cell to be recognized and eliminated by the immune system in much the same way as cells that contain foreign proteins induced by infection with a pathogen such as a virus or bacterium specifically targeted for destruction.

Most precancerous cells bearing neo-antigens are thought to be efficiently eliminated by immune surveillance. Evidence for this comes from studies of immunodeficient mouse mutants, which are

more likely to acquire carcinogen-induced cancer than wild-type mice. This raises the crucial question of how did the tumors that we do see escape immune surveillance? In the following section, we discuss how the local environment within the tumor suppresses the immune system, and how monoclonal antibodies that recognize molecules enriched on the surface of tumor cells can redirect certain immune system cells to kill the tumor cell. Finally, we discuss newer ways of genetically engineering T cells of the immune system so they express on their surface monoclonal antibodies that enable these cells to recognize and kill many types of tumor cells.

The Tumor Microenvironment and Immunoediting Limit the Ability of the Immune System to Detect and Kill Established Tumors

As solid tumors develop, they recruit a variety of other cell types in addition to the cancerous cells themselves. These cells, which include fibroblasts from the organ in which the tumor develops, as well as a wide variety of different types of cells from the immune system, transform the tumor into a type of tissue, rather than simply a clone of isolated cancer cells. The non-cancer cells that surround the malignant tumor cells comprise what is called the tumor *microenvironment*. Cancer cells influence the behavior of the fibroblasts and immune cells in the microenvironment, and

conversely, cells in the tumor microenvironment influence the way the cancer cells evolve, particularly with respect to immune recognition.

During the very early stages of tumor development, inflammatory immune cells including neutrophils, macrophages, and dendritic cells respond to the presence of these abnormal tumor cells by releasing cytokine molecules that inhibit tumor cell proliferation and angiogenesis and send signals to activate various types of T cells. A particularly important class of T cells involved in killing tumor cells are called *cytotoxic T cells*; they are distinguished by the presence of the CD8 protein on their surface (see Chapter 24). To understand the mechanisms by which cytolytic T cells potentially can kill tumor cells and why this often does not occur, we must review some of the salient points by which T cells in general, and cytotoxic T cells in particular, recognize foreign molecules.

Cytotoxic T Cells Target Tumor Cells

As detailed in Chapter 24, T cells recognize not intact proteins but peptide fragments of these proteins of about 10 amino acids in length. Cells constantly degrade representative samples of all cellular proteins — whether they are made in the cell or brought in by pathogens — to peptides of this size, which are then bound to one of several two-chain glycoprotein complexes called *MHC molecules*. These MHC-peptide complexes are then transported to the cell

surface where they can be detected by passing T cells. In the case of peptides derived from mutant proteins that are displayed on the surface of the tumor cells themselves, these are generally bound to class I MHC molecules (see Figure 24-26). Because most tumor cells, like all nucleated cells in the body, express MHC class I molecules on their surface, cytotoxic T cells can directly recognize tumor cells expressing neo-antigens. However, only a small fraction of neo-antigens appear to be detected by T cells as foreign. Therefore, only tumors with a high mutational burden, such as melanomas, reach the threshold for recognition as a suitable target for destruction by immune surveillance.

Engagement of tumor cells bearing foreign tumor antigens on their MHC class I molecules by cytotoxic T cells can result in killing of the tumor cells by two different mechanisms. Some activated cytotoxic T cells release granules containing the proteins perforin and granzyme-B. Perforin and granzymes enter a target cell through endocytosis, and perforins punch holes in the endocytic vesicle membrane, allowing granzymes to enter the cytoplasm. Granzymes are serine proteases and cleave and activate apoptotic molecules within the target cells, causing programmed cell death (see Figure 24-36). Other cytotoxic T cells express on their surface FasL, a natural ligand molecule related to TNF. When FasL on the cytotoxic T cell binds to its receptor, the Fas protein on the tumor cells, an apoptotic signaling pathway in the tumor cells becomes activated, which also leads to tumor cell death. During early stages of solid tumor development, these processes effectively eliminate many of

the incipient tumor cells. As a consequence, the developing tumor and the immune system reach balance where the tumor does not grow but is not completely eliminated.

Tumor Cells Escape from Killing by Cytotoxic T Cells

Even as growth plateaus, tumor cells continue to divide in a steady state balance with immune-mediated killing. Eventually, the tumor cells may acquire additional oncogenic mutations that allow them to escape the immune system and to begin to proliferate again. This immune evasion can occur through a variety of mechanisms. In some cases, the tumor cells send signals that change the types of immune cells that are recruited into their microenvironment, so that pro-inflammatory macrophages and neutrophils are replaced with anti-inflammatory, wound-healing neutrophils and macrophages, which secrete a different set of cytokines that fail to recruit T cells into the tumor. In other cases, the tumor down-regulates the expression of class I MHC proteins, rendering the tumors invisible to cytotoxic T cells. Sometimes the tumor cells up-regulate the expression of anti-apoptotic proteins like *BCL2* that blunt the killing effects of granzymes, perforins, and FasL.

Two other mechanisms that tumor cells use to survive involve a direct attack by tumor cells on the signaling mechanisms that T cells use to kill them. In order for cytotoxic T cells to become licensed to

kill cells bearing foreign antigens, they require a second co-stimulatory activation signal in addition to recognition of peptide fragments of tumor antigens presented on class I MHC molecules (see Figure 24-35 for more detail). This second check before killing probably exists to prevent cytotoxic T cells from accidently killing normal cells and causing autoimmune injury to normal tissues. One of the best understood co-stimulatory signals comes from CD28, a cell-surface protein on T cells that binds to either of two cell-surface proteins that are on the surface of antigen-presenting cells, CD80 and CD86 (Figure 25-24, step **3**). Only when T cells see MHC-bound foreign tumor antigens and receive a signal through CD28 do they become activated for cytokine production and T-cell–mediated killing. Within a day after activation, T cells up-regulate a different cell-surface protein, CTLA-4, which competes with CD28 for binding to CD80 and CD86 and sends a negative signal that inhibits T-cell responses (Figure 25-24, step **4**). This is important because it restrains the extent of immune activation in order to limit inflammatory tissue damage. Because of the up-regulation of CTLA-4 on anti-tumor cytotoxic T cells, they become self-inhibited in killing tumor cells.

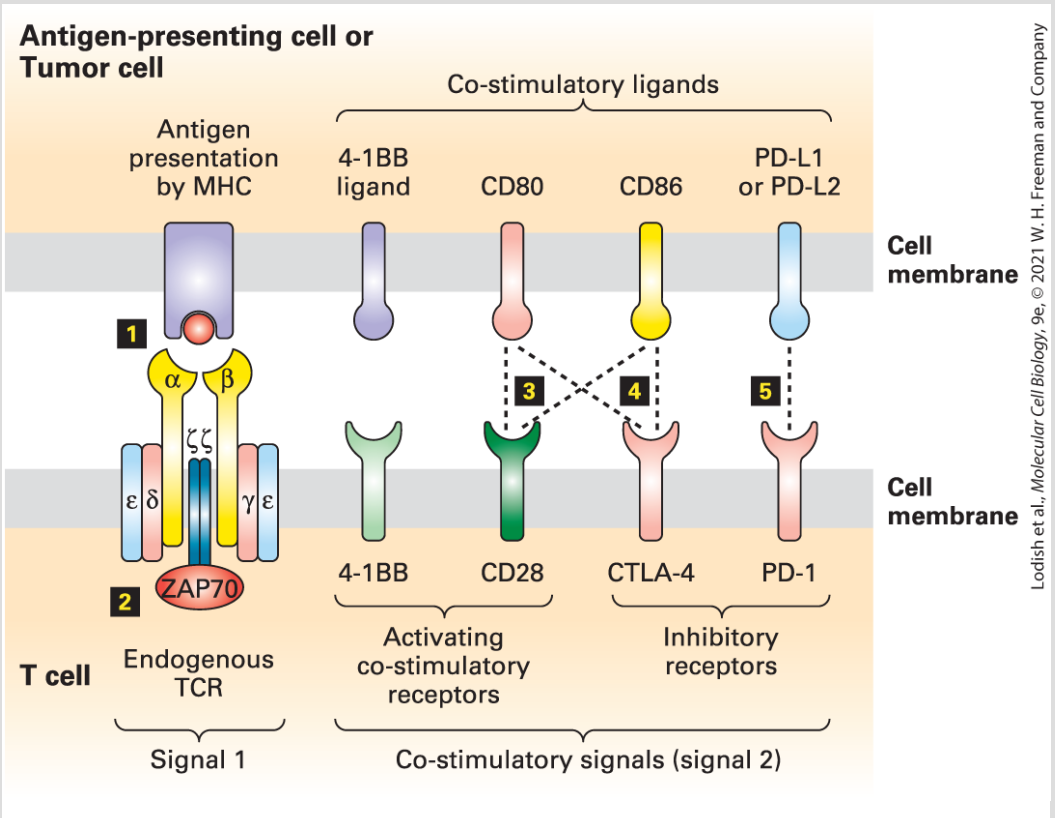


FIGURE 25-24 Overview of T-cell signaling. Activation of a T cell is initiated by the T-cell receptor (TCR), which contains one α and one β chain, each of which differs slightly in sequence in different T cells of an individual. Each αβ TCR binds to a single or group of related target peptides bound ("presented") on the surface of an MHC protein on an antigen presenting cell (step **1**). TCRs also contain signaling ζ chains as well as γ, δ, and ε chains. Binding of the TCR to a MHC-peptide complex triggers activation of the ZAP70 protein tyrosine kinase attached to the cytosolic domain of the ζ chains (step **2**); in a cytotoxic T cell this triggers a signal transduction pathway (signal 1) essential for activation of the ability of the T cell to kill the antigen-presenting cell. Activation of the T cell also requires a second signal, triggered by binding of one of several proteins on the antigen-presenting cell, predominantly CD80 or CD86, to the CD28 protein on the surface of the T cell (step **3**), triggering activation of a second signal transduction pathway that together with the first leads to activation of the T cell. Among the proteins induced during T-cell activation is CTLA-4; binding of CTLA-4 to CD80 or CD86 on the antigen-presenting cell (step **4**) activates another signaling pathway leading to inactivation of the T cell. If the antigen-presenting cell expresses on its surface either the PD-L1 or the PD-L2 protein,

these can bind to the PD-1 protein on the T cell (step **5**) and trigger a signaling pathway that induces apoptosis of the T cell, preventing it from killing the antigen-presenting cell.

A second T-cell surface receptor protein that acts as a negative regulator of T-cell function is PD-1 (*p*rogrammed *d*eath-1). PD-1 has a structure similar to CTLA-4 and when activated negatively regulates T-cell function by inducing apoptosis of activated T cells. The two ligands of PD-1, PD-L1 and PD-L2, are in the same protein family as CD80 and CD86 and are present on the surface of antigen-presenting cells; the ligands thus induce apoptosis of activated cytotoxic T cells, preventing inappropriate killing of antigen-presenting cells by T cells (Figure 25-24, step **5**). Many tumor cells overexpress PD-L1 and thus escape death from cytotoxic T cells by in effect masquerading as antigen-presenting cells.

Both the CTLA-4 pathway and the PD-1/PD-L1 signaling system are called **immune checkpoint pathways** because they normally keep the immune system in check in order to prevent the development of autoimmune diseases. Both pathways are co-opted by tumor cells in order to suppress anti-tumor immunity. This entire process in which tumor cells are initially eliminated by the immune system, reach a steady state plateau, and then eventually escape from immune control by a variety of molecular mechanisms is referred to as **immunoediting**, since the final tumor outgrowths have been shaped by its response to the innate and adaptive immune system.

Activation of the Immune System Presents a Major Opening for Cancer Therapy

As we have seen, the immune system has the capacity, largely based on activation of cytotoxic T cells, to recognize and eliminate precancerous cells, especially those that express abundant neo-antigens. We have also seen that cancer cells that are able to escape immune surveillance and continue to proliferate to form tumors have acquired the capacity to suppress killing by cytotoxic T cells. One avenue for cancer therapy is the reactivation or reengineering of the immune system to recognize and eliminate mature tumors. Here we describe two of the most promising approaches for cancer immunotherapy.

Monoclonal Antibodies Against Immune Checkpoint Molecules to Treat Cancers

One very successful method is to use antibodies that block CTLA4 from binding to CD80 and CD86 and thus prevent cytotoxic T cells from receiving inhibitory signals from CTLA-4. A CTLA-4-blocking antibody, ipilimumab, was approved for the treatment of advanced melanoma. It and similar monoclonal antibodies that block immune checkpoint activation have proven very effective in treating cancers such as melanomas that have a large mutational burden and thus have many neo-antigens recognized by many different T cells.

Monoclonal antibodies directed against either PD-L1 or PD-1 are also effective in treating several types of advanced cancers, particularly those tumors that contain many mutant proteins such as melanomas and certain lung cancers. However, when administered alone, monoclonals against CTLA-4, PD-1, or PD-L1 are effective in treating only a fraction of patients with most cancers. Increasingly, combination therapies, in which these monoclonals are combined with other antibodies and drugs are increasing the efficacy of this type of cancer immunotherapy and may eventually allow even some terminal cancer patients to survive.

Engineering T Cells to Recognize Tumor Antigens

We have just seen that cytotoxic T cells, when reactivated by inhibition of the immune checkpoint, have the capacity to recognize and kill some tumors with a large burden of mutations. However, most tumors, probably because they express an insufficient number of neo-antigens, do not respond to monoclonal antibodies to checkpoint inhibitors such as CTLA-4 or PD-1. An alternative approach is to genetically engineer T cells in the laboratory that can directly attack tumor cells bearing a marker protein on their surface and then infuse them into patients where they can perform their tumor-killing function. These T cells contain specially engineered T-cell receptors, called *c*himeric *a*ntigen *r*eceptors, and the modified cells are called **CAR T cells**. To understand how they work, we need

to first review signaling pathways that connect T-cell receptor binding to anti-tumor activity.

The cytosolic domains of the a and b subunits of the T-cell receptor do not directly activate a signal transduction pathway. Rather, binding of the extracellular TCR domain to a MHC-peptide complex triggers activation of ZAP70 protein tyrosine kinases attached to the cytosolic domains of the two adjacent ζ chains (see Figure 25-24, *bottom left*). The idea behind CARs was to engineer the T-cell receptor to trigger activation of ZAP70 in response to the presence of a tumor-specific antigen rather than the MHC-peptide complex. The first generation of CARs accomplished this by employing the antigen-binding V_H and V_L domains of a monoclonal antibody that binds to a protein enriched on the surface of a tumor cell (e.g., CD19 on a B-cell tumor) linked to the cytosolic domain of the ζ subunit through a transmembrane segment. T cells expressing such a chimeric receptor do activate ZAP70 when they encounter a tumor cell expressing the CD19 protein on its surface (Figure 25-25a), but they fail to kill the tumor cell. Realizing that this failure could be due to the absence of crucial co-stimulatory signals (signal 2 in Figure 25-24), researchers designed chimeric antigen receptors in which the cytosolic domain contained signal-transducing segments from the cytosolic domain of the ζ subunit as well as from one or more of the co-stimulatory receptors such as CD28 or 4-1BB (see Figure 25-24). In clinical trials, such CARs targeting CD19 were expressed using lentiviral vectors in normal T cells isolated from a

large number of patients with several types of human B-cell tumors (Figure 25-25b). Retransfusion of these engineered CAR T cells back into these patients in many cases led to complete remission of the cancer. Currently, researchers are developing CAR T cells expressing CARs targeting a number of proteins enriched on the surface of many types of human tumors. These are being tested against a number of malignancies for which there are presently no cures.

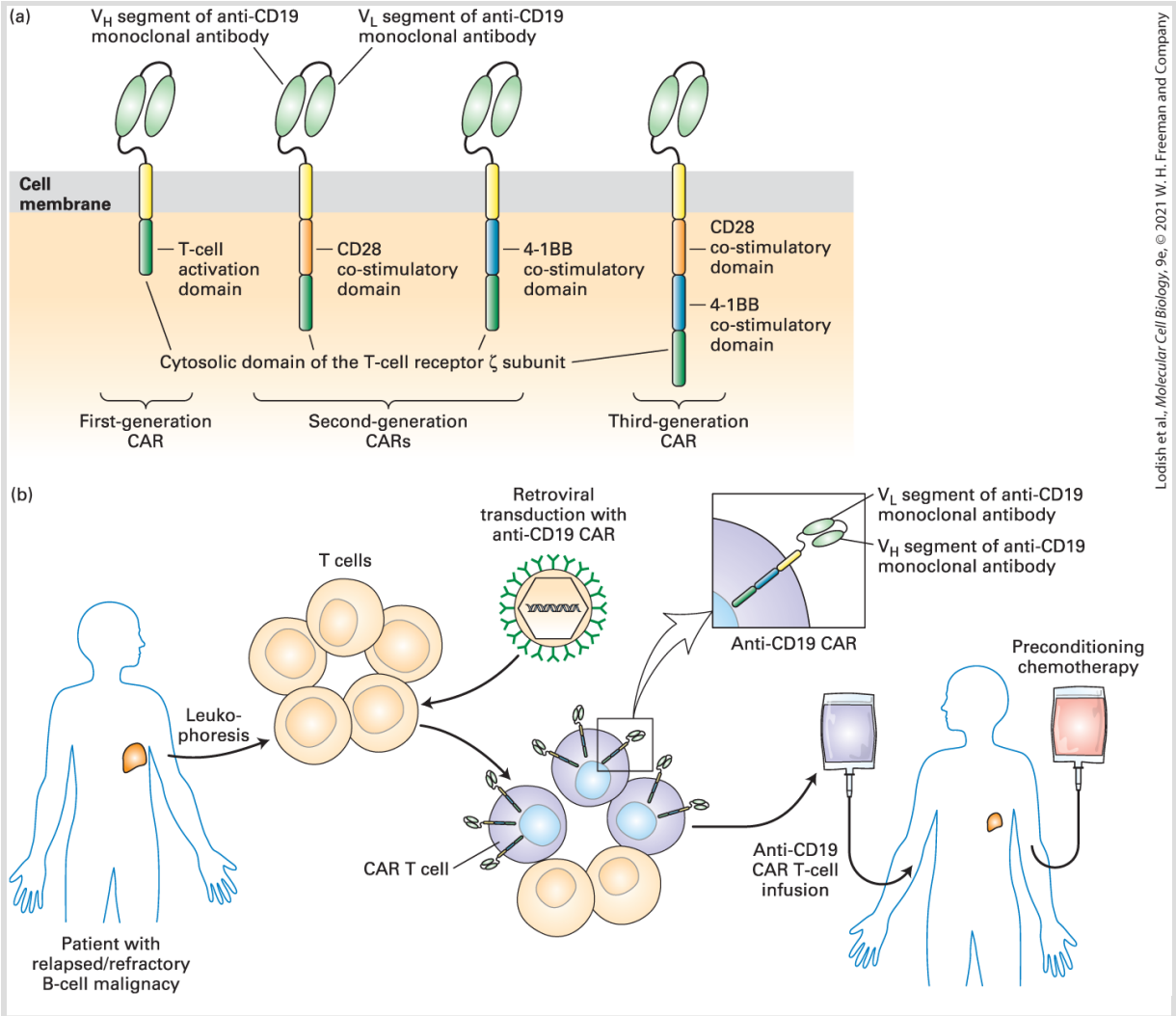


FIGURE 25-25 Three generations of CAR T cells. (a) Chimeric antibodies (CARs) expressed in T cells. The first generation CAR T cells to treat B-cell malignancies expressed on the T-cell surface of a chimeric protein containing the V_L and V_H transmembrane domain and the

cytosolic domain of the ζ subunit of the T-cell receptor (second panel). Recall that binding of the TCR to a peptide MHC complex normally triggers activation of the ZAP70 protein tyrosine kinase attached to the cytosolic domain of the ζ chains. Second-generation CARs contained, in addition, the cytosolic signaling domains from one of the co-stimulatory receptors that form the 4-1BB protein or the CD28 protein (see Figure 25-24). Subsequent CARs incorporated both of these co-stimulatory domains and signaling domains from other proteins as well. These co-stimulatory domains activate multiple signaling pathways that trigger rapid and extensive activation of T-cell killing of target tumor cells. (b) Generation of CAR T cells. A cDNA encoding a chimeric antibody receptor is cloned into a lentiviral vector — similar to the retrovirus vector depicted in Figure 6-35 except derived from the human HIV retrovirus. Leukapheresis, a procedure to extract T cells from a patient with a B-cell malignancy, is performed, and the T cells are infected in culture with the lentivirus vector. After culture to allow integration of the viral DNA into the genome and expression of the chimeric antibody, the T cells are transfused back to the patient. The production of CAR T cells normally takes 17–21 days.

Complications of Cancer Immunotherapy

One caveat to all of these immune-activating treatments — both monoclonal antibodies against immune checkpoint molecules and CAR T cells — is that they unleash the full force of the immune system not only against the tumor, but often against the adjacent normal tissues that are not cancerous. Consequently, a major side effect of these therapies is a *cytokine storm* that leaves the patient with potentially devastating tissue damage. Successful application of cancer immunotherapy will require finding a balance between activation of the T cells to recognize tumors with other treatments that are less injurious of normal tissues.

KEY CONCEPTS OF SECTION 25.4

Evasion of Programmed Cell Death and Immune Surveillance Processes

- Precancerous cells often have sustained DNA damage and are normally eliminated by programmed cell death or apoptosis.
- Oncogenic driver mutations that suppress apoptosis allow continued tumor cell proliferation even in the presence of sustained DNA damage.
- The most common driver oncogenic driver mutations are in p53, a transcription factor that activates the expression of genes for cell cycle arrest or apoptosis. Cancer cells that have loss of function mutations in p53 continue to proliferate damage to their genomes and thus have a propensity to accumulate mutations.
- Immune cells in the tumor microenvironment detect mutant proteins and shape the developing tumor through a three-step process involving tumor cell elimination, equilibrium between the tumor cells and the immune system, and late escape of tumor cells from immune control.
- Cytotoxic T cells recognize fragments of mutant proteins presented by MHC proteins on the surface of tumor cells. However, the killing activity of the T cells is usually abated in part because of activation of inhibitory receptors on the T-cell surface, such as CTLA4 and PD-1.
- Monoclonal antibodies to CTLA4 and PD-1 and to PD-L1, the ligand of PD-1, prevent activation of these inhibitory receptors and enhance the ability of the T cells to kill the tumor cells. These monoclonals are used to treat cancers such as melanomas in which many mutations have accumulated.
- CAR T cells are normal T cells in which lentivirus vectors have been used to introduce a chimeric protein, a CAR, containing the V_H and V_L domains of a monoclonal antibody targeting a protein on the surface of a tumor cell, a transmembrane segment, and cytosolic domains from the TCR ζ submit and from any of several co-stimulatory receptors. CAR T cells are efficacious in killing many types of human tumors and in curing the cancer.

End of Chapter

Achieve

Visit Achieve to access study tools and to learn more about the content in this chapter:

- Perspectives for the Future
- Classic Experiment 25.1: An Experiment That Led to the Identification of the RAS Oncogene
- Analyze the Data
- Additional study tools, including videos, animations, and quizzes

Key Terms

angiogenesis
benign
Burkitt's lymphoma
cancer stem cell
carcinogen
carcinoma
CAR T cells
epithelial-to-mesenchymal transition (EMT)
immune checkpoint pathways
immunoediting
leukemia

loss of heterozygosity (LOH)
lymphomas
malignant
metastasis
multi-hit model
mutagen
oncogene
oncogenesis
p53 protein
Philadelphia chromosome
proto-oncogene
Ras protein
Rb protein
retrovirus
sarcoma
telomerase
transformation
tumorigenesis
tumor microenvironment
tumor suppressor gene
Warburg effect

Review the Concepts

1. Despite differences in origin, cancer cells have several features in common that differentiate them from normal cells. Describe these.

2. What characteristics distinguish benign from malignant tumors?
3. Which important characteristic of tumor cells did Otto Warburg discover?
4. Because of oxygen and nutrient requirements, cells in a tissue must reside within 100 μm of a blood vessel. Based on this information, explain why many malignant tumors often possess gain-of-function mutations in one of the following genes: β*FGF*, *TGF*-α, and *VEGF*.
5. Ninety percent of cancer deaths are caused by metastatic rather than primary tumors. Define *metastasis.* Explain the rationale for the use of the cancer treatment agent batimastat, an inhibitor of matrix metalloproteases and of the plasminogen activator receptor.
6. What is the importance of the EMT during metastasis?
7. What hypothesis explains the observations that incidence of human cancers increases exponentially with age? Give an example of data that confirm the hypothesis.
8. Distinguish between proto-oncogenes and tumor suppressor genes. To become cancer promoting, do proto-oncogenes and tumor-suppressor genes undergo gain-of-function or loss-of-function mutations? Classify the following genes as proto-oncogenes or tumor-suppressor genes: *p53, ras, BCL-2, JUN, MDM2*, and *p16*.
9. Describe how mutations in genome maintenance factors promote tumorigenesis. Why would inactivation of a mismatch repair gene cause colon cancer?

10. Hereditary retinoblastoma generally affects children in both eyes, while spontaneous retinoblastoma usually occurs during adulthood only in one eye. Explain the genetic basis for the epidemiological distinction between these two forms of retinoblastoma. Explain the apparent paradox: loss-of-function mutations in tumor-suppressor genes act recessively, yet hereditary retinoblastoma is inherited as an autosomal dominant.
11. Explain the concept of loss of heterozygosity (LOH). Why do most cancer cells exhibit LOH of one or more genes? How does failure of the spindle assembly checkpoint lead to loss of heterozygosity?
12. Many malignant tumors are characterized by the activation of one or more growth-factor receptors. What is the catalytic activity associated with transmembrane growth-factor receptors such as the EGF receptor? Describe how a point mutation that converts a valine to glutamine within the transmembrane region of the HER2 receptor leads to activation of the relevant growth-factor receptor.
13. Describe the common signal transduction event that is perturbed by cancer-promoting mutations in the genes encoding RAS and NF1. Why are mutations in *RAS* more commonly found in cancers than mutations in *NF1*?
14. Describe the mutational event that produces the *MYC* oncogene in Burkitt's lymphoma. Why does the particular mechanism for generating oncogenic *MYC* result in a lymphoma rather than another type of cancer? Describe another mechanism for generating oncogenic *MYC*.

15. Pancreatic cancers often possess loss-of-function mutations in the gene that encodes the Smad4 protein. How does this mutation promote the loss of growth inhibition and highly metastatic phenotype seen in pancreatic tumors?
16. Why are mutations in the p16-arf-p15 locus so dangerous?
17. Explain how epigenetic changes can contribute to tumorigenesis.
18. Several strains of human papilloma virus (HPV) can cause cervical cancer. How does the viral E7 protein contribute to host cell transformation?
19. Loss of p53 function occurs in the majority of human tumors. Name two ways in which loss of p53 function contributes to a malignant phenotype. Explain how the carcinogen benzo(*a*)pyrene can cause loss of p53 function.
20. The active form of p53 is a tetramer of four identical subunits. However, a mutant version of just one subunit can inactivate the entire tetramer. Explain how this dominant negative interaction increases the probability that mutations in p53 are selected for during cancer progression.
21. Explain why only about 1 percent of mutations in a tumor cell have an effect on a cellular protein and thus may cause production of a neo-antigen.
22. CAR T cells are genetically engineered to kill cancer cells by introduction of a chimeric protein composed of the antigen-binding domain of a monoclonal antibody linked to the cytosolic ζ subunit of the T-cell receptor. Explain the

function in targeting and killing cancer cells of the antibody part and the T-cell receptor part of the chimeric protein.

Index

Note: Key terms are in bold, figures are denoted by "f" following the page number, and tables are denoted by "t" following the page number.

A

A kinase-associated protein (AKAP), 690f
A site, 213
A tubule, 799
AAA ATPase family, 602, 815
ABC proteins, 487–491, 489t, 490f, 492f
ABC superfamily, 481, 481f
ABO blood-group antigens, 460–461, 461f, 461t
Acceptor, 713
Acceptor stem, 208
Acetic acid, titration curve, 59, 59f
Acetyl CoA, 464, 532
    oxidation of, 533–534, 533f
    pyruvate conversion to, 531–532, 533f
    structure of, 532f
Acetyl lysine, 48f
Acetylation, 818
Acetylcholine, 1065, 1066, 1069, 1074–1075
Acetylcholine-gated cation channels, 1074–1075, 1074f

Acetylcholinesterase, 1073
Acetyl-coenzyme A (acetyl CoA), 532. *See also* **Acetyl CoA**
Acid hydrolases, 19
Acidic activation domains, 349–350
Acidic amino acids, 47
Acidic keratins, 839
Acids, 58, 58f
Acquired immune deficiency syndrome (AIDS), 223
Actin, 755
Actin cross-linking proteins, 770
Actin filaments, 752f
    assembly of, 763–769
        Arp2/3 complex, 764–766, 765f, 766f, 768f
        formins, 763–764, 764f
        polymerization, 766–767, 767f
        toxins and, 769
    binding proteins, 758t
    capping proteins, 762–763, 762f
    in cell locomotion, 789f
    dynamics of, 758–763
    growth of, 760–761, 760f
    hippo kinase cascade pathway, regulation of, 984–985, 986f
    myosin head steps along, 777–779, 778f
    myosin V carries vesicles along, 785, 786f
    in nonmuscle cells, 783–784
    organization of, 790–791
    phagocytosis, 769f
    polymerization, 758–759, 759f

signal-induced changes to, 792f
structures of
F-actin, polarity of, 757–758, 758f
F-actin polymers, 756–757, 757f
G-actin monomers, 756–757, 757f
overview of, 755–756
treadmilling, 761–762, 761f
Actin nucleation, 765, 765f
Actin-activated ATPase activity, 775
Actin-based cellular structures
adapter proteins, 770–773, 772f
cross-linking proteins, 770, 771f
organization of, 770–773
Actin-binding proteins, 762, 782–783, 783f
Actin-nucleating proteins, 763
Action potentials, 664, 782, 998, 1045, 1045f
definition of, 1044
inputs for, 1076, 1076f
magnitude of, 1052
myelinated axons, 1060–1061, 1061f
propagation of, 1052–1064, 1052f
propagation unidirectionally without diminution, 1055, 1056f
in reflexes, 1047f
sequential opening and closing of voltage-gated channels, 1052–1055, 1053f
Activated bonds, 209
Activation (A) loop, 112
Activation domain, 345, 349–350, 350f, 359–361

Activation energy, 63–64, 63f, 97, 97f
Activation loop, 111, 710, 714f
Activation-induced deaminase (AID), 1116, 1116f
Activator, 326f, 327, 345, 346f, 358
Activator (Ac) elements, 292
Active asymmetric kinase domain dimer, 712–713, 714f
Active sites, 97–99
    in DNA transcription, 201
    serine and, 99–102
    of trypsin, 98f
Active transport, 473
Active zone, 1066
Acute inflammatory demyelinating polyneuropathy, 1062
Acyl group, 52
Acylation, 459
ADAM family, 738, 739f, 740
Adaptation, 670
Adapter protein (AP) complexes, 631t, 642–643
Adapter proteins, 717–718, 770–773, 772f, 905, 909f
Adaptive immune system
    collaboration of immune-system cells
        antigen-presenting cells, activation of, 1143
        cancer, 1146–1147
        high-affinity antibodies, 1143–1145, 1144f
        toll-like receptors, 1140–1143, 1141f
        vaccines, 1145–1146
    definition of, 1093
    inflammation and, 1100–1102

innate immune response, interplay with, 1101f
in layers of immune defenses, 1094f
lymph nodes, initiation in, 1097f
secreted Ig and, 1115–1116
as third line of defense, 1102–1103
in tissues and lymph nodes, 1096–1098
Adenine, 48, 49, 180, 180f
Adenosine, 48f
Adenosine diphosphate (ADP), 7, 8f
Adenosine monophosphate (AMP), 517, 517f
Adenosine triphosphate (ATP), 7, 8f, 512
generation by oxidation of fatty acids, 534–535, 536f
glycolysis, rate of, 517–518
hydrolysis of, 64–65, 64f, 585–586, 636
light absorption, 561
in photosynthesis and respiration, 65–66
synthesis by ATP synthase, 550f
synthesis by chemiosmosis, 551f
synthesis by proton-motive force, 550–558
ATP-ADP exchange and phosphate transport, 557, 557f
binding-change mechanism, 553f
F_1 γ subunit, rotation of, 553–554, 554f
mechanism of, 550–551
multiple protons and, 554–555
synthesis of, 515, 561, 569
Adenovirus, 221, 222f
Adenylyl cyclase, 683
activation or inhibition of, 683–691, 683f

anchoring proteins, 688–689
cAMP activates PKA, 683–685, 685f, 686f
cAMP-mediated activation of PKA, responses to, 687–688, 687f
feedback mechanisms, 689–691
gene transcription, activation of, 688, 689f
hormone-induced, 684f
PKA, hormone-induced activation, 685–687, 685f, 686f
receptor-ligand complexes, 683
signal amplification, 687
Adherens belt, 755, 783
Adherens junctions, 773, 912f, 913–918, 913t, 916f–917f
Adhesion junctions, 986
Adhesion molecules, 911–926
Adhesion receptor, 902, 904f
Adhesion receptor-mediated signaling pathways, 909f
Adrenaline, 663
Adsorption, 221
Adult stem cells, 143–144. *See also* Stem cells
Adult subventricular zone (SVZ), 1049
Adult-onset diabetes. *See* Type II diabetes
Aequorin, 675
Aerobic, 66
Aerobic oxidation (aerobic respiration), 513, 513f, 515
Afferent neurons, 1046
Affinity, 43, 96, 116, 665–667, 670–672, 671f
Affinity chromatography, 121, 672–673
Affinity maturation, 1113

Aggrecan, 938
Aggrecanases, 942
Agonist, 672
Agrin, 1067
AIRE (autoimmune regulator), 1135
Akt, 725
Alanine, 46
Alcoholic fermentation, 519f
Algae, 785, 787f
Aliphatic amino acids, 46
Alkylation, 338
Alleles, 228, 229, 229f, 232f, 259–260
Allelic exclusion, 1114
Allogeneic tissue, 1134
Allosteric effector, 107
Allosteric proteins, 107
Allosteric-binding site, 107
Allostery, 57, 107
All-*trans*-retinal, 699, 699f, 701, 1063, 1063f
A-loop, 111, 112–113, 112f, 113f
Alpha (α) helix, 72f, 74, 74f, 347, 347f, 348f
    in transmembrane proteins, 455–458, 456f, 457f
Alpha carbon (C_α) atoms, 45
Alphaproteobacterium, 523–524
α subunit, 679–681, 680f, 681f
α1-antirypsin, 601
α-actinin, 770
α-SNAP, 636

αβ-tubulin dimers, 797–799, 799f, 802f, 804
Alphoid DNA, 319
Alport's syndrome, 931
Alprenolol, 672f
Alternating access model, 473
Alternative pathway, 1099, 1100f
Alternative polyadenylation sites, 399, 421
Alternative RNA splicing. *See* Alternative splicing
Alternative splice site, 280, 402
Alternative splicing, 204, 204f, 387, 400, 402f, 404, 404f
Alu elements, 297f
Alzheimer's disease, 95
Amanita phalloides, 769
Amino acids, 7, 45–48
    common, 46f
    common modifications, 48f
    essential, 47
    linear arrangement of, 72–73, 72f
    mTORC1 kinase activation, 966–968, 967f
    Na^+-linked symporters and, 501–503
    peptide bonds, 72–73, 73f
    phosphorylated, 673–674, 674f
    sequences of, 88–89
    transcellular transport, 507–508
Aminoacyl-tRNA, 207, 210, 213–216
Aminoacyl-tRNA synthetases, 207, 208f, 209–210, 213
AMPA receptors, 1089f, 1090
Amphipathic, 34, 52, 444

Amphiphilic, 34
Amphitelic attachment, 881, 882f
Amplification, 669, 1164. *See also* Signal amplification
of DNA sequences, 242–245, 245f, 246f, 316–318
Amyloid fibrils, 94f, 95
Amyloid state, 94–95
Amyloidoses, 95
Amyloids, 94–95, 94f
Amylopectin, 559
Amylose, 559
Amyotrophic lateral sclerosis (ALS), 998
Anaerobic respiration, 515, 516
Anaphase, 827f, 828
Anaphase A, 828, 833, 834f
Anaphase B, 828, 833
Anaphase-promoting complex (cyclosome; APC/C), 828, 864, 864f, 869f, 884–886, 886f
Anchoring collagens, 930t, 932–933
Anchoring junction, 911, 912–913, 913t
Anchoring proteins, 688–689
Anemias, 772
Aneuploid cells, 140, 229
Aneuploidy, 826, 893, 898f, 1152
Anfinsen, Christian, 88
Angiogenesis, 1155
Angiomotin, 985, 987f
Angulin, 921
Animals

cloning of, 996–997
models for cancer development studies, 1174–1176, 1176f
Anion, 39
Anion antiporter, 504–505
Anion Exchanger 1 (AE1), 505, 506f
Aniridia, 31f, 327
Ankyrin, 771
Annular phospholipids, 458, 458f
Annulus, 954, 954f
Antagonist, 672
Antagonistic interactions, 1021
Antenna complexes, 561
Anterograde, 625
Anterograde transport, 639–641, 640f, 641f, 810–811, 811f, 812f, 817f
Antibiotic resistance, 299–300
Antibiotics, 211–212
Antibody, 1095. *See also* Immunoglobulins; Immunology
cancer immunotherapy, 1147
definition of, 96, 1093, 1102
diversity of, 1109–1116
fluorescence-tagged, 142f
high-affinity, 1143–1145, 1144f
ligand-binding of, 96, 96f
monoclonal (*See* Monoclonal antibodies)
organelles, isolation of, 174
somatic hypermutation, 1113
Antibody assays, 122
Antibody probes, 356f

Antibody staining, 332f
Antibody-affinity chromatography, 120f, 121
Antibody-dependent cell-mediated cytotoxicity (ADCC), 1108
Anticodons, 205, 208–209, 208f, 218
Antidiuretic hormone, 479
Anti-fibronectin antibody treatment, 908f
Antigen processing and presentation, 1123
Antigen-presenting cell (APC), 1098, 1117, 1143
Antigens, 96, 674, 1095
    cytosolic, 1123–1125, 1124f
    definition of, 1094
    presentation by B cells, 1127f
    presentation of by phagocytic cells, 1126f
    protein fragments and, 1123
    t cells and, 1181–1182
Antigen-specific receptors, 1132–1134, 1133f
Anti-mitogens, 869
Antiparallel dimer, 934
Antiparallel strands, 180–181
Anti-pausing factor, 340
Anti-peptide antibodies, 1143
Antiporters, 473, 500–507
    anion, 504–505, 505f, 506f
    cation, 504
    Cl^-/HCO_3^-, 504
    H^+-linked, 1068–1070
    Na^+/H^+, 504
    $Na^+HCO_3^-/Cl^-$, 504

Na^+-linked Ca^{2+}, 504
proton/sucrose, 506
three Na^+/one-Ca^{2+}, 504
Antisense inhibition, 421
Antisera, 1103
Antiserum, 1103
Antitermination complex, 340f
Antithrombin III (ATIII), 936, 937, 937f
APC gene, 1174, 1175f
APC/C^{Cdc20}, 884–886, 886f, 892–893
Apical, 507, 911
Apical Par complex, 1022–1023, 1023f
Apical regions, 648–649, 648f
Apical surface, 27, 911–912, 911f, 912f
Apicomplexa, 27
Aplysia californica, 430f, 1087, 1088f
Apolipoproteins, 650
Apoptosis. *See also* Programmed cell death
Bax and Bak proteins, 1035–1037, 1036f
definition of, 31, 1027
evasion of, 1177–1178
extrinsic pathway, 1038f, 1039
p53 protein, 1177–1178
phosphatidylserine and, 1032, 1033f
programmed cell death by, 1028–1029, 1028f
proteasomes and, 104
regulation of, role of mitochondria, 1034–1035
in vertebrates, 1037–1039

Apoptosome, 1037f
Apoptotic pathway, 1029–1031, 1030f
Aquaporin, 457, 478–480, 479f
Aquifex aeolicus, 503f
Arabidopsis thaliana, 325f, 524, 560, 1013–1014, 1014f
Archaea, 2f, 14, 331
ARF protein, 631, 632–633
Arginine, 47, 100f, 974–975
Arginyl-rTNA-protein-transferase (ATE), 974, 975f
Argonaute protein, 420, 422
Arms, 850
Arp2/3 complex, 763, 764–766, 765f, 766f, 768f, 791
Arrestin, 690, 701, 702–703, 702f
Arrhythmia, 781
Artificial membranes, 477–478
Ascorbic acid, 934
Ash 1 protein, 428, 429f, 430f
Asparagine, 47
Aspartic acid, 47
Assays, 121, 122
Assembly, 221
Aster, 826
Astral microtubules, 828, 829f
Astrocyte intermediate precursor cells (aIPCs), 1050f
Astrocyte syncytia, 1048
Astrocytes, 1047–1048, 1049f, 1065, 1066f, 1077
Asymmetric carbon atom, 36–37
Asymmetric cell division, 32, 990, 1015–1026

cell polarization before cell division, 1017–1018
general features of, 1018f
in neuroblasts, 1026f
par proteins and, 1020–1022, 1020f, 1021f, 1025–1026, 1025f
in yeast, 1018–1019
Asymmetric kinase dimer, 713, 714f
Ataxia telangiectasia and Rad3-related protein (ATR), 889–890, 890f
Ataxia telangiectasia mutated (ATM), 889–890, 890f, 892
Atherosclerosis, 43, 411, 466
ATM (ataxia telangiectasia mutated), 889–890, 890f, 892
Atoms, covalent bonds and, 36–37
ATP synthase, 173, 481, 514, 551–553, 552f, 554–555, 556f
ATP synthesis, 514, 514f
ATP/ADP exchanger, 557
ATP-AMP ratio, 968–969, 968f
ATPase fold, 756
ATP-powered pumps, 473, 480–484, 481f
ATR (ataxia telangiectasia and Rad3-related protein), 889–890, 890f
Atractylis gummifera, 557
ATRIP protein, 890
Attenuation, 1145
AUG codon, 212–213
Augmin complex, 800
AU-rich elements, 419
Aurora B, 832, 833f
Aurora kinase, 875, 876, 879–882, 888
Autism spectrum disorder, 406–407, 406f

Autocrine signaling, 663
Autoimmunity, 1094, 1109
Autonomously replicating sequences (ARSs), 318, 319f
Autophagic pathway, 657–659, 658f
Autophagic vesicle, 657
Autophagosome, 655, 657, 658–659
Autophagy, 19, 104, 529, 655, 657, 966
Autoradiography, 124–126, 625–626
Autosomal dominant diseases, 254, 254t, 255f
Autosomal dominant polycystic kidney disease (ADPKD), 825
Autosomal recessive diseases, 254t
Autosome, 313
Auxins, 953
Avian leukosis virus, 1161
Avidity, 1105
Avidity model of T-cell selection, 1135
Axon guidance, 1045
Axon terminal, 1043f, 1071f
Axon termini, 1045
Axonal transport, 810
Axonemal dynein, 820, 822f
Axoneme, 820
Axons
    anterograde transport of vesicles, 810–811, 811f
    definition of, 527
    information flows through neurons, 1044–1045
    myelinated, 1060–1061, 1061f
    in neurons, 798f, 800

transport, rate of, 810f
Axoplasm, 810

B

B cells, 1127f
antigens, presentation of, 1127f
antigens, recognition of, 1095
development, 1109–1116, 1114f
development of, 1113–1115, 1114f, 1135f
high-affinity antibodies, 1143–1145, 1144f
lymphocytes, 1096f
naive, 146, 1106–1107
proliferation and differentiation of, 1132–1134, 1133f
BAC (bacterial artificial chromosomes), 240
Backward slippage, 286, 287f
Bacteria
circadian rhythms in, 981
ribosomes, structure of, 212f
transposable DNA elements, 290–291, 291f
use of in research, 14
Bacteriophage, 219, 221
Bag of marbles (Bam) protein, 1004
Bak proteins, 1035–1037, 1036f
Balbiani rings, 413, 415f
Ball-and-chain inactivation model, 1059f
Ball-and-stick model, 78, 78f
Band 3 anion transporter, 505

Banding patterns, on chromosomes, 313, 317f
Band-shift assays, 343, 344f
Barbed end, 757
Bardet-Biedl syndrome, 825
Barth's syndrome, 547
Basal, 911
Basal body, 798, 800, 800f, 820, 821f
Basal lamina, 27, 28f, 142, 142f, 912, 926–932
    functions of, 927–928, 927f
    laminin, role of, 928, 929f
    major protein components of, 928f
    sheet-forming type IV collagen, 928–931, 931f
Basal layer, 839
Basal state, 666
Basal surface, 27, 911–912, 911f, 912f
Base excision repairs, 192, 192f
Base pairs, 180, 181, 182f, 208–209
Basement membrane, 928, 931, 1156
Bases, 8, 48, 49f, 58
Basic amino acids, 47
Basic helix-loop-helix (bHLH), 77, 349
Basic keratins, 839
Basic zipper (bZIP), 349
Basolateral, 507, 912
Basolateral regions, 648–649, 648f
Basolateral surface, 27
Bax proteins, 1035–1037, 1036f
B-cell receptor (BCR), 1115

Bcl-2 family, 1034, 1035f
BCR-ABL protein kinase, 1162–1163, 1162f
"Beads-on-a-string" arrangement of nucleosomes, 301, 302f
Beige-fat cells, 558
Bence-Jones proteins, 1106
Benign, 1155
Benign familial neonatal convulsions, 1055
Benign tumor, 1155
Bent closed conformation, 944, 945f
Benzo(a)pyrene, 1158, 1158f
Beta (β) barrel, 75
Beta (β) sheet, 72f, 74–75, 75f, 94f
Beta (β) strands, 74, 75f, 94f
Beta (β) turn, 75–76, 75f
Beta-blockers, 672
β islet cells, 960–961, 961f, 998–999, 1000f, 1002
β_1-adrenergic receptor, 672
β_2-adrenergic receptor, 672, 678f
β2-microglobulin, 1117
β-adrenergic receptors, 676
β-arrestin, 691f
β-catenin, 741, 742, 742f, 744
β-galactosidase, 327, 328, 328f, 329f
β-interferon enhancer, 352, 352f
β-like globin gene family, 282–284
β-oxidation, 542
β-propeller domain, 653
β-thalassemia, 1011

BH3 domain, 1031
BH3-only protein, 1031, 1037–1039
bHLH proteins, 350
bHLH transcription factor, 971–972
Bidirectional growth, 189–190, 189f
Bidirectional synapses, 1076
Bile acids, 449
Bile salts, 462
Binding, proteins and, 71, 72f
Binding assays, 670–671, 671f
Binding specificity, 665, 718
Binding-change mechanism, 553
Biochemical energetics
    ATP generation during photosynthesis and respiration, 65–66
    forms of energy, 60–61
    free energy, 61–63
    hydrolysis of ATP, 64–65
    NAD^+ and FAD couple biological reaction, 66–67
    reaction rate and activation energy, 63–64
    transformation of energy, 61
Bioinformatics, 5, 248, 249–250
Biological complexity, 252
Biological fluids, pH values of, 57–58, 57f
Biological processes, studying in culture, 146
Biology, 1
Biomarkers, 133
Biomembranes, 51–54
    lipids

classes of, 447–449, 448f
composition of, 450–453, 452f
excess, storage of, 454, 454f
lateral mobility, 449–450, 451f
major components, 450t
microdomains, cholesterol and sphingolipid clusters in, 453–454
phospholipids, sphingolipids, and cholesterol, 464–468
cytosolic proteins and, 464–465
enzymes and, 464
ER membrane, 465, 466f
flippases and, 465–466
movement, 464–465
synthesis, 464, 466–467, 466f, 467f
transport of, 467–468, 468f
structure of, 442–468, 443f
lipid bilayer, 444–454, 444f
variation in, 447f
Biomolecular condensates, 21, 84–86, 85f
Biomolecules, common functional groups, 38t
Bi-orientation, 831, 832, 881, 898
Biosynthetic pathways, 234, 235f
BiP chaperone, 598, 600, 600f
Bitter taste, 1081
Bivalent bonding, 1103
Bivalents, 895
Black Death, 109–110
Blackburn, Elizabeth, 318
BLAST (basic local alignment search tool), 250

Blastocoel, 993
Blastocyst, 272, 993, 997
Blastopore, 30
Blastula, 854
Blindness, 699
Blood, 902
Blood cells, 1009, 1010f
Blood glucose level, regulation of
    insulin and glucagon, 960
    insulin inhibits glucose synthesis, 963–964
    intracellular vesicles, fusion of, 961–963, 962f
    β islet cells, insulin secretion from, 960–961, 961f
Blood vessels, tumor growth and, 1155
Blood-brain barrier, 1048, 1049f
Bloom's syndrome, 1159t
Blue native (BN)-PAGE, 545, 546f
Blunt-ended DNA, 243f
B-lymphocytes, 144–145
Bonds, high-energy, 65
Bone marrow, 1010f, 1012f
Bone marrow transplantation, 1009–1011, 1011f, 1012f
Bone morphogenetic proteins (BMP), 733, 737
Botulinum toxin, 1070
Botulism, 1070–1071
Boundary element, 308, 311
Boundary membrane, 522
Boveri, Theodor, 1152
Brain, 1042f, 1060f

Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, 1043
Brain stem, 1043, 1043f
Branchlets, 1049f
Branch-point A, 392, 395f, 407f
BRCA1 gene, 195, 891, 1159
BRCA2 gene, 195, 1159
Breast cancer, 146, 195, 672, 1153f, 1159, 1159t, 1176–1177
Breast Cancer Resistance Protein, 490
Brenner, Sydney, 31
Bright-field light microscopy, 148f–149f, 150, 151f
Bromodomains, 308, 358
Brown-fat tissue, 558
Buffering capacity, 59
Buffers, 58–60, 59f
Bundle sheath cells, 573
Burkitt's lymphoma, 1171, 1172f
bZIP proteins, 350

C

C. elegans, 28, 31, 274, 275f, 715, 1020f, 1021f
    apoptotic pathway, 1029–1032, 1029f
    CED-3 protease, 1031f
    cell asymmetry in, 1020
    cell death, 1029f, 1031f
    cell lineage in, 1020f
    germ-line stem cells, 1004–1005

C lobe, 713, 714f
C_2H_2 zinc finger, 348–349
C_3 pathway, 570
C_4 pathway, 573, 573f
C_4 zinc finger, 348–349, 349f
C_a backbone trace, 78–79, 78f
Ca^{2+} ATPases, 482–483, 483f, 485f
Ca^{2+} ions
    concentration of, and feedback loops, 696, 697f
    cytosolic elevation of, 694f
    DAG activation of protein kinase C, 697
    glycogenolysis, regulation of, 697, 697f
    membrane lipid phosphatidylinositol 4,5-biphosphate,
    hydrolysis of, 693–694, 693f
    neurotransmitters, release of, 1070
    release from ER triggered by IP_3, 694–695, 695f
    rise in, cellular responses to, 692t
    as second messengers, 692–697
    skeletal muscle contraction, regulation of, 782–783, 783f
    store-operated channel in plasma membrane, 695–696, 696f
Ca^{2+} pump, mechanism of action, 483–484
CAD domains, 696
Cadherin, 904, 904f, 913–918, 914f, 917f
Caenorhabditis elegans, 902, 991
Cajal bodies, 438, 439f
Calcium, 108, 109f, 1072f
Calcium signaling, 530
Callose, 955

Calmodulin, 108, 109f, 165f, 483, 686
Calnexin, 598, 599, 601f, 692
Calorie, 61
Calreticulin, 598, 599, 601f, 692
Calvin cycle, 570–573
Cambrian explosion, 520
cAMP
    activates protein kinase A, 683–685, 685f, 687–689
    cellular responses to rise in, 688t
    extracellular signaling, amplification of, 687f
    glycogen metabolism, regulation of, 686f
    in glycogenolysis, 697, 697f
    levels of, 690f
cAMP-dependent protein kinase, 684
cAMP-response element (CRE), 688
Campylobacter jejuni, 491, 492f
Cancer
    aerobic glycolysis, 1153–1154, 1154f
    age and, 1173f
    breast, 146, 195, 672, 1153f, 1159, 1159t, 1176–1177
    causative mutations, 260
    cervical, 224
    colon, 1152f, 1159t, 1173–1174, 1175f
    common mutations, 1166t
    diagnosis, 1176–1177
    evolutionary progression of, 1150f
    genetic and genomic basis of, 1157–1168
        carcinogens, 1157–1158

epigenetic changes, 1167–1168
familial syndromes, 1158–1159, 1159t
genome sequencing, 1160, 1161f
inherited predispositions, 1163–1164
microRNAs and, 1166–1167
oncogenes, 1160–1161
oncogenic drivers, 1162–1165, 1165f
somatic mutations, 1159–1160
tumor suppressor genes, 1165–1166
immune system, activation of, 1181–1183
immune system and, 1146–1147
immunotherapy, 1147
lung, 1158
melanomas, 193, 1181
mTOR pathway mutations, 969
overview of, 1149–1151
programmed cell death, evasion of
immune system and, 1179
oncogenic driver mutations, 1177–1178
p53 protein, 1178–1179
tumor microenvironment and immunoediting, 1179–1181
skin, 193
somatic mutations, 1146–1147
targeted therapy, 1162–1163
TGF-β signaling and, 737
treatments for, 1176–1177, 1181–1183
tumor cells vs. normal cells, 1151–1156
tumorigenesis, 1168–1177

animal models, 1174–1176, 1176f
colon cancers, 1173–1174, 1175f
growth control pathways, 1170–1171
molecular cell biology, 1176–1177
multi-hit model, experimental reconstruction of, 1172–1173
nuclear transcription factors, 1171–1172
receptor mutations, 1168–1169, 1169f
signaling pathways, aberrations in, 1172
signal-transducing proteins, 1169–1170
types of, 1149
Cancer stem cell, 1154
Capacitors, 494
Capping proteins, 762–763, 762f
Capsid, 219, 220–221
Capsomeres, 220f
Capture, 805
CapZ protein, 762, 764, 765, 782f
CAR T cells, 1181, 1182, 1182f
Carbohydrates, 49
Carbon, pathway in photosynthesis, 571f
Carbon atoms, covalent bonds and, 36–37, 36f
Carbon dioxide
from acetyl CoA, 533–534, 533f
transport by erythrocytes, 504–505, 505f
Carbon fixation, 66, 561, 562f, 570–572, 572f
Carbonic anhydrase, 504
Carboxyl-terminal domain (CTD), 332–333, 332f, 387–389, 388f
Carcinogen, 1157–1158

Carcinoma, 1149
Cardiac hypertrophy, 740
Cardiac muscle contractions, 504
Cardiolipin, 467, 547
Cargo molecules, 629
Cargo proteins, 623, 625, 633
Carnitine, 535
Carotenoids, 563
Carriers, 254. *See also* Transporters
Cartilage, 938f
Cas genes, 269, 271f, 272, 273f
Caspases, 1031–1032, 1037
Catabolism, 66, 516
Catalase, 20, 536
Catalysis, 71, 72f
Catalysts, 55, 64, 97
Catalytic loop, 113f
Catalytic RNA, 436–437, 436f
Catalytic sites, 97, 98f
Cataracts, 925
Catastrophe, 803, 804f
Catecholamines, 1069f
Cathepsins, 1128
Cation, 39
Cation antiporter, 504
CD3 complex, 458f
$CD4^+$ T lymphocytes, 219
CD40 ligand, 1138

Cdc (cell division cycle) mutants, 232, 233–234, 233f, 236, 241, 242f, 853
Cdc14 phosphatase, 886
Cdc20, 884
Cdc25, 863
Cdc25B phosphatase, 876
Cdc25C phosphatase, 876
CDC28, 867
Cdc42 protein, 790–792, 791f, 793f, 843–844, 844f, 1016, 1017f
Cdc45, 873
CDK2, 861f
CDK-activating kinase (CAK), 862
cDNA library, 242, 243f
Ced-3 gene, 1029f, 1030–1031, 1030f, 1031f
CED-3 protease, 1031f
Cell adhesion, 788–790
    in motile and nonmotile cells, 943–951
        igCAMS, 949
        integrin-mediated adhesion, regulation of, 944–948, 946f–947f
        integrins and, 943–944, 943f, 945f
        leukocyte movement into tissues, 949–951, 950f
        muscular dystrophy, 948–949, 948f
    plant vs. animal, 955–956
Cell body, 1044
Cell cortex, 755, 789f
Cell culture, 855
Cell cycle, 21, 22f, 825. *See also* Eukaryotic cell cycle

Cell death, 1027–1039, 1027f
apoptosis, 1028–1029, 1028f
apoptosis in vertebrates, 1037–1039
apoptotic pathway, 1029–1031, 1030f
in *C. elegans*, 1029f, 1031f
caspases and, 1031–1032, 1037
environmental stresses, 1037–1039
evolutionarily conserved proteins, 1029–1031, 1030f
extrinsic apoptosis and necroptosis pathways, 1038f, 1039
mitochondria, in apoptosis regulation, 1034–1035
neurotrophins and, 1032–1034, 1034f
overview of, 991f, 992
phosphatidylserine, 1032, 1033f
pro-apoptotic proteins, 1035–1037
trophic factors, 1037
Cell differentiation
cleavage of embryo, 992–993
embryonic stem cells, 995f
ES and iPS cells and human cells, 998–1000
multipotent somatic stem cells, 1003–1004, 1003f
patterns of, 1004f
in planarians, 1002–1003, 1002f
reversal of epigenetic changes, 996–997
Cell division, 30, 30f. *See also* Asymmetric cell division
Cell energetics, 512–573
ATP and NADPH, 570–573
ATP synthesis, 514, 514f
chemiosmosis, 514

chloroplasts and photosynthesis, 559–566
core antenna complexes, 564
DNA in chloroplasts, 560
light absorption, 561
light-harvesting complexes, 564, 565f
overview of, 562f
reactive oxygen species, 564–566
stage 4, 570–573
stages 1 and 2, 561–564
thylakoid membranes, 559–560, 561
citric acid cycle and fatty acid oxidation, 531–537, 532f
ATP generation by mitochondrial oxidation of fatty acids, 534–535, 536f
conversion of pyruvate to acetyl CoA, 531–532
fatty acids, peroxisomal oxidation of, 536, 536f
high-energy electrons, generation of, 533–534, 533f
NAD^+ and NADH concentrations, 534
net result of, 538t
contact sites of mitochondria and ER membrane, 527–531
mitochondrial function and dynamics, 529–531, 530f
electron transport, 514
electron-transport chain, 537–549
energy release and electron carriers, 538–540, 538f
multiprotein complexes, 540–544
NADH and $FADH_2$, oxidation of, 537
overview of, 541f
proton pumping and, 537–538
proton-motive force, 548–549

reactive oxygen species, 547–548, 547f
reduction potentials favor electron flow from NADH to O_2, 544–545, 545f
stoichiometry of proton pumping, 548
supercomplexes, 545–547, 546f
glycolysis, 515–520, 516f
anaerobic, 518–520, 519f
net result of, 538t
rate of, 517–518, 517f
stage I, 516–517
light energy to generate oxygen, NADPH, and ATP, 566–569
mitochondria, structure of, 520–526, 521f
evolution of, 523–524, 524f
membranes, 520–523
mtDNA, 524–526, 525f
proton-motive force, 514, 514f, 550–558
ATP synthesis, mechanism of, 550–551
ATP-ADP exchange and phosphate transport, 557, 557f
brown fat and heat generation, 558
F_0 c ring rotation, 555
F_1 γ subunit, rotation of, 553–554, 554f
mitochondrial oxidation, rate of, 558
multiple protons and, 554–555
multiprotein complexes, 551–553, 552f
Cell growth signals
integrating with nutrient and energy levels
active mTORC1 complex, 965–966, 965f
autophagy, inhibiting, 966

glycolysis, stimulating, 966
mTORC1 kinase activation, 966–969, 967f
ribosomal RNAs and tRNAs, synthesis of, 966
Cell junction, 902, 905, 912, 912f, 913t
Cell line, 140
Cell lineage, 990
Cell migration, 787–793
actin-based structures involved in, 789f
chemotactic molecules, 793, 793f
force generation coordinated with cell adhesion and membrane recycling, 788–790
small GTP-binding proteins, 790–791, 791f, 793f
steps in, 788f
step 1: membrane extension, 788
step 2: cell-substratum adhesions, 788
step 3: cell-body translocation, 788–789
step 4a: breaking cell attachments, 789–790
step 4b: recycling membrane and integrins, 790
wounded-cell monolayer assay, 792f
Cell murder, 1038f
Cell polarity, 753–754, 1015–1026
cell polarization before cell division, 1017–1018
general features of, 1018f
intrinsic polarity program, 1016, 1017f
par proteins and, 1022–1023, 1023f
planar cell polarity pathway, 1023–1025, 1024f
positive feedback loops involving Cdc42, 1016, 1017f
in yeast, 1018–1019

Cell proliferation, 959
Cell senescence, 140
Cell signaling, 665f, 755f. *See also specific types*
Cell strain, 140
Cell surface mannose receptor, 1098
Cell surface membrane, 664–665
Cell wall, 952, 952f
Cell-adhesion molecules (CAMs), 27, 140, 902, 903f, 904–905, 904f
    cell-cell adhesions and, 904, 905f
    evolution of, 908–909
    mechanotransduction and, 909–910, 910f
Cell-binding domain, 939f
Cell-cell adhesion, 903–910, 903f
    cell-adhesion molecules (CAMs), 904–905, 904f
    definition of, 902
    epithelial cells, surfaces of, 911–912
    igCAMS, 949
    matrix junctions and adhesion molecules, 911–926
        cadherins in adherens junctions and desmosomes, 913–918, 914f
        epithelial cells, surfaces of, 911f
        gap junctions, 923–925, 924f
        junctions, type of, 912–913, 912f
        tight junctions, 920–923, 920f, 921f, 922f
        tunneling nanotubes, 925, 926f
    model for generation of, 905f
Cell-cell adhesion receptors, 902, 903f, 904
Cell-free assays, 628–629, 629f, 644f

Cell-matrix adhesion, 903–910, 903f
basal lamina, 926–932
connective tissue, 932–942
definition of, 902
epithelial cells, surfaces of, 911f
matrix junctions and adhesion molecules, 911–926
epithelial cells, surfaces of, 911–912
gap junctions, 923–925, 924f
integrins in, 918–919, 919t
junctions, type of, 912–913, 912f
tight junctions, 920–923, 920f, 921f, 922f
tunneling nanotubes, 925, 926f
Cell-matrix adhesion receptor, 902
Cells
biological processes, studying, 146
birth, lineage and death of, 991f
in culture, 139–148
for animal cells, 139–148
biological processes, studying, 146
drugs used in research, 146–147
flow cytometry, 140–142
hybridomas produce monoclonal antibodies, 144–146
life span of, 140
stem cells, 143–144
two-dimensional and three-dimensional, 142–143
drugs used in research, 146–147, 147f
electron microscopy (*See* Electron microscopy)
flow cytometry, 140–142

integrating into tissues, 901–956
adhesive interactions in motile and nonmotile cells, 943–951
basal lamina, 926–932
cell-cell and cell-EM adhesion, 903–910, 903f
cell-cell and cell-EM junctions and adhesion molecules, 911–926
overview of, 901–903, 903f
plant tissues, 951–956
light microscopy (*See* Light microscopy)
mitochondria
functions of, 522t
structure of, 526
organelles, isolation of, 172–175
primary, 140
transformed, 140
variety of, 4f
Cell-surface receptors, 664, 665f
perlecan and, 931–932
study of, 670–675
binding assays, 670–671, 671f
cellular response to signaling molecule, 671
external signals, sensitivity to, 671–672
purification of, 672–673
signaling molecules, chemical analogs of, 672
Western blotting, 673–674
types of, 707f
Cellular communication, 661

Cellular differentiation, 359, 366–368, 372
Cellular environment, 958–987
blood glucose level, regulation of, 960–964
cell growth signals, integrating with nutrient and energy levels, 964–969
cholesterol and unsaturated fatty acids, changes in levels of, 970–972
circadian rhythms, 978–982, 980f–981f
low oxygen, responding to, 972–975, 973f
overview of, 958–960
physical environment, sensing and responding to, 982–987
temperature, elevated, 976–978, 977f
Cellular membranes, 446f
Cellular oxidative stress, 547
Cellular respiration, 66
Cellulose, 50, 953
Cellulose synthase, 953
Central dogma (of molecular biology), 178
Central nervous system, 1043, 1043f, 1047, 1048–1051, 1061–1062, 1061f, 1065, 1068. *See also* Nervous system
Central stalk, 811
Centralspindlin, 836, 887f, 888
Centrifugation, 116–117, 117f, 173–174, 173f
Centriole, 798, 800, 879
Centromere, 288f, 311f, 318, 354–355, 828f, 829, 850
Centromere sequences, 318–319, 319f, 320f
Centromeric heterochromatin, 287
Centrosome (spindle pole body), 800, 800f, 801f, 824f, 879

duplication and separation, 879, 880f
duplication of, 825–826, 826f
mitotic spindles and, 850
Centrosome cycle, 879, 880f
Centrosome disjunction, 826, 879
Centrosome maturation, 879
Cerebellum, 1043, 1043f
Cerebrum, 1043, 1043f, 1044
Cervical cancer, 224
CFTR correctors, 492
CFTR gene, 33, 254, 255f, 638
CFTR potentiator, 492–493
cGMP phosphodiesterase (PDE), 700
Chain elongation, 213–216, 215f, 388–389, 388f
Chance, Britton, 545
Channel-inactivating segment, 1054
Channelopathies, 1055
Channelrhodopsin, 162, 1063, 1063f, 1064f
Channels, 472f, 473
Chaperones, 90, 599f, 692
chaperonins, 92–94, 93f
Hsp70, 90–92
Hsp90, 92
molecular, 90
protein folding in vivo, 89–94, 91f
proteins, folding and assembly of, 598–600, 599f
Chaperonins, 90, 92–94, 93f
Charcot-Marie-Tooth disease, 925, 1062

Chargaff, Erwin, 180
Checkpoint pathway, 889, 892–893, 893f
Checkpoints, 849, 890f, 891f
Chemical adducts, 193
Chemical bond energy, 35, 35f
Chemical building blocks, 35, 35f
Chemical equilibrium, 35, 35f, 54–60, 55f, 56f
Chemical libraries, 146, 147f
Chemical potential energy, 60–61
Chemical reactions, 54–60
    buffers and, 58–60, 59f
    dissociation constants, 56–57
    equilibrium, 55
    extent of, 55
    hydrogen ions in, 58
    pH values of biological fluids, 57–58, 57f
    rates of forward and reverse reactions, 55, 55f
    steady state in cells, 55–56, 56f
chemical building blocks of cells, 45f
    amino acids, 45–48
    nucleic acids, 48–49
    phospholipids, 51–54
    polysaccharides, 49–51
Chemical signals, 1044, 1082f
Chemical synapses, 1045, 1046f, 1076
Chemiluminescence, 122, 123f
Chemiosmosis, 514
Chemiosmotic coupling, 514

Chemiosmotic hypothesis, 550
Chemistry
    covalent bonds, 36–44
    foundations of, 34–69
    key concepts, 35f
    noncovalent interactions, 36–44
Chemokine, 1101, 1139
Chemotactic signals, 1100
Chemotaxis, 793
Chiasmata, 895, 897f
Chicken, web-footed, 1027f
Chimeric antibodies, 1182f
Chimeric proteins, 139, 605, 606f
Chiral carbon atoms, 37
Chirality, 37
Chironomus tentans, 413, 415f
Chk2, 892
Chlamydia, 14
Chlamydomonas reinhardtii, 25, 820, 822f, 823, 1063
Chloride channels, 509–510, 509f
Chlorophyll *a*, 561, 563f, 567
Chlorophylls, 561
Chloroplasts, 21, 513
    DNA in, 560
    functions of, 21
    light absorption, 561
    membranes and, 446–447, 446f
    photosynthesis and, 559–566

thylakoid membranes, 559–560
structure of, 560f
targeting proteins to, 609–610, 610f
targeting sequences, 578t
Cholesterol, 43, 449
bHLH transcription factor, 971–972
in biomembranes, 449, 452, 452f
biosynthesis, 970
changes in levels of, 970–972
endoplasmic reticulum SCAP protein, 970–971, 971f
in membrane microdomains, 453–454
PCSK9 gene and, 259–260
synthesis and intracellular movement, 464–468
synthesis of in cytosol and ER membrane, 466–467, 467f
transport of, 467–468, 468f
Cholesteryl esters, 53
Choline, 52
Chromatid, 311f, 313, 850
Chromatin, 278, 300, 354
condensation and function, 304–311
condensed, 326f
as disordered chain of nucleosomes, 303–304
nucleosomes and, 302f
regulators, 996
structure of, 300–303, 304f
Chromatin immunoprecipitation, 336f, 338, 371, 374f
Chromatin-mediated repression, 354
Chromatin-remodeling complexes, 358–359, 359f, 1167

Chromatography, 119–121
Chromatophores, 512
ChromEMT, 301–302, 303, 304f
Chromodomain, 308
Chromogenic enzyme reactions, 121
Chromoshadow domain, 308
Chromosomal passenger complex (CPC), 832, 833f, 835, 835f
Chromosomal translocations, 315f
Chromosome conformation capture methods, 309–311, 310f
Chromosome congression, 831–832, 831f
Chromosome homologs, 850f
Chromosome paint probes, 308
Chromosome painting, 313–316, 315f
Chromosome territories, 278f, 308–311
Chromosomes, 16, 17, 17f, 257f
    alignment of, 831–832
    attachments, stable and unstable, 882f
    duplication of during S phase, 850f
    evolution of, 313–316
    metaphase structure, 311, 313
    morphology and functional elements in Eukaryotes, 313–322
        centromere sequences, 318–319
        chromosome painting, 313–316
        interphase, 316–318
        in metaphase, 313
        replication and stable inheritance of, 318
        telomeric sequences, 319–321
    primate, evolution of, 316f

in prometaphase, 829–831, 831f
segregation of, 882–883, 883f, 884, 885f, 897–899, 897f
structure of, 278f
Chronic lymphoblastic leukemia (CLL), 1177
Chronic myelogenous leukemia (CML), 1162
Chronic progressive external ophthalmoplegia, 526
Chronobiology, 978
Chymotrypsin, 100f
Cilia, 15
Cilium, 447, 820–825
beating, production of, 820
image of, 798f
intraflagellar transport, 820–823, 822f, 823f
primary, 823–825, 824f
structural organization of, 821f
Circadian clock (circadian oscillator), 959
Circadian rhythms, 978–982, 980f–981f
in bacteria, 981
negative feedback loops, 979–981
suprachiasmatic nucleus, 981–982, 982f
Circulating tumor cells, 1156
Circulatory system, 1096–1098, 1096f
Circumferential belt, 783–784
Cis (lateral) interactions, 905
Cis activation, 379
Cis configuration, 87, 88, 88f
Cis double bonds, 53
cis-Golgi cisterna, 625

cis-Golgi network, 625, 626, 627, 628, 629f, 637f, 640–641, 640f
Cisterna, 625, 639
Cisternal maturation, 624f, 625, 639–641, 640f, 641f
Cistron, 280
Citric acid, 533
Citric acid cycle, 515, 531–537, 532f
ATP generation by mitochondrial oxidation of fatty acids, 534–535, 536f
conversion of pyruvate to acetyl CoA, 531–532
fatty acids, peroxisomal oxidation of, 536, 536f
high-energy electrons, generation of, 533–534, 533f
NAD^+ and NADH concentrations, 534
net result of, 538t
overview of, 533f
stage II, 531–532
stage III, 531
stage IV, 531
CKIs (cyclin kinase inhibitors), 861
Clamp domain, 331–332, 332f
Clamp loader, 188, 188f
CLASPs, 807, 809f
Class switching, 1116
Classical cadherin, 905f, 914–918, 916f–917f
Classical pathway, 1099, 1100f
Class-switch recombination, 1116f
Clathrin, 174, 175f, 630, 631t, 642–644, 643f, 644f
Clathrin-coated pits, 651
Clathrin-coated vesicles, 630, 635, 644, 644f

Claudin, 920
Claudin16 gene, 922
Clausin14 gene, 923
CLAVATA3 (CLV3) gene, 1014, 1014f
Cleavage, 992–993
Cleavage and polyadenylation specificity factor (CPSF), 423, 423f
Cleavage furrow, 835
Client proteins, 90, 91f, 92. *See also* Substrate
CLIP segment, 1122, 1128
CLN3, 867, 868
C-lobe, 111, 112f
Clonal selection theory, 1106–1107, 1106f
Clone, 138, 239. *See also* DNA cloning
Cloned DNA, 239, 245–246
Cloning, 996–997
Clonogenic neoblasts, 1002
Closed complex, 338f
Closed loops, 1047
Clotting factor IX gene, 298
Cl^-/HCO_3^- antiporter, 504
Cluster analysis, 263, 263f
Clustered protocadherin, 914, 915
cNeoblasts, 1002
Co-activator, 349–350, 350f, 358, 362f, 688, 689f
Coated vesicles, 631–633, 631t, 632f
Coatomers, 638
Coding region, 279
Coding sequences, 298

Codons, 178, 206, 206t, 208–209, 208f
Coenzyme A (CoA), 464, 532, 532f
Coenzyme Q (CoQ), 540, 540f
Coenzymes, 66, 67f, 102
Cofactor, 102
Cofilin, 761–762
Cognate antigens, 1103
Cognate proteins, 343
Cognates, 209
Cohen, Stanley, 711
Cohesinopathies, 875
Cohesins, 303, 825, 850, 873, 897–898, 897f
Coiled coil, 74, 76, 76f
Co-immunoprecipitation (co-IP), 123f, 124
Coincidence detection, 766
Colchicine, 146, 805, 806f, 823
Collagen, 47, 906t, 927
    in connective tissue, 941f
    definition of, 932
    sheet-forming type IV, 928–931, 931f
    triple helix, 929, 930f, 931f
    type of, 930t
    type I, 935
    type I and type II, 934–935
    type II, 935
    type IV, 928, 928f, 931
    type IX, 935
    type VI, 934, 934f

Collagen α chains, 929, 930f, 931, 933f
Collagenases, 942
Colon cancers, 1152f, 1159t, 1173–1174, 1175f
Colony, 140
Colony-forming units–erythroid (CFU-E), 728f
Colorless native (CN)-PAGE, 545
Combinatorial diversity, 918
Commensal bacteria, 1095
Compensating mutations, 392
Competitive inhibitors, 102
Complement, 1098
Complement system, 1098–1100, 1100f
Complementarity-determining regions (CDRs), 96, 96f
Complementary, 181
Complementary DNA (cDNA), 241–242, 263–264
Complementary matching, 8
Complementary RNA, 200–202
Complementary-determining regions (CDRs), 1106, 1107f
Complementation analysis, 233–234, 234f
Complex traits, 258
Complex transcription units, 280–281, 282f
Complexin, 1072
Concentration gradient, 61
Condensates, 345
Condensin, 303, 311, 882
Condensin complex, 883f
Condensin I, 883
Condensin II, 883

Conditional mutations, 232, 233f
Cones, 698, 1077
Confocal microscopy, 155–156, 157f, 158f
Conformations, 70, 78–79, 129–130
Conjugating plasmids, 299
Connective tissue, 902, 906, 907f, 932–942
    elastic fibers, 941–942, 941f
    fibrillar collagens, 932–934, 933f, 934f
    fibronectins, 938–941, 939f, 940f
    hyaluronan, 937–938
    metalloproteases, 942
    proteoglycans, 935–937, 935f, 938f
    type I and type II collagens, 934–935
Connexin, 923–925
Connexon, 924f, 925
Conservation, 248
Constant region, 1108
Constitutive, 346
Constitutive secretion, 625
Constitutive secretory vesicles, 646
Constitutive splice sites, 402
Contact inhibition, 1154
Contact sites, 605
Contractile bundles, 783
Contractile ring, 755, 784, 835, 835f, 887
Contractile vacuole, 478
Convergent evolution, 1013
Coomassie blue dye, 122

Cooperative DNA binding, 351–352, 351f
Cooperativity, 107, 108f
Co-oriented, 898
COPI, 630, 631t, 632, 632f, 640f
COPII, 630, 631f, 631t, 632–633, 632f, 637–638, 637f, 640f
Copy-and-paste mechanism, 293
$CoQH_2$-cytochrome *c* reductase (Complex III), 540, 542–543, 543f
Core antenna complexes, 561, 564
Core element, 380
Core gate, 106
Core photosystems, 561
Core PIC, 338
Core promoter elements, 334f
Core promoter initiation complex, 338
Co-repressor, 350, 357, 358
Cortical granules, 992
Co-Smad, 735, 737
Co-transcription, 387
Cotranslational translocation, 581–582, 581f
Cotransport, 500–507
Cotransporter, 473, 474
Coupled transport, 500
Covalent bonds
    atoms, electronic structure of, 36–37, 36f
    electronegativity, 37–39
    noncovalent interactions and, 36–44
    strength of, 39, 39f
Cowpox virus, 1146

CoxII gene, 524
Coxsackievirus and adenovirus receptor (CAR), 921
CoxVa, 606
CPE-binding protein (CPEB), 423, 423f
CpG island, 334–336
Crambin, 41f
CREB protein, 349–350, 350f
CRE-binding (CREB) protein, 688, 689f
Crick, Francis H. C., 8, 180, 289
CRISPR, 269–272
    eukaryotic genomes, 270
    genome editing, 268
    genome-wide, 270
    inducible activation, 270–272
    pharming, 560
    single-nucleotide mutations, 271f
CRISPR interference, 273f
CRISPR-Csas9, 273f
Crista junctions, 521f, 522, 556f
Cristae, 20, 556f
Critical cell size, 868
Critical concentration (C_c), 759, 760f, 802–803
Critical concentration for the (–) end, 760
Critical concentration for the (+) end, 760
Critical micelle concentration (CMC), 462, 463f
Crohn's disease, 1154
Cross-bridge cycle, 780
Cross-exon recognition complex, 401

Crossing over, 195, 895
Cross-linking proteins, 770, 771f
Cross-presentation, 1123, 1125
Cross-β sheet, 95
Cryoelecron microscopy, 130, 131f, 169–170
Cryoelectron tomography, 170, 171f
Cryptic splice sites, 405
Crypts, 1005
C-terminal domain, 80, 331f
C-terminus, 73
Cubitis interruptus (Ci), 745
Cues, 1016
Culture
    for animal cells, 139–140
    biological processes, studying, 146
    definition of, 138
    in vivo environment, emulating, 142–143
    life span of, 140
    stems cells and, 143–144
    two-dimensional and three-dimensional, 142–143
Culture medium, 139–140
Cut-and-paste mechanism, 291–292, 292f
Cyan fluorescent protein (CFP), 160, 161f, 162, 162f
Cyanide, 544
Cyanobacteria, photosystems in, 565f
Cyclic AMP, 687f
Cyclic electron flow, 568–569, 568f
Cyclic guanosine monophosphate (cGMP), 700

Cyclic photophosphorylation, 568
Cyclin box, 861
Cyclin D, 849
Cyclin D1 gene, 1170
Cyclin E, 859
Cyclin fold, 861
Cyclin-dependent kinases (CDKs), 190
    cell cycle progression, regulation of, 858f
    centrosomes and, 825, 826f
    cyclins determine activity of, 857, 859–861, 859f
    inhibitors of, 863
    kinases and, 849
    as master controllers, 848
    regulation by activating and inhibitory phosphorylation, 861–863, 863f
    regulatory cyclin subunit and, 857–858, 857f
Cyclins, 107, 859–861
    CDK, activity of, 859f
    cyclin D, 849
    discovery of, 860f
    levels of, 863–865, 864f
Cyclooxygenases, 102
Cyclosporine, 1134
Cysteine, 47
Cysteine dioxygenase, 974, 975f
Cystic fibrosis, 254t, 255f, 638
Cystic fibrosis transmembrane regulator (CFTR), 491–493, 493f
Cytochalasin D, 769

Cytochrome *c*, 174
Cytochrome *c* oxidase (Complex IV), 540, 544
Cytochrome *bf* complex (Cyto *bf*), 562f, 568, 568f
Cytochrome P-450 enzymes, 1157, 1158f
Cytochromes, 539
Cytogenetic maps, 257f
Cytokeratins, 797f, 839
Cytokine receptors, 717, 724, 727–732
    signal transduction pathways, 706f–707f
    signaling suppression, 730, 732f
    structure and activation of, 729f
Cytokine storm, 1183
Cytokines, 385, 417, 419, 727–732, 1013, 1100
    binding of to receptors, 728–730
    definition of, 705
    extracellular growth factors, 1013
    interferons, 727–728
    receptor proteins, 706–707
    regulation by, 727–728
    signal transduction, 732f
    signaling suppression, 730
    T cells and, 1138
Cytokinesis, 755, 784, 784f, 827f, 828, 835–836
    in eukaryotic cell cycle, 850–852, 887–888, 887f
Cytolytic T lymphocyte (CTL), 1118
Cytoplasm, 10, 12f
Cytoplasmic dynein, 814, 817f
Cytoplasmic interitance, 525

Cytoplasmic mechanisms, 417–430
cytoplasmic polyadenylation and, 422–423
microRNAs and, 419–421
mRNA, concentration of, 417
mRNA, degradation of, 417–422, 418f
mRNA, localization of, 428–430, 430f
protein synthesis and, 423–424
RNA interference (RNAi), 421–422
sequence-specific RNA-binding proteins and, 424–426
surveillance mechanisms, 426–428
Cytoplasmic mRNPs, 387
Cytoplasmic polyadenylation, 422–423, 423f
Cytoplasmic polyadenylation element (CPE), 423
Cytoplasmic sleeve, 955
Cytoplasmic streaming, 785, 787f
Cytosine, 48, 49, 180, 180f
Cytoskeleton, 284, 442, 754
components of, 754f
coordination and cooperation between elements, 843–844
definition of, 7
of epithelial cell, 753f
functions of, 15
of migrating cell, 753f
muscular dystrophy and, 948f
regulation of by cell signaling, 755f
signal-induced changes to, 792f
types of filaments, 15f
Cytosol, 466–467, 467f, 601–602, 664–665

Cytosolic domain, 455
Cytosolic enzymes, 516–517
Cytosolic face, 446, 447f, 473f, 495
Cytosolic leaflets, 452–453
Cytosolic materials, degradation of, 654–659
Cytosolic pH, regulation of, 471, 504, 508–509, 509f
Cytosolic proteins, 464–465, 617f, 657–659
Cytosolic side, 540
Cytosolic stress granules, 21
Cytotoxic T cells, 1117–1118, 1118f, 1119f, 1120, 1137, 1137f, 1145f, 1179–1181

D

D gene segments, 1132
D. melanogaster, 28
DAG. *See* 1,2-diaglycerol
Dalton, 73
Danio rerio, 32
Dark adaptation, 701, 701f
Dark reactions, 561
Darwin, Charles, 1, 3f, 956
Darwinian evolution, 1157
Daughter cells, 32, 178, 230f, 991f, 992
Daughter strands, 186–187, 187f
De novo DNA methyl transferases, 372–373
Deadenylation-dependent pathway, 418, 418f
Deadenylation-independent decapping pathway, 419

Deafness, 925
Deamination, 191, 191f
Death domain (DD), 1039
Death proteins, 1039
Death receptors, 1039
Deconvolution microscopy, 155–156, 156f
Decoration experiments, 757
Dedifferentiation, 1009
Defensins, 1102
Degenerate codons, 206
Degradation, of proteins, 104–107
Dehydration reaction, 44, 45f
Delayed rectifier K^+ channels, 1054
Delayed response genes, 869
Deletion, 1164
Demyelinating disease, 1062
Denaturation, 89, 182–184, 183f
Dendrites, 800, 844, 1043f, 1044–1045, 1093f
Dendritic cells, 1101
Dendritic spines, 1087f
Dense bodies, 840
Density, centrifugation and, 116–117, 117f
Density gradient, 116, 135f
Deoxynucleotides (dNTPs), 244
Deoxyribonucleic acid (DNA), 2, 177
    carcinogens, 1157–1158
    chemical directionality, 180, 180f
    in chloroplasts, 560

damage to, 12
damage response system, 889–892, 890f, 891f
double-helical structure of, 179–185, 181f, 182f
complementary antiparallel strands, 180–182
denaturation, 182–184, 183f
torsional stress, 184, 184f
familial syndromes, 1158–1159, 1159t
model of, 9f
in Neanderthals, 5, 6f
overview of, 8–11
repair and recombination, 190–199
base excision, 192
double-strand breaks, 193–194, 196f, 198f
excision-repair systems, 191–192
homologous recombination, 194–198
mismatch excision, 192–193
mutations, 191
nucleotide excision, 193
replication, 179f, 185–190
bidirectionality, 189–190, 189f
CDK inhibitors, degradation of, 869–871, 869f, 870f
daughter strands, 186–187
initiation of, 871–873, 872f
proteins, cooperation of, 187–189
templates and primers, 185–186
sequence, 9
somatic mutations, 1159–1160
synthesis of, 185–187, 186f, 187f

transcription, 179f, 199–205
alternative RNA splicing, 204
functional mRNAs, 202–204
stages of, 201–202
template DNA to complement RNA, 200–202
translation, 179f
Deoxyribonucleoside triphosphate monomers (dNTPSs), 179f
Deoxyribonucleotide triphosphates (dNTPs), 246, 247f
Deoxyribose, 48f
Dephosphorylation, 110, 110f, 886, 887f
Depolarization, 1045, 1045f
Depolarized state, 700
Depurination, 192
Dermal tissue, 952
Desensitization, 672, 737
Desensitization response, 714
Desmin, 840
Desmocollin, 918
Desmoglein, 918
Desmoplakin, 918, 918f
Desmosomal cadherin, 914, 918
Desmosome, 912f, 913–918, 913t, 918f
Desmotubule, 955
Destruction complex, 741
Destruction proteins, 1029
Detergents, 461–463, 462f, 463f
Detyrosylated microtubules, 818, 819f
Deubiquitinases (Dubs), 107

Deubiquitinylation, 113–114
Deuterostomes, 30
D-Glucose, 50
Diabetes
    treatments for, 998–999, 1000f
    type I, 964, 998–999
    type II, 259, 960, 964, 998–999
Diabetes insipidus, 480
Diabetes mellitus, 960, 964
Di-acidic sorting signal, 638
Diacylglycerol, 447
Diapedesis, 951
Dicer, 272, 274, 420, 421–422, 421f
Dichroic mirror, 152
Dictyostelium, 793
Differential centrifugation, 116, 173, 173f
Differential protein expression, 324
Differential-interference-contrast (DIC) microscopy, 150–151, 151f, 811f
Dihydrofolate reductase (DHFR), 605, 606f
Dimerization, 709–711, 712–713, 712f, 714f
Dimerization arm, 711, 712f
Dimethylallyl pyrophosphate (DMPP), 466–467
Dioscorides, 557
Diploid cells, 229, 229f, 231f
Dipole, 38, 38f
Dipole moment (μ), 38
Direct isolation, of DNA sequences, 244–245, 245f, 247f

Direct repeats, 291, 291f, 292f, 293f, 369
Direct-acting carcinogens, 1157
Disaccharides, 50, 51f
Diseases
    abnormal RNA splicing, 405–408, 406f, 407f
    amyloids and, 94–95, 94f
    cloned DNA and, 257–258
    common inherited, 254t
    genetic defects and, 258
    human linkage studies and, 256–258
    iPS cells in treatments, 998, 999f
    lysosomal storage diseases, 645
    mitochondrial DNA mutations, 526
    mobile DNA elements and, 298
    monogenic, 253–255, 255f
    primary cilium, defects in, 825
    protection against, 259–260
    protective alleles, 259–260
    protein folding and, 94–95
Dishevelled genes, 1025
Disintegrin, 738
Dislocation, 601, 602
Disruption construct, 268
Dissociation (Ds) elements, 292
Dissociation constant (K_d), 56–57, 99, 666
Distant enhancers, 341–342
Distant poles, 832
Disulfide bond, 47, 77

Disulfide bonds (—S—S—), 597–598, 598f
Diversity, 1094
DNA chips, 262
DNA clone, 239
DNA cloning, 237–248, 240f
cDNA libraries, 241–242
disease and, 257–258
DNA fragments, insertion of, 238
E. coli plasmid vectors, 239–240
gene expression and, 260–268
cDNA, sequencing, 263–264
cluster analysis, 263
E. coli expression systems, 264–265
plasmid expression vectors, 265–267
in situ hybridization, 261
polymerase chain reaction, 242–245
sequencing, 245–246
yeast genomic libraries, 240–241
DNA fingerprint, 277, 287–288, 288f
DNA library, 240
DNA ligase, 187, 238, 239f
DNA looping, 353f
DNA methylation, 372–373, 1167
DNA microarray, 261–263, 262f
DNA polymerase, 177, 178, 185–186, 186f, 190
DNA polymerase I, 185
DNA polymerase α, 187, 188f, 189f
DNA polymerase β, 192f

DNA polymerase δ, 187–188, 188f, 189, 189f
DNA polymerase ε, 188, 188f, 189f
DNA polymerase η, 193
DNA polymorphisms, 255–257, 258, 288, 1120
DNA probes, 356f
DNA replication, 230f, 231, 287f, 320f
DNA sequencing, 248–253, 257f
    biological complexity, 252
    comparisons across species, 250–252, 251f, 252f
    DNA cloning and characterization, 237–248
    evolution of chromosomes, 313–316
    gene identification, 249
    genes, functions and evolutionary origins of, 250
    methods of, 246f, 247f
    mutations, functional consequences of, 249–250
    polymerase chain reaction, 242–246
DNA transposons, 289, 290–292, 290f, 292f, 298, 299f
DNA-binding domain, 345, 347–349, 348f, 371f
DNA-PK (DNA-dependent protein kinase), 889–890, 890f
DNase, 306, 306f
DNase I footprinting, 343–345, 344f
DNase I hypersensitive sites, 366–368, 367f
Dobbs, Arthur, 956
Dobzhansky, Thoedosius, 1
Dolichol phosphate, 595
Domain architecture, 79
Domains, 77, 79–80, 80f
Dominant, 229, 229f, 230–231, 231f

Dominant-active protein, 790
Dominant-negative mutations, 229
Dominant-negative protein, 790
Donor, 713
Dopamine, 1073
Dorsal root ganglion cells, 1078
Dosage compensation, 308–309
Double bonds, fatty acids and, 54f
Double helix, 8, 9f, 179–185
Double mutants, 234–235, 235f, 236
Double-label fluorescence microscopy, 153, 155f
Double-positive cells, 1135
Double-sex genes, 404
Double-strand break, 889, 890f
Double-stranded RNA (dsRNA), 274, 275f
Doublet microtubules, 799, 799f
Down syndrome, 898
Down-regulation, 654
Downstream, 200–201, 200f, 327
Downstream promoter element (DPE), 334
DPC (deleted in pancreatic cancer), 737
Dravet's syndrome, 1055
Driver mutations, 260, 1160
Drosha, 420
Drosophila melanogaster, 31, 31f, 232–233, 316–318, 317f, 323
    cell cycles in, 850
    circadian clock in, 979, 980f–981f
    genetic studies of, 745

germarium, 1005f
germ-line stem cells, 1004–1005, 1005f
hedgehog signaling, 706, 745, 746–747, 746f
hippo kinase cascade pathway, 983–984, 983f
olfactory receptor neurons, 1083–1085, 1085f
retinal neurons in, 404–405, 405f
sexual differentiation, 403–404, 403f
toll-like receptors, 1140
Ubx gene mutation, 325f
voltage-gated ion channels, 1055
Wnt signaling, 741
Drugs
chemical analogs of signaling molecules, 672, 672f
export of by ABC proteins, 487–490, 490f
tubulin polymerization, drugs affecting, 805
Dscam isoforms, expression of, 404–405, 405f
Dual-specificity kinases, 82, 720
Dual-specificity phosphatases (DUSPs), 722
Duchenne muscular dystrophy (DMD), 32–33, 254t, 255, 255f, 298
Duplicated genes, 282
Duplication, 1164
Duchenne Muscular Dystrophy (DMD), 405
Duty radio, 779
Dynactin, 816
Dynamic instability, 802–805, 803f
Dynamin, 527, 633, 643–644, 644f, 1073
Dynamitin, 816, 816f
Dynein, 754, 809–819

in anaphase B, 833
cilia and flagella, 820
domain structure of, 815f
kinesins, cooperation with, 817f
LIS1 protein, 817f
organelle transport, 814–816, 815f, 816f
power stroke of, 815f
regulation of, 817f
Dynein/dynactin-dependent pathway, 834, 835f
Dystroglycan, 948–949
Dystrophic epidermolysis bullosa, 928
Dystrophin, 298, 773, 948
Dystrophin glycoprotein complex (DGC), 948f, 949

E

E site, 213
E2F transcription factor (E2F), 867
E3-ligases, 864
Early endosomes, 634–635
Early gene expression, 221
Early response genes, 720, 869
Edman degradation, 128–129
EEA1 (early endosome antigen 1), 635
EF hand, 77
Effective stroke, 822f
Effector, 889
Effector neurons, 1046

Effector protein, 662, 664, 665f, 681, 682t, 790, 958
Effector specificity, 665
Efferent neurons, 1046
EGF domain, 298
EGF receptor heterodimers, 712
EGF receptor homodimers, 711
EGF receptor, ligand binding to, 712–713, 714f
Eggs, 854, 990
eIF2, 424
eIF2B, 424
eIF4E-binding proteins (4E-BPs), 965
80S initiation complex, 213
Elastases, 942
Elastic fibers, 941–942, 941f
Elastin, 932, 941
Electric potential, 61, 496f
Electric potential gradients, 487f
Electrical energy, 60
Electrical signals, 1044
Electrical synapses, 923, 1076–1077
Electrochemical gradient, 472, 472f
Electrochemical signaling, 1046
Electron carriers, 538–540, 538f
Electron density map, 129
Electron microscope grid, 167–168
Electron microscopy, 166–172, 167f, 169f
Electron shuttles, 534
Electron transfer flavoprotein (ETF), 542

Electron transfer flavoprotein:ubiquinone oxidoreductase (ETF:QO), 542
Electron transport, 514
Electronegativity, 37–39
Electrons, 36–37, 36f
Electron-transport chain, 514, 515
electron-carrying prosthetic groups, 539f, 539t
energy release and electron carriers, 538–540, 538f
multiprotein complexes, 540–544
$CoQH_2$-cytochrome *c* reductase (Complex III), 542–543, 543f
cytochrome *c* oxidase (Complex IV), 544
NADH-CoQ reductase (Complex I), 540–541, 542f
succinate-CoQ reductase (Complex II), 541–542, 542f
NADH and $FADH_2$, oxidation of, 537
overview of, 541f
proton pumping and, 537–538
proton-motive force, 548–549
reactive oxygen species, 547–548, 547f
reduction potentials favor electron flow from NADH to O_2, 544–545, 545f
stoichiometry of proton pumping, 548
supercomplexes, 545–547, 546f
Electrophoresis, 117–119
Electrophoretic mobility shift assay (EMSA), 343–345, 344f
Electroporation, 265
Electrospray (ES) ionization, 127f
Elongation, 323, 370–371
Elongation complex, 202

Elongation factors (EFs), 213–214, 339–340, 339f
Elongation phase, 759, 759f
Embryo, 261f, 854–855, 992–993, 993f, 994f
Embryogenesis, 985–986, 991
Embryoid bodies, 994, 995f
Embryonic stem (ES) cells, 140, 991, 993–994, 994f. *See also* Stem cells
    differentiated cell types, formation of, 995f, 998–1000, 1000f, 1001f
    overview of, 991–992
    pluripotency, factors controlling, 994–996, 996f
Emerson, Robert, 561
Emery-Dreifuss muscular dystrophy (EDMD), 842
Emphysema, 601
ENaC channels, 1082
ENCODE (Encyclopedia of DNA elements), 379
End-binding proteins, 807–809, 809f
Endergonic reactions, 61, 62f, 64
Endocrine signaling, 663
Endocytic pathway, 623, 624, 624f
Endocytic vesicles, 755
Endocytosis, 18, 623f, 767–768, 768f, 790, 1066. *See also* Receptor-mediated endocytosis
Endoglycosidase D, 627, 627f
End-on attachments, 881f
Endonucleolytic pathway, 419
Endoplasmic reticulum, 17–18, 18f, 576f, 578t
    cellular cholesterol, level of, 970–971, 971f

feedback loops in, 696, 697f
membrane proteins, classes of, 587f
membrane proteins, insertion into, 586–594
in nuclear envelope reassembly, 887f
synthesis of integral membrane proteins, 587–588
transport to Golgi complex, COPI vesicles and, 638–639
transport to Golgi complex, COPII vesicles and, 637–638
vesicle-mediated protein trafficking, 637f
Endoribonucleases, 433f
Endosomes, 18, 19f, 653–654, 653f
Endosymbiont/endosymbiont hypothesis, 20–21, 446, 524f, 550
Endosymbionts, 523
Endothelial cells, 912, 1049f
Endothelium, 27, 28f, 950f
Endothermic reactions, 62
End-product inhibition, 107
Energy, in biological systems, 60–61
Energy charge, 517
Energy coupling, 65
Engulfment proteins, 1029
Enhanceosome, 352f
Enhancers, 28, 279, 327–328, 329f, 333, 341–342, 352, 730
Entamoeba histolytica, 25
Enthalpy (*H*), 62
Entropy (*S*), 40, 42, 62
Entry, 221
Environment, sensing, 1077–1086
mechanoreceptors, 1078–1079, 1078f, 1079f

odor, 1083–1084, 1083f
olfactory receptor neurons, 1084–1086, 1085f
pain receptors, 1079
primary tastes, 1080–1083
Enzyme inhibitors, 102
Enzymes, 7
activation energy and, 97f
active sites, 97–99
as catalysts, 97
cholesterol synthesis and, 466–467, 467f
in common pathways, 103
definition of, 71
fatty acids and, 464
free energy reactions, 100f
multienzyme complexes, 103, 103f
reaction mechanism, 99f
Enzyme-substrate (ES) complex, 98, 99f, 100f, 101f, 112f
Eosin, 151
Epidermal growth factor (EGF), 654, 709, 711, 712f, 713f, 715, 790, 1156f
Epidermal growth factor (EGF) domain, 79, 80f, 672
Epidermolysis bullosa simplex, 840–841, 841f
Epigenetic code, 308
Epigenetic marks, 372
Epigenetic memory, 308
Epigenetic modifications, 29
Epigenetics, 372
animal cloning and, 996–997

regulation of transcription, 372–380
DNA methylation, 372–373
long noncoding RNAs, 377–379
methylation of specific histone lysines, 373–375
by polycomb and trithorax complexes, 375–377
tumorigenesis, 1167–1168
Epilepsies, 1055
Epimerases, 50
Epimers, 50
Epinephrine, 663, 672f, 683, 697
Epithelial cell, 753f, 911–912, 911f, 1022–1023, 1023f
Epithelial tissue, 902
Epithelial-to-mesenchymal transition (EMT), 915, 917, 917f, 1156
Epithelium, 16f, 142, 142f, 752, 753f, 902
cell junctions, 912f
origins of, 27
planar cell polarity pathway, 1023–1025, 1024f
stems cells and, 1005–1007, 1006f, 1007f
Epitope, 96, 145, 267, 1108
Epitope tagging, 122, 154, 267
Epstein-Barr virus, 146
Equilibrium constant (K_{eq}), 55, 63
Equilibrium density-gradient centrifugation, 117, 174, 174f
ER lumen, 580f, 597–598
ER membrane, 465, 466f
cholesterol synthesis, 466–467, 467f
mitochondrial contact sites, 527–531
secretory proteins and, 579–581

targeting proteins to and across, 579–586
pulse-chase experiments, 579–580
ERAD (ER-associated protein degradation), 601, 602, 972
Erasers, 110
ERF transcription factors, 975f
ERK kinase, 968–969, 968f
ERM (ezrin-radixin-moesin) family, 772f, 773
ERMES, 529
Ero1, 597
ERVs (endogenous retroviruses), 293, 295
Erythrocytes, 504–505, 728f
Erythropoietin (Epo), 674f, 727, 728f, 972–973, 1013
Erythropoietin receptor (EpoR), 728f, 732f
Escherichia coli, 12f
expression systems, 264–265
membranes and, 447
plasmid vectors, 238, 239, 239f, 240–241, 264f
porins and, 459, 459f
use of in research, 14
ESCRT proteins, 656
Essential amino acids, 47
Essential fatty acids, 52
Essential light chain, 775
Esterification, 49
Estrogen receptor response element (ERE), 369, 370f
Ethylene response factors (ERFs), 974, 975f
Ethylmethane sulfonate (EMS), 229
Eubacteria, 2f, 14

Euchromatin, 302f, 304f, 307f, 354
Eukaryotes, 2f
biomembranes of, 442, 443f
cell cycle, 21
cell cycle proteins, 22f
chloroplasts, 21
chromatin condensation and function, 304–311
chromatin structure, 300–304
chromosomes, morphology and functional elements, 313–322
CRISPR and, 270
cytoskeleton, 15
DNA, major classes of, 279t
endoplasmic reticulum, 17–18
endosomes, 18
evolution of, 251f
gene density in, 283f
gene structure and organization, 278–285
introns and mRNA encoding, 279–280
nonprotein-coding genes, 284, 285f
solitary vs. family, 282–284
tandemly repeated arrays, 284
transcription units, 280–282
genome comparisons, 252f
genomes of, 249–250
glyoxisomes, 19–20
Golgi complex, 18
lysosomes, 19
mitochondria, 20–21

nucleus, 15–17
organelle-like structures, 21
peroxisomes, 19–20
pre-mRNA processing
    3′ cleavage/polyadenylation, 398–399, 398f
    5′ cap added, 387–388, 388f
    chain elongation, 388–389, 388f
    overview of, 387–399, 387f
    regulation of, 400–411
    RNA splicing, 391–398, 391f, 392f, 393f, 394f, 395f, 396f
    RNA-binding domains and, 389–391, 390f
    spliceosomes, 393–398, 395f
    splicing cycle, 395–397, 396f
    transesterification, 391–393, 392f
protein degradation in, 104
proteins and, 73
protein-sorting pathways, 577f
in research, 23f
ribosomes, structure of, 212f
structure and function of, 14–22
transcription in, 11f, 324, 326f, 327–333
    carboxy-terminal domain, 332–333
    clamp domain, 331–332
    nuclear RNA polymerases, 329–331
    regulation of, 342–343
    regulatory elements, 327–329
transposable (mobile) DNA element, 291–292
unicellular, 22–27

alga *Chlamydomonas reinhardtii*, 25
parasites, 25–27
yeasts, 22–25
vacuoles, 19
Eukaryotic cell cycle, 21, 847–899
APC/C activates separase through securin ubiquitinylation, 884–886, 886f
chromosome segregation, 882–883, 883f, 884, 885f
cohesins, cleavage of, 884, 885f, 886f
cytokinesis creates two daughter cells, 887–888, 887f
DNA replication
CDK inhibitors, degradation of, 869–871, 869f, 870f
duplicated strands become linked, 873–875, 874f
initiation of, 871–873, 872f
fundamental processes of, 888f
G_1 controls entry into S phase, 849
G_2 phase, 849–850
G_2/M transition and mitosis, 875–883
centrosomes duplication and separation, 879, 880f
mitotic spindles, 879–882, 881f
nuclear envelope breakdown, 876–879, 878f
positive feedback loops, 876, 877f
meiosis, 894–899
chromosome segregation, 897–898, 897f
germ cell formation, 894
vs. mitosis, 894–895, 895f
overview of, 896f
sister kinetochores, co-orienting, 898–899

mitosis, exit from, 886
model organisms, 852–856
M-phase, 850–852
overview of, 847–852, 848f
progression and control, 856–866
CDK inhibitors, 863
CDKs, cyclins determine activity of, 859–861, 859f
CDKs, regulation by activating and inhibitory phosphorylation, 861–863, 863f
cyclin levels, regulation of, 863–865, 864f
cyclin-CDK activity, regulation of, 862t, 863
cyclin-dependent kinases, 857–858, 857f, 858f
feedback loops in, 865
new CDK substrates and functions, 865–866, 866f
phosphoserine/threonine-binding domains, 865
regulation, surveillance mechanisms in, 889–893
DNA damage response system, 889–892, 890f, 891f
spindle assembly checkpoint pathway, 892–893, 893f
S phase, 850f
study of, 852–856
biochemical characterization, 854–855, 854f
budding and fission yeasts, 852–853, 853f
cell cycle regulation in mammals, 855
tools, 855–856
transition from G_1 into S phase
in budding yeasts, 867
control of, 868f
cyclin-CDK complexes, 867–868

E2F transcription factor (E2F), 867–868
extracellular signaling, 868–869
in Metazoans, 867–868
Eukaryotic cells, 5, 12, 13f
Eukaryotic translation initiation factors (eIFs), 213
Evanescent wave, 158, 160f
Evolution, 1, 2f
of chromosomes, 313–316
of primate chromosomes, 316f
of protein families, 81, 81f
timeline of, 3t
transposable (mobile) DNA element and, 298–300
Evolutionary tree, 29f
Excision-repair systems, 191–192
Excitable cells, 1045
Excitatory postsynaptic compartments, 1068f
Excitatory receptor, 1076
Excitatory signals, 1044
Excitatory synapse, 1065
Exergonic reactions, 61, 62f, 64
Exocyst, 634
Exocytosis, 625
Exon junction complexes, 397
Exon recognition, 401f
Exon shuffling, 298, 298f, 299f, 409
Exon skipping, 405
Exonic splicing enhancers, 401, 401f, 404
Exonic splicing silencers, 404

Exon-junction complex (EJC), 426, 426f, 427f
Exons
alternative splicing, 204, 204f
definition of, 202
gene duplication and, 281f
shuffling, 298, 298f, 299f, 409
splicing of, 391, 392f
transcription units and, 280–282, 282f
Exonuclease activity, 186
Exoplasmic domain, 455
Exoplasmic face, 446, 446f, 447f, 495, 587
Exoplasmic leaflets, 452–453
Exoplasmic side, 540, 543f
Exoribonucleases, 433f
Exosome, 409–410, 409f
Exothermic reactions, 62
Exportin, 162, 163f
Exportin 1, 618
Exportin 5, 420
Expressed sequence tag (EST), 249
Extended closed conformation, 944, 945f
Extended open conformation, 944, 945f
Extensin, 953
External face, 446
External signal, 671f
Extracellular cues, 894
Extracellular domain (ECD), 709–711, 710f, 711f
Extracellular matrix (ECM), 27, 143, 143f, 663, 733–735, 902

adhesion, signaling, and other functions, 905–908
basal lamina, 926–932
functions of, 927–928, 927f
laminin, role of, 928, 929f
major protein components of, 928f
perlecan, 931–932
sheet-forming type IV collagen, 928–931, 931f
cell density in, 907f
connective tissue, 932–942
elastic fibers, 941–942, 941f
fibrillar collagens, 932–934, 933f, 934f
fibronectins, 938–941, 939f, 940f
hyaluronan, 937–938
metalloproteases, 942
proteoglycans, 935–937, 935f, 938f
type I and type II collagens, 934–935
defective skeletal development and, 908f
functions of, 907t
hippo kinase cascade pathway, regulation of, 984–985, 985f
muscular dystrophy and, 948–949, 948f
proteins, 906t
Extracellular signaling, 663f, 868–869
Extracellular signaling molecules, 661, 705f
Extravasation, 950, 950f, 1139, 1156
Extrinsic hormonal signals, 1004
Eyeless, 31, 31f
Eyepiece, 150
Eyespot, 1063

Ezrin, 772f

F

F factor, 240
F_0F_1 complex, 521f, 532f, 550
F_1 ATPase, 553
Fab domain, 768
Facilitated diffusion, 473
Facilitated transport, 473, 474–480
Facilitated transporters, 473
F-actin, 757–758, 758f, 771f
F-actin polymers, 756–757, 757f
Facultative anaerobes, 518
FAD (flavin adenine dinucleotide), 66–67, 66f, 67f, 102
$FADH_2$, oxidation of, 537
Familial adenomatous polyposis (FAP), 1174
Familial hypercholesterolemia (FH), 651
Families, 81
Fanconi anemia, 1159t
Fas ligand, 1039
Fatty acid-binding proteins (FABPs), 464–465, 465f
Fatty acids, 51–52
    biosynthesis, 970
    double bonds and, 54f
    enzymes and, 464
    essential, 52
    movement, cytosolic proteins and, 464–465

oxidation of, 531–537, 532f
peroxisomal oxidation of, 536, 536f
in phospholipids, 52t
phospholipids, incorporation into, 465
polyunsaturated, 52
saturated, 52
unsaturated, 52, 464, 970–972
Fatty acyl group, 52
Fatty acyl-CoA dehydrogenase, 542
Fc domain, 768
Fc receptor (FcR), 768, 1108
F-class pumps, 480–481, 481f
Feedback circuits, 1048f
Feedback inhibition, 107
Feedback loops, 696, 697f, 737–738, 857f, 865, 876, 877f
Feedback mechanisms, 689–691, 721–722
Feedback repression, 669–670, 689, 959
Feed-forward activation, 518
Fermentation, 518, 519f
Fertilization, 991, 992
Fetal aneuploidy, 898f
FG-nucleoporin, 614–616, 615f
FH1 domain, 763
FH2 domain, 763, 764f
Fibril-associated collagens, 930t, 932, 934f
Fibrillar adhesions, 943
Fibrillar collagen, 930t, 932–934, 933f, 934f
Fibrillin-1 gene, 941

Fibrils, 94f, 95, 932, 933–934, 935
Fibroblast, 140, 204f, 262, 262f, 309f, 932, 939
Fibroblast growth factor (FGF), 709, 710, 710f, 716f
Fibronectin
alternative splicing, 204, 204f
antibodies, 908f
cell-binding region of, 940f
exons and, 280
functions of, 938–941
multi-adhesive matrix proteins, 904f, 906, 906t, 927
organization of, 939f
Fibrous proteins, 82
Fibulins, 928
Fight-or-flight response, 683, 685
Filamin, 770
Filopodia, 755, 788
Fimbrin, 770
Fingerprint, DNA, 277, 287–288, 288f
First polar body, 854f
Fischer, Emil, 98
Fiske, Cyrus, 64
Fission (mitochondrial), 527–529, 528f
Fission yeast, 852
5′ cap, 203, 333, 387–388
5′ capping, 387
5′ end, 180, 180f, 185–186, 333, 388, 388f
5′ methylated cap, 203f
5′ → 3′ direction, 185–186, 186f, 187, 187f, 200f, 201

Flagella, 15
Flagellum, 447, 820–825
beating, production of, 820
intraflagellar transport, 820–823, 822f, 823f
structural organization of, 821f
Flanking sequence, 244f
Flavin adenine dinucleotide (FAD), 517
Fleming, Alexander, 146
Flippases, 453, 465–466, 491
Flow cytometer, 141
Flow cytometry, 140–142, 856f
Flowers, 952
Fluid mosaic model, 443, 443f
Fluid-phase endocytosis, 768
Fluorescence in situ hybridization (FISH), 277, 313, 315f
Fluorescence microscopy, 86, 138, 148f–149f, 152–153, 153f, 155f
Fluorescence recovery after photobleaching (FRAP), 159–160, 161f, 450, 451f, 829
Fluorescence-activated cell sorter (FACS), 141, 141f, 142f
Fluorescent staining, 152
Fluorescently labeled dNTPs, 246, 247f
Fluorochromes, 152
Fluoxetine, 1073
Fmr1 gene, 1091
Focal adhesion kinase (FAK), 909f
Focal adhesion/focal contact, 913
Focal adhesions, 773, 788, 790
Focal complexes, 943

Focal contacts, 913t
Folded protein domains, 350f
Force generation, 788–790
Formaldehyde, 151
Formin, 763–764, 764f
Förster resonance energy transfer (FRET), 160–162, 161f, 162f, 680
48s initiation complex, 213
43S preinitiation complex, 213
FOS gene, 1171, 1171f
4-Hydroxyproline, 974
Fractions, 119
Fragile X syndrome (FXS), 1091
Frameshift mutation, 207, 249, 271f
Framework regions, 1106
Franklin, Rosalind, 8, 180
FRAP, 159–160
Free Ca^{2+}, 674–675
Free energy, 61–63, 97f
Free energy reactions, 100f
Free-energy change (ΔG), 61–63, 62f
FRET biosensor, 160, 162, 162f
Frizzled (Fz), 741, 1025
Frontal lobe, 1043f, 1044
Fructose 2,6-bisphosphate, 517f, 518
Fruit flies, 232–233
Functional complementation, 240–241, 242f
Functional domains, 79, 345, 346f
Functional light-chain gene, 1109–1111

Functional mRNA, 202–204, 203f
Functional RNAs, 284, 285f
Fungi, 852
Fura-2, 152
Furin, 647
Fusion (mitochondrial), 527–529, 528f
Fusion proteins, 371f

G

G protein–coupled receptor (GPCR), 661f, 662, 676–682, 1083–1084, 1083f
adenylyl cyclase, activation or inhibition of, 683–691, 683f
anchoring proteins, 688–689
cAMP activates protein kinase A, 683–685
cAMP-mediated activation of PKA, 687–688, 687f
feedback mechanisms, 689–691
gene transcription, activation of, 688, 689f
hormone-induced, 684f
PKA, hormone-induced activation, 685–687, 685f, 686f
receptor-ligand complexes, 683
signal amplification, 687
β_2-adrenergic receptor, 678f
binding of β-arrestin to, 691f
in cell adhesion, 904
classes of, 677f
effector proteins, regulation of, 681
exchange of GTP to GDP, 679–681, 680f, 681f

glucagon receptor, 679f
human hormones, identification of, 681–682
pharmaceutical importance, 677t
structure of, 676–678, 676f
G_1 cyclin-CDK, 857
G_1 phase, 21, 849
G_1/S phase cyclin-CDK, 857
G_2 phase, 21, 849–850
G-actin, 758
G-actin monomers, 756–757, 757f, 759f, 760f
Gain-of-function mutations, 1166f
GAL4 gene, 345, 346f
Galactose, 501–503, 502f
Gametes, 255, 256f, 990
Gamma-amino butyric acid (GABA), 1065
γ-Carboxyglutamate, 48f
γ-secretase, 740
γ-tubulin ring complex (γ-TuRC), 800, 801f, 879, 880f
Gangliosides, 449
Gap junction, 911, 912, 913t, 924f, 1076–1077
astrocytes and, 1049f
formation of, 1077
functions of, 912f, 913, 923–925
GARP protein, 734
Gasdermin pore structure, 1142f
Gated cation channels, 1078–1079, 1078f, 1079f
Gated channels, 473
GATOR1, 967, 967f

Gaucher's disease, 560
GCaMP, 165f, 166
GDP/GTP exchange, 109
Gel filtration chromatography, 119–121, 120f
Gelatinases, 942
Gel-shift assays, 343, 344f
Gelsolin, 763
GenBank, 250
Gene control, 324
Gene conversion, 197
Gene density, 283f, 286
Gene deserts, 286
Gene duplication, 281f
Gene expression, 178, 221
    cloned DNA fragments and, 260–268
        cDNA, sequencing, 263–264
        cluster analysis, 263
        DNA microarray, 261–263
        E. coli expression systems, 264–265
        plasmid expression vectors, 265–267
        in situ hybridization, 261
    control of, 705–750
        growth factors and receptor tyrosine kinases, 709–715
        JAK/STAT signaling pathway, 727–732
        overview of, 323–324, 705–709, 707f
        phosphoinositide signal transduction pathway, 724–726
        proteasomal degradation of signaling component, 740–750
        Ras/MAP kinase signal transduction pathway, 716–723

site-specific protein cleavage, 738–740
TGF-β family of growth factors, 733–738
heterochromatin and, 354–355
long-term memories and, 1090–1091
regulation of, 324, 324f
Gene family, 282–284
Gene hyperactivation, 273f
Gene inactivation, 273f
Gene interaction network, 237f
Gene knockout, 272, 331
Gene repression, 373–375, 376f
General import pore, 605
General transcription factors (GTFs), 201, 334, 335, 336–339, 337f
Gene-rich regions, 286
Genes, 2, 9, 199
altering function of, 268–275
CRISPR, 269–272
homologous recombination, 268–269
RNA interference (RNAi), 272–274
somatic cell recombination, 272
chromosomal organization, 286–289
definition of, 279
eukaryotic structure and organization, 278–285
evolutionary origins, 250
functions of, 250
genetic analysis of, 228–237
human traits and, 253–260
cancer, causative mutations, 260

cloned DNA and disease, 257–258
DNA polymorphisms, 255–256
genetic risk factors, identifying, 258–259
human linkage studies, 256–257
monogenic diseases, 253–255
multiple genetic defects, 258
protective alleles, 259–260
identification of, 249
identification of, in genomic DNA sequences, 249
structure of, 278f
tagging, 267
Genetic analysis, 228–237, 228f
complementation tests, 233–234
conditional mutations, 231–232
double mutants, 234–235
genetic suppressions, 235
global analysis, 236
recessive and dominant mutant alleles, 229
recessive lethal mutations in diploids, 232–233
segregation of mutations, 230–231
synthetic lethal mutations, 235–236
Genetic code, 10, 178, 206–207, 206t
Genetic complementation, 233–234, 234f, 281
Genetic diseases, 32–33
Genetic diversity, 194–198
Genetic heterogeneity, 258
Genetic mapping, 255–257, 256f, 257f
Genetic modifications, 227

Genetic screens, 231–232
Genetic suppressions, 235, 236f
Genome, 10t, 16, 221
Genome editing, 269–272
Genome-wide association study (GWAS), 259
Genomic analysis, 228f
Genomic libraries, 240–242, 242f
Genomics, 5, 28
Genotype, 228, 229
GEO (Gene Expression Omnibus), 379
Gephryin, 1068, 1068f
Germ cells, 894
Germ line, 1004
Germarium, 1004, 1005f
Germinal centers, 1113
Germ-line stem cells, 1004–1005, 1005f
Get1/Get2, 590–591
Get3 dimer, 590–591
GFAP (glial fibrillary acidic protein), 840
GGN2 (general control non-derepressible 2), 424
Gibbs, J. W., 61
Gibbs free energy, 61
$Glc_3Man_9(GlcNAc)_2$, 595, 597f, 601f
Glia, 1044–1051
Glial cells, 219, 998, 1026f, 1047
    myelin sheath and, 1047–1048
    myelin sheaths, production of, 1061–1063, 1061f, 1062f
    neural stem cells, 1048–1051, 1050f

synapses, formation of, 1065
Glial fibrillary acidic protein (GFAP), 1049f
Globular proteins, 82
Glomerular basement membrane, 931
Glomeruli, 1084
G-loop, 111, 113, 113f
Glucagon, 677, 960
Glucagon receptor, 677, 679f
Glucocorticoid receptor response element (GRE), 369, 370f
Glucocorticoid receptor response (GR), 371f
Glucomannans, 953
Gluconeogenesis, 688, 960, 963–964
Glucose
anaerobic vs. aerobic metabolism, 519f
conversion to pyruvate, 516–517
facilitate transport of, 474–480
fermentation of without oxygen, 518–520, 519f
metabolism, 518
Na^+-linked symporters and, 501–503
rehydration therapy and, 508
synthesis inhibited by insulin, 963–964
transcellular transport, 507–508, 508f
uptake of, 961–963, 962f
Glucose tranporters. *See* GLUT proteins
Glucose uniporter, 502
GLUT proteins, 477
GLUT1, 470, 475–477, 475f, 594, 961
GLUT1 uniporter, 475–477, 475f, 476f

GLUT2, 961, 961f
GLUT4, 174, 961–963, 962f
GLUT4 storage vescile (GSV), 961
Glutamate, 47, 1065
Glutamic acid, 47
Glutamine, 47
Glycine, 47, 75f, 929
Glycogen, 50
    catabolism, 683, 685–687
    degradation of, 684f
    degradation pathway, 687
    glucose, storage as, 963–964
    metabolism, regulation of, 686f
    synthesis of, 684f, 963
Glycogen phosphorylase (GP), 685
Glycogen phosphorylase kinase (GPK), 668, 686, 697, 697f
Glycogen synthase (GS), 685–686, 686f
Glycogenolysis, 683, 697, 697f
Glycolipids, 449, 460–461
Glycolysis, 66, 515–520, 516f
    aerobic, 1153–1154, 1154f
    anaerobic, 518–520, 519f
    in blood glucose regulation, 963–964
    overview of, 513f
    rate of, 517–518, 517f
    stage I, 516–517
    stimulating, 966
Glycolytic pathway, 516–517, 516f, 538t

Glycophorin A, 456, 456f
Glycophorin C, 771
Glycoprotein, 460, 595, 626
folding and stability of, 596–597
in plant cell walls, 952–953, 952f
Glycosaminoglycan (GAG), 51, 932, 935–937, 935f, 937f
Glycosidases, 596
Glycosidic bonds, 44, 51
Glycosphingolipids, 465
Glycosylation, 595
Glycosylphosphatidylinositol (GPI), 460, 460f, 592, 593f, 920
Glycosylphosphatidylinositol (GPI) membrane anchor, 648
Glycosyltransferases, 595
Glyoxisomes, 19–20
Goat–anti-rabbit preparation, 153
Goethe, Johann von, 72
Golgi, Camillo, 1086
Golgi cisterna, 629f
Golgi complex, 18f, 169f
anterograde transport, and cisternal maturation, 639–641, 640f, 641f
coated vesicles and, 623f
functions of, 18
localization of, 797f
retrograde transfer within, 638–639
transport from ER, 637–638
Golgi matrix proteins, 962f
Golgi stain, 1086–1087, 1087f

Goodpasture's syndrome, 931
Gout, 805
GPCR/cAMP/PKA pathway, 689–691
GPI-anchored proteins, 593f
G protein–coupled receptor, 678
G protein–coupled receptor kinases (GRKs), 690, 691f
Grafts, 1117–1118
Grana, 21, 559
Granulocyte colony–stimulating factor (G-CSF), 264, 264f, 727
Granzyme-mediated cells, 1137f
Granzymes, 1101f, 1137, 1137f, 1180
Grazing incidence-structured illumination microscopy, 530f
Green fluorescent proteins (GFP), 122, 153f, 154–155, 164f, 165, 267, 267f, 626, 626f, 641f, 810, 914f, 997
Greider, Carol, 318
GroEl/GroES, 92–94, 93f
Ground tissue, 952
Group I introns, 408, 436, 436f
Group II introns, 408–409, 408f, 436, 436f
Group VII ethylene response factors (ERFs), 974
Groups, 81
Growth control pathways, 1170–1171
Growth factors, 663, 709–715, 790, 792f
Growth hormone (GH), 666f, 727
Growth-factor receptors, 968–969, 968f
GRPase-activating protein (GAP), 701
GRPases-activating protein (Ran-GAP), 617, 617f
GSK3, 963

GTP analogues, 680
GTP hydrolysis, 644f
GTPase superfamily, 108, 218, 668, 717
GTPase switch, 109f, 631–633, 717
GTPase-activating proteins (GAPs), 669, 717, 790, 790f
GTPases, 91–92, 108–109
GTP-binding proteins, 668–669, 674, 675f
Guanine, 48, 49, 180, 180f
Guanine nucleotide dissociation inhibitors (GDIs), 109, 633, 790, 790f, 1016, 1017f
Guanine nucleotide exchange factors (GEFs), 109, 109f, 617–618, 617f, 716f, 717, 790, 790f
Guanosine diphosphate (GDP), 108–109, 668–669, 669f, 679
Guanosine triphosphate (GTP), 108–109, 668–669, 669f, 679, 702, 718, 719f
Guillain-Barre syndrome (GBS), 1062
Gustducin, 1081
$G_{\beta\gamma}$ subunit, 679–680, 680f, 684f

H

H&E stain, 151, 152f
H^+/K^+ ATPase, 509, 509f
H1, 300
H-2 complex, 1117
H2A, 300
H2B, 85f, 300
H3, 300

H4, 300
HA synthase, 937
Habituation, 1087
Haeckel, Ernst, 72
Hair cells, 402–403
Half-channels, 555
Half-life, 124
Haltere, 325f
Hansemann, David von, 1152
Haparan sulfate, 710f
Haplo insufficient, 229
Haploid cells, 229, 231
Haplotype, 259f
Hartwell, L. H., 232, 233, 848, 860
Hb^{s} gene, 229, 253
HD allele, 254, 255f
Head domain, 774, 778f, 811, 814
Heart muscle, 780–781, 783
Heat-shock elements (HSEs), 977, 977f
Heat-shock factors, 977–978
Heat-shock gene, 370
Heat-shock proteins (HSP), 976
Heat-shock response, 959, 976–978, 977f
Heavy chains, 96, 96f, 777, 1103, 1104f, 1107f
Heavy meromyosin (HMM), 775
Heavy-chain locus, 1111–1113
Hedgehog (Hh), 740, 744, 744f, 1172
    activation of signaling pathways, 706f–707f, 708

feedback regulation, 745–746
identification of, 706
receptors, 745
signaling in *Drosophila*, 746f
signaling in vertebrates, 746–747
HeLa cell line, 140, 170f, 390f, 439f, 525f, 528f, 697f, 855f
Helicase, 186, 338
Helicobacter pylori, 1154
Helix-loop-helix, 76f, 77
Helix-turn-helix, 77, 347
Helper T cells, 1117, 1120, 1138, 1145
Hemagglutinin (HA), 80f, 83, 599f, 600
Hematopoietic stem cell (HSC), 1009–1013, 1010f, 1011f
Hematoxylin, 151
Heme, 539, 539f
Heme groups, 102
Heme-regulated inhibitor (HRI), 424
Hemicellulose, 953
Hemidesmosomes, 912f, 913, 913t, 918–919
Hemoglobin, 107–108, 108f, 283
Hemophilia A, 254t, 296, 298
Henderson–Hasselbalch equation, 58, 59, 102
Henry VIII, King, 805
Heparin, 936
Hepatocytes, 143, 990, 1002
Heptad repeats, 76
Heptads, 76
HER, 711

HER family of receptors, 713f
HER1, 711, 712f, 715
HER2, 146, 672, 712, 1176–1177
Herceptin, 146
Herd immunity, 1146
Hereditary spherocytic anemias, 772
Heritability, 258
Hermaphrodite, 31
Herpes simplex type I virus (HSV-I), 341
Heterochromatin
    chromatin, 304f
    definition of, 17, 301
    formation of, 354–355
    histone H3 lysine 9 methylation in, 307, 373–375
    histone-tail post-translational modification in, 307f
    in nucleosome interactions, 302f
    repressed transgene, association with, 374f
Heterochromatin protein 1 (HP1), 308
Heterodimeric integrins, 904f
Heterodimeric nuclear receptors, 369
Heterodimeric transcription factors, 350, 351f, 352f
Heteroduplex, 197
Heterogeneity, 302f
Heterogeneous nuclear RNA (hnRNA), 385t, 389
Heterogeneous organs, 1154
Heterogeneous ribonucleoprotein particles (hnRNPs), 389, 415f
Heterologous desensitization, 690
Hetero-oligomerization, 728

Hetero-oligomers, 711
Heterophilic binding, 904
Heteroplasmy, 526
Heterotrimeric G protein, 668, 678f, 679–681, 680f, 681f, 682t, 706f–707f
Heterotypic adhesion, 904
Heterozygotes, 232–233
Heterozygous, 229
Hexokinase, 517
Hexoses, 50, 50f
Hif-1α, 959, 972, 973–974, 973f
Hif-1β, 973, 973f
High-density lipoproteins (HDL), 43
High-energy electrons, 514f, 533–534, 533f, 561–564
Highly expressed genes, 364f
High-salt solutions, 461–463
High-throughput LC-MS/MS, 133–134, 134f
Hippo kinase cascade pathway, 983–986, 983f, 985f, 986f, 987f
Hippo pathway, 960
Hippocampus, 1049, 1088, 1089f
Histidine, 47
Histone acetyl transferases (HATs), 306
Histone acetylase, 358
Histone acetylation, 305–306, 358, 359f
Histone code, 305, 308, 373
Histone deacetylases (HDACs), 306, 355
Histone deacetylation, 355–358, 357f
Histone H3, 373–375, 375f

Histone lysine demethylases, 373
Histone methyl transferase (HMT), 308
Histone octamer, 300
Histone tails, 304–311, 305f, 307f, 354, 357
Histone-modifying enzymes, 1167
Histones, 300, 301f, 303, 305f, 354
    methylation of specific histone lysines, 373–375, 375f
    post-translational modifications, 374t
HIV rev protein, 415–416, 416f
HLA complex, 1117
HMG (high-mobility group) proteins, 312
HMG-CoA reductase, 466
HML loci, 354–355, 354f
HMR loci, 354–355, 354f
hnRNA. *See* Heterogeneous nuclear RNA (hnRNA)
hnRNP proteins, 384, 387, 389, 390f
Holliday, Robin, 197
Holliday structure, 197, 197f, 198, 199f
Holoenzyme, 360, 360f
Homeodomain, 347–348
Homeostasis, 56, 959
Homeostatic chemokines, 1139
Homodimeric nuclear receptors, 369, 729f
Homogenate, 173, 173f
Homolog, 81
Homologous, 25, 31f
Homologous desensitization, 690
Homologous recombination

DNA repair and, 194–198, 198f, 891
exon shuffling and, 298f
gene duplication and, 281f
in meiosis, 894
mobile DNA elements and, 298
mutant alleles in yeast genes, 268–269, 269f
in somatic cell recombination, 272
Homologous sequences, 251
Homologs, 81
Homology, 81
Homo-oligomers, 711
Homophilic binding, 904
Homophilic crossbridges, 904f
Homotypic adhesion, 904, 907f
Homozygous, 229, 235
Hooke, Robert, 138
Hormones, 368, 663, 681–682
Horvitz, H. Robert, 1029
Host cells, 221–224
Host defense collagens, 930t, 933
Host defenses, 1095–1103
adaptive immunity as third line of defense, 1102–1103
circulatory and lymphatic systems, 1096–1098, 1096f
inflammation, 1100–1102
innate and adaptive immune systems, 1096–1098, 1097f
innate immunity as second line of defense, 1098–1100
mechanical and chemical boundaries, 1098
pathogens and, 1095–1096

Host range, 219
Housekeeping genes, 334
Hox genes, 375, 377f
HOXD13 gene, 324, 325f
Hps93, 610
Hsp40, 92
Hsp60, 92
Hsp70, 90–92, 91f
Hsp90, 91f, 92
Human Cell Atlas Project, 902
Human genome, 249, 279t
Human growth hormone (HFH), 588, 594f
Human immunodeficiency virus (HIV-1), 223, 339–340, 415–416, 443f, 656–657, 657f, 951
Human kinome, 82
Human linkage studies, 256–257
Human papillomaviruses (HPV), 224
Human T-cell lymphotrophic virus (HTLV), 223
Human traits
    complex, 258
    genes and, 253–260
        cancer, causative mutations, 260
        cloned DNA and disease, 257–258
        DNA polymorphisms, 255–256
        genetic risk factors, identifying, 258–259
        human linkage studies, 256–257
        monogenic diseases, 253–255
        multiple genetic defects, 258

protective alleles, 259–260
Humans, ribosomes, structure of, 212f
Hunt, Tim, 860
Huntington's disease, 254, 254t, 255f, 287
Hutchinson-Gilford progeria, 842
Hyaluronan, 937–938
Hyaluronate, 937
Hyaluronic acid (HA), 937
Hybridization, 184, 261
Hybridomas, 144–146, 145f
Hydration shell, 40
Hydrocarbons, 42–43
Hydrochloric acid, 508–509
Hydrogen bonds, 40–41, 41f, 181–182, 181f
Hydrogen ions, 58
Hydrolysis, 44, 64–65, 64f, 101f
Hydropathic index, 593
Hydropathy profile, 593, 594f
Hydrophilic, 34, 444
Hydrophilic amino acids, 47
Hydrophobic, 34, 444
Hydrophobic amino acids, 46
Hydrophobic core, 581
Hydrophobic effect, 42–43, 42f, 444f
Hydroxyl *O*-linked oligosaccharaides, 936f
Hydroxyproline, 929
Hyperacetylated chromatin, 306
Hyperacetylation, 357f

Hypercholesterolemia, 254t
Hyperchromicity, 183
Hyperpolarization, 1054
Hypertonic solutions, 478
Hypertrophic cardiomyopathies, 780
Hypervariable regions, 1106, 1107f, 1132
Hypoacetylated chromatin, 306
Hypoacetylation, 355
Hypoglycemia, 960
Hypothalamus, 727, 981, 1064f
Hypotonic solutions, 478
Hypoxia-responsive-elements (HRE), 972

I

I-cell disease, 645
Icosahedron, 220–221
Identities, 81
I-domain, 919
IFT complex A and B, 823
Ig, membrane-bound and secreted, 1115–1116, 1115f
IgA molecule, 1105, 1105f
IgCAMs, 949
IgG molecule, 1105, 1105f
IL-2 gene, 1134
Imatinib, 1162
Imidazole, 47
Imipramine, 1073

Immune checkpoint pathways, 1181
Immune system. *See* Immunology
Immunity, 1093
Immunoaffinity chromatography, 121
Immunoblotting (IB), 122, 123f, 1095. *See also* Western blotting
Immunoediting, 1179–1181
Immunoelectron microscopy, 169
Immunofluorescence microscopy, 153–154, 1095
Immunoglobulin (Ig) domains, 96, 96f
Immunoglobulin (Ig) superfamily, 904, 904f
Immunoglobulin fold, 1107f, 1108
Immunoglobulins, 96, 600, 1095, 1103–1108. *See also* Antibody
    constant region, 1108
    definition of, 1102
    fold, 1107f, 1108
    heavy and light chains, 1103
    isotypes of, 1104–1106, 1104f, 1116
    naive B cells, 1106–1107
    somatic gene rearrangement, 1110f
    structure of, 1104f, 1107f
Immunological checkpoints, 1147
Immunology, 1093–1147
    adaptive response, collaboration of immune-system cells, 1140–1147
        antigen-presenting cells, activation of, 1143
        cancer, 1146–1147
        high-affinity antibodies, 1143–1145, 1144f
        toll-like receptors, 1140–1143, 1141f

vaccines, 1145–1146
antibody diversity and B-cell development, 1109–1116
during adaptive response, 1115–1116, 1115f
functional light-chain gene, 1109–1111
heavy-chain locus, 1111–1113
immunoglobulins isotypes, 1116
pre-B-cell receptor, 1113–1115, 1114f
somatic hypermutation, 1113
cancer and, 1181–1183
chemical defenses, 1094f
helper T cells, subsets of, 1138
host defenses, 1095–1103
adaptive immunity as third line of defense, 1102–1103
circulatory and lymphatic systems, 1096–1098, 1096f
inflammation, 1100–1102
innate and adaptive immune systems, 1096–1098, 1097f
innate immunity as second line of defense, 1098–1100
mechanical and chemical boundaries, 1098
pathogens and, 1095–1096
immunoglobulins, 1103–1108
constant region, 1108
fold, 1107f, 1108
heavy and light chains, 1103
isotypes, 1104–1106, 1104f
naive B cells, 1106–1107
structure of, 1104f, 1107f
innate lymphoid cells, 1138–1139
leukocytes, 1139

mechanical defenses, 1094f
MHC and antigen presentation, 1117–1128
antigen presentation, and protein fragments, 1123
class I MHC pathway presents cytosolic antigens, 1123–1125, 1124f
class II MHC molecules presents antigens to endocytic pathway, 1125–1128, 1126f
cytotoxic T cells, 1118
grafts and, 1117–1118
MHC molecules, characteristics of, 1120–1122
organization of, 1118f
T cells guided by MHC molecules, 1118–1120
overview of, 1093–1095
T cells
activation, signals required for, 1136–1137, 1136f
CD4 or CD8 lineage, 1135–1136
CD8 co-receptor, 1137, 1137f
cytokines, secretion of, 1138
MHC molecules, recognition of, 1134–1135
proliferation and differentiation, 1132–1134, 1133f
T-cell receptors, 1130f, 1131f
similarity to immunoglobulin genes, 1131–1132
structure of, 1129–1130
variable residues, encoding of, 1132
three layers of, 1094f
Immunoprecipitation (IP), 122–124, 673
Immunoproteasome, 1123

Immunoreceptor tyrosine-based activation motif (ITAM), 1114, 1130, 1134
Immunotherapy, 1147, 1181–1183
Impila, 557
Import receptor, 604–605
Importin, 616, 617f, 830
In situ hybridization, 257f, 258, 261, 261f, 288f, 309f, 318, 365
In vivo transcription, 332f, 334
In vivo transfection assay, 345f
Inbreeding, 232–233
Indian muntjac, 313, 314f
Indirect immunofluorescence microscopy, 153, 154f
Indirect-acting carcinogens, 1157
Induced fit, 43, 98
Induced pluripotent stem (iPS) cell, 991, 997–1000, 1000f
Induced regulatory T cells, 1138
Inflammasome, 1098, 1099f, 1108, 1141–1142, 1142f
Inflammation, 1100–1102, 1138–1139
Inflammatory chemokines, 1139
Inflammatory response, 1101, 1101f
Inflammatory T cells, 1138
Inhibitory interneuron, 1047f
Inhibitory postsynaptic compartments, 1068f
Inhibitory receptor, 1076
Inhibitory synapse, 1044, 1065
Initiation, 323, 370–371, 380–382
Initiation codon, 206, 214f
Initiation proteins, 1029

Initiation stage of transcription, 201, 201f
Initiator, 334
Initiator caspase, 1032
INK4s (inhibitors of kinase 4), 863
Innate immune system, 1094f, 1096–1102, 1101f
Innate lymphoid cells (ILCs), 1098, 1138–1139
Inner cell mass (ICM), 991, 993–994, 994f, 995f
Inner ear hair cells, 402–403
Inner mitochondrial membrane, 20, 534, 548–549, 557
Inner nuclear membrane, 16
Inner-arm dyneins, 820
Inner-membrane proteins, 606–607, 607f, 608f
Inositol 1,4,5-trisphosphate (IP_3), 693–694, 724
Insertion sequences, 290–291, 291f
Inside-out signaling, 909, 909f, 919
Insig-1(2)/SCAP/SREBP pathway, 971
Insula, 1081
Insulin, 960–961, 963–964
Insulin receptor, 711f
Insulin-dependent diabetes. *See* Type I diabetes
Integral membrane proteins, 82, 455, 587. *See also* Transmembrane proteins
Integrase, 293
Integrated retroviral DNA, 293f
Integrin-linked kinase (ILK), 909f
Integrins
    adhesion, regulation of, 944–948
    adhesive structures, clustering into, 943f

cell-adhesion molecules, 904
cell-matrix adhesions, mediation of, 918–919
conformations of, 945f
definition of, 735
expression, 948
fibronectins, binding to, 939f, 940f
focal adhesions and, 788
recycling of, 790
regulation of integrin-mediated adhesion, 944–948, 946f–947f
role of in cell adhesion, 943–944
signaling pathways, 909f
in vertebrates, 919t
Intercalated disks, 780
Interferons, 727–728
Intergenic regions, 286
Interleukin, 1138
Interleukin 1 (IL-1), 706f–707f, 747
Interleukin 2 (IL-2), 728, 729f, 1138
Interleukin 4 (IL-4), 1138
Intermediate filament-associated proteins (IFAPs), 843
Intermediate filaments, 15, 15f, 754, 754f, 796, 837–842
assembly from subunit dimers, 838, 838f
classes of, 840t
cytoplasmic, expression of, 839–841, 839f
dynamism of, 839
functions of, 913
image of, 798f
lamins, 841–842, 842f

localization of, 838f
overview of, 797f
plectin cross-links, 843f
proteins and cellular organization, 843
structure of, 838f
Intermediate-repeat DNA, 289
Intermembrane space, 20
Intermembrane-space proteins, 607–609, 607f, 609f
Internal aqueous space, 445–447
Internal face, 446
Internal ribosome entry site (IRES), 213
Interneurons, 1044, 1046, 1047f
Internode, 1062
Interphase cells, 823–825, 824f, 827f
Interspersed repeats, 286, 289, 298f
Interstitial fluid, 1096f
Intersystem crossing, 563
Intestinal stem cells, 1005–1007, 1006f, 1007f
Intestinal villi, 901f
Intracellular cues, 894
Intracellular ionic environment, 480–484
Intracellular movements, 766–767, 767f
Intracellular vesicles, 961–963, 962f
Intraflagellar transport (IFT), 820
Intravital imaging, 158, 159f
Intrinsic polarity program, 1016
Intrinsic regulator signals, 1004
Intrinsically disordered proteins, 82, 83f, 90, 361

Intrinsically disordered region (IDR), 82, 85f, 418
Intronic splicing enhancers, 404
Intronic splicing silencer, 403
Introns
in alternative RNA splicing, 204, 204f
definition of, 202
exon shuffling and, 298f
in multicellular eukaryotes, 279–282, 298
splicing of, 391, 391f, 397
Invadopodium, 1156
Invagination, 18
Invariant chain, 1122
Invariant natural killer T cells (iNKT), 1136
Inverted repeat, 290–291, 291f, 369
Ion channels
molecular selectivity filter, 496–498
novel, 499–500
patch clamps and, 498–499, 499f
Ion concentrations, 152–153, 482t
Ion-exchange chromatography, 120f, 121
Ionic bonds (ionic interaction), 39–40, 40f
Ionic detergents, 462, 462f
Ionic gradients, 482
Ionizing radiation, 193
Ionophore, 548
Ionotropic glutamate receptors, 1075
Ionotropic receptors, 1065
IP_3. *See* Inositol 1,4,5-trisphosphate (IP_3)

IP_3 receptors, 694f
IP_3/DAG pathway, 693, 694f, 697f
IP_3-gated Ca^{2+} channels, 694–695, 695f
Iron Response Element Binding Proteins (IRE-BPs), 425
Iron Response Elements (IREs), 425, 425f
Iron-sulfur clusters, 539–540, 539f
IS element. *See* Insertion sequences
I-Smad, 735
Isoelectric focusing (IEF), 119, 119f
Isoelectric point (pI), 119
Isoform, 204, 909
Isoleucine, 46
Isopentenyl pyrophosphate (IPP), 466–467, 467f
Isopeptide bond, 106, 114f
Isoproterenol, 672, 672f
Isotonic medium, 478
Isotype, 1104
Ivacaftor, 492–493

J

J gene segments, 1109–1111, 1132
JAK kinase, 727, 729f, 730
JAK protein tyrosine kinases, 728–730
JAK/STAT signaling pathway, 727–732
Jun N-terminal kinases (JNKs), 720
Junction adhesion molecule (JAM), 921, 921f, 949
Junctional imprecision, 1110–1111

Junk DNA, 278
Juxtaparanode, 1063

K

K^+ channels, 1054–1055, 1054f, 1057f, 1058f
K^+ ions, 495–496, 497f
K^+-channel protein isoforms, 402–403
Kandel, Eric, 1087
Karyomeres, 887f
Karyotype, 313, 314, 314f, 1152f
Katanin, 809
KDEL receptor, 638, 639f
KDEL sorting signal, 638
Kearns-Sayre syndrome, 526
Kendrew, John, 129
Keratin, 839–840, 839f, 841f
Keratinocytes, 839, 997
KH domain, 389–390
Killer proteins, 1029
Kilocalorie, 61
Kinase, 86, 92, 110, 673
Kinase cascade, 717
Kinase domain, 111, 112f
Kinase-associated proteins (AKAPs), 688–689
Kindlins, 946
Kinesin, 809
Kinesin proteins, 754, 809–819

in anaphase B, 833
dyneins, cooperation with, 817f
head domain, 814f
superfamily of, 812, 812f
Kinesin-1, 810–811, 811f, 812–814, 812f, 813f, 814f
Kinesin-2, 812f
Kinesin-5, 147, 147f
Kinetic energy, 60
Kinetochore
centromere sequences and, 319
definition of, 826
functions of, 829–831, 831f
microtubule interaction, 320f, 832, 833f, 850
mitotic spindles and, 879–882, 892, 893f
structure of, 830f
Kinetochore microtubules, 828, 829f
King, Mary-Claire, 891
Kinome, 110
KKXX sorting signal, 639
Kleisins, 303
Knee-jerk reflex, 1047, 1047f
Knock-in strategy, 1174
Knockout mice, 321, 715
Knockout strategy, 1174
Kornberg, Arthur, 185
Koshland, Daniel, 98
Kozak, Marilyn, 213
Kozak sequence, 213

Krebs cycle, 533

L

Lac promoter, 264, 264f
Lacks, Henrietta, 140
Lactic acid fermentation, 519f
Lactose, 51f
LacZ gene, 264f
Lagging strand, 187, 187f, 189f
Lamellipodium, 755, 788
Laminin, 906, 906t, 927, 928, 928f, 929f
Laminopathies, 842
Lamins, 16, 839, 841–842, 842f
Lampbrush chromosome, 384
Langerhans cells, 1140
Large rRNA, 211
Large T-antigen, 187, 188f, 189f
Laser-scanning confocal microscope (LSCM), 156
Late endosome, 625
Late gene expression, 221
Latent TGF-β binding protein (LTBP), 734
Lateral, 911
Lateral surface, 27, 911–912, 911f
Latrunculin, 769
LC-MS/MS, 133–135, 134f, 135f
LDL receptor, 651
LDLR gene, 651, 652f

Leading edge, 755, 767
Leading strand, 187, 187f, 189f
Leaflets, 445
Leaves, 560f, 952
Leber's hereditary optic neuropathy, 526
Lectin, 460, 904
Lectin domain, 950
Legionnaire's disease, 109
Lentiviruses, 266, 266f
Leptin receptor positive, 1011, 1012f
Let-7 gene, 419
LEU gene, 319f
Leucine, 46
Leucine zipper, 77, 349
Leucine-rich repeats, 1140
Leukemia, 223, 1149
Leukemia inhibitory factor (LIF), 995f, 996
Leukocyte-adhesion deficiency, 951
Leukocytes, 949–951, 950f, 1096, 1139
Lever arm, 777
Levi-Montalcini, Rita, 711, 1033
LG domains, 928
Lgr5 gene, 1007–1008, 1008f, 1009f
Lgr5^{+} intestinal stem cells, 1007–1009, 1008f, 1009f
Lidocaine, 1059
Ligand, 56, 57f, 95, 662
    affinity of receptors, 670–672, 671f
    binding, 96f

binding of, 95–97
receptors binding to, 665
receptors binding to with high affinity, 665–667
Ligand domain, 160, 162
Ligand-binding curve, 670
Ligand-binding domain, 371f
Ligand-binding site, 96, 96f
Ligand-gated channel, 1074
Ligand-induced dimerization, 1103
Light adaptation, 701, 701f
Light chains, 96, 96f, 775, 777, 779f, 1103, 1104f, 1107f
Light energy, 566–569
Light meromyosin (LMM), 775
Light microscopy, 148–166
deconvolution and confocal, 155–156
development of, 139, 148f–149f
fluorescence, 152
FRAP, 159–160
FRET, 160–162
immunofluorescence, 153–154
light-sheet, 165–166
nanometer resolution, 162–163
optogenetics, 162
phase-contrast and differential-inference-contrast, 150–151
resolution, 150
specimen requirements, 151
super-resolution, 163–165
tagging with fluorescent proteins, 154–155

TIRF, 158
tissue preparation, 151, 152f
two-photon excitation, 157–158
Light reactions, 561
Light-activated ion channels, 1063–1064, 1063f
Light-harvesting complexes (LHCs), 512, 561, 562f, 564, 565f
Light-sheet microscopy, 165–166, 165f
Lignins, 953
Limb development, 327–328, 329f, 421f
Limit of resolution, 150
Lin-4 gene, 419
LINC complexes, 842f
Linear electron flow, 568
LINEs (long interspersed elements), 295–296, 296f, 297f, 298, 299f
Link domain, 938
Linkage, 255
Linkage disequilibrium, 259, 259f
Linkage mapping, 255–257, 256f, 257f
Linker, 815
Linker DNA, 301, 301f, 302f, 303
Linker domain, 811
Linker histone, 300
Linker scanning mutagenesis, 340, 341f, 342
Linkers, 243f, 246f
Lipid bilayer, 442, 444–454, 444f, 445f, 455
Lipid droplets, 454, 454f
Lipid raft, 453
Lipid-anchored membrane protein, 455

Lipid-binding motifs, 461, 462f
Lipids, 43
classes of, 447–449, 448f
composition of, 450–453, 452f
covalently attached, 459–460, 460f
excess, storage of, 454, 454f
lateral mobility, 449–450, 451f
major components, 450t
Lipmann, Fritz, 64
Lipoprotein complexes, 650–651, 650f, 652f
Lipoproteins, 43, 53, 650
Liposome, 444
Liquid chromatography, 119–121, 120f, 329f
Liquid-liquid phase separation, 84, 85f
LIS1 protein, 817f
Lissencephaly, 816
Listeria, 766–767, 767f, 773, 788
Live cell imaging, 149, 151f
Live cell microscopy, 151
Living systems, 5f
Lohmann, Kurt, 64
Long interspersed elements (LINEs), 295–296, 296f, 297f, 298, 299f
Long noncoding RNAs (lncRNAs), 284, 324, 377–379
Long nonprotein-coding RNA, 309
Long terminal repeats (LTRs), 292, 292f
Long-term depression (LTD), 1088, 1089f
Long-term memory, 1088f, 1090–1091
Long-term potentiation (LTP), 1088, 1089f

Loss of heterozygosity (LOH), 1164
Loss-of-function mutations, 1166f
Lou Gehrig's disease, 998
LOV domain, 162, 163f
Low-density lipoproteins (LDL), 43, 649, 650, 650f, 651f, 652f, 970
Low-molecular-weight G proteins. *See* Monomeric G protein
LoxP sites, 272
LoxP-Cre recombination system, 272, 274f
LTR retrotransposons, 292–295, 292f, 293f, 296
Luciferase, 267
Lumen, 446, 580f
Luminal sorting signals, 633
Lung cancer, 1158
Luteinizing hormone (LH), 696
Luteinizing hormone-releasing hormone (LHRH), 696
Lymph, 1097, 1097f
Lymph nodes, 1096–1098, 1096f, 1097f, 1101f
Lymphatic system, 1096–1098, 1096f
Lymphocytes, 385, 1096, 1097f, 1098
Lymphomas, 1149
Lynch syndrome, 1159
Lysine, 47, 100f, 114f, 373–375, 375f
Lysophosphatidic acid (LPA), 792f
Lysosomes
    acidity of, 486–487, 487f
    degradation of, 714–715
    degradation of membrane proteins and cytosolic materials, 654–659, 655f

autophagic pathway, 657–659, 658f
multivesicular endosomes, formation of, 656–657, 656f
plasma membrane, retroviruses bud from, 656–657, 656f, 657f
segregation of membrane proteins by multivesicular endosomes, 655–656
functions of, 19, 19f, 172
lysosomal storage diseases, 645
mannose 6-phosphate and, 644–645, 645f
Lysozyme, 1098
Lytic cycles, 221–222, 222f

M

M (mitotic) phase, 21, 848
M domain, 581
Macromolecules, 7, 12, 44, 45f, 57f
Macronucleus, 318
Macrophage, 1096, 1098, 1100
Madin-Darby canine kidney (MDCK) cells, 142f, 143, 143f, 452, 648, 914, 914f, 917f
Maintenance methyl transferase, 373
MAIT cells, 1128
Major histocompatibility complex (MHC), 1095, 1179
class I MHC molecules, 1118–1125, 1121f, 1122f, 1124f, 1125f
class II MHC molecules, 1093f, 1118–1120, 1121f, 1122, 1125–1128, 1126f
cytotoxic T cells, 1118f, 1119f

grafts and, 1117–1118
organization of, 1118f
polymorphisms, 1120–1122
recognition of by T cells, 1134–1135
T cells, guidance of, 1118–1120
Malaria, 25–27, 26f
Malate-aspartate shuttle, 534, 535f
Malignant, 1155
Malignant tumor, 1155
Mammals
cell cycle regulation in, 855
central nervous system, 1043f
early development, 992–1001, 993f
embryo, cleavage of, 992–993, 993f
fertilization, 992
hippo kinase cascade pathway, 983–984, 983f
suprachiasmatic nucleus, 981–982, 982f
x-chromosome activation in, 308–309
Mannose 6-phosphate (M6P), 644–645, 645f, 646f
Mannose 6-phosphate receptors, 644
Mannose-binding lectin pathway, 1099, 1100f
Manor groove, 181
MAP kinase, 706f–707f, 717, 718–722, 719f, 721f, 723f
Marfan syndrome, 941
Marine sponges, 907f
Mass, centrifugation and, 116–117, 117f
Mass spectrometry (MS), 126–128, 126f, 127f, 133–136, 865–866, 866f
Mast cell, 1102

Master splicing factors, 404
Master transcription factors, 29, 30f, 372, 708f, 709, 995, 996f
MAT locus, 428
The Materials of Medicine (Dioscorides), 557
Maternal mRNA, 992
Mating factor, 24f
Mating-type genes, 354f
Mating-type switching, 429f
Matriglycan, 936, 948
Matrisome, 926–927
Matrix, 20, 135, 602
Matrix junctions, 911–926
Matrix metalloproteases (MMPs), 942
Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry, 126f
Matrix-targeting sequences, 603
Maturases, 409
Maturation-promoting factor (MPF), 855, 860, 861
Maximum transport rate (V_{max}), 475, 475f
MCM (minichromosome maintenance) proteins, 190
MCM helicase, 871, 872f
MCTPs, 955
Mechanical energy, 60
Mechanochemical motor proteins, 775–777
Mechanoreceptors, 958, 1078–1079, 1078f, 1079f
Mechanosensation, 956
Mechanosensor, 909, 910f, 1078
Mechanotransducers, 1079f

Mechanotransduction, 909–910, 910f, 955–956
Media, for animal cell cultures, 139–140
medial-Golgi cisterna, 625, 626, 628, 629f, 639, 640, 640f
Mediator, 326f, 354, 359–361, 360f, 361f, 365
Mediator of transcription complex, 354
Medium, 671
Medulloblastomas, 747
Meiosis, 190, 195, 230, 230f, 255, 256f, 894–899
    chromosome segregation, 897–898, 897f
    cohesin function during, 897f
    germ cell formation, 894
    vs. mitosis, 894–895, 895f
    overview of, 896f
    sister kinetochores, co-orienting, 898–899
Meiosis I, 854, 894–895, 897–898, 897f
Meiosis II, 854, 894, 898
MEK kinases, 720, 722, 723f
Melanocytes, 843
Melanomas, 193, 1181
Melanophores, 818f
Melanosomes, 818, 818f, 843
Melatonin, 981
Melting. *See* Denaturation
Melting temperature (T_m), 183, 183f
Membrane attack complex, 1098–1099
Membrane contact sites (MCSs), 529, 530f
Membrane depolarization, 1056–1059
Membrane nucleoporins, 614

Membrane potential, 472, 1045, 1045f
Membrane proteins
beta (β) strands in porins, 458–459, 459f
classes of, 472–474, 472f, 473f, 587f
covalently attached lipids, 459–460, 460f
degradation of, 654–659, 655f
functions of, 455–463
insertion into ER, 586–594
lipid-binding motifs, 461, 462f
orientation of, 592f
removal by detergents or high-salt solutions, 461–463, 462f, 463f
segregation by multivesicular endosomes, 655–656
structure of, 455
synthesis of on ER, 587–588
topology of, 593–594
Membrane recycling, 788–790
Membrane transport proteins, 71, 442, 470–471
Memory, 1094
formation and storage of, 1086–1091
gene expression and, 1090–1091
hippocampus, 1088
long-term memories, 1088f
synapses, number and strength of, 1086–1088
synaptic plasticity, 1088–1090, 1089f
Memory cells, 1138
Mendelian genetics, 230
Menten, Maud Leonora, 98
Meristem, 953, 1013–1015, 1014f

Merotelic attachment, 881
Mesenchymal cells, 915
Mesenchyme, 993
Mesophyll cells, 573
Messenger ribonuclear protein complex (mRNP), 618–619
Messenger RNA (mRNA), 2, 11f, 71, 205
cDNA library and, 242, 242f
decoding by tRNAs, 205–210
definition of, 385t
degradation of, 417–422
encoding single proteins, 279–280
export from nucleus by Ran-independent mechanism, 618–620, 619f
in situ hybridization and, 261
localization of, 428–430
nuclear export of, 413, 413f, 414f, 415f
quality control, 413
regulation of
by miRNAs, 421
protein synthesis and, 423–424
roles of in protein synthesis, 205, 205f, 206–207
structure of, 217f
transcription and, 87
surveillance mechanisms, 426–428, 426f, 427f
translation of
and cytoplasmic polyadenylation, 422–423
cytoplasmic polyadenylation and, 423f
enhanced rate, 965–966

iron-dependent regulation, 425f
sequence-specific RNA-binding proteins and, 424–426
transport of
across nuclear envelope, 412–416
in yeast, 430f
unspliced viral, transport of, 415–416, 416f
Metabolic cooperation, 923
Metabolic coupling, 83, 103, 923, 925
Metabolic intermediates, 516
Metal shadowing, 167–168, 168f
Metalloproteases, 738, 739f, 740, 942
Metaphase, 827f, 828
Metaphase chromosomes, 302, 303, 311f
Metaphase plate, 831–832, 850
Metastasis, 1151, 1155–1156, 1156f
Metazoans, 1f, 5, 27–30, 902, 908
epigenetic repression in, 377–379
epithelia, 27
evolution and cell functions, 27–28
multicellularity, 27
in research, 31–33
structure of, 27–28
transcription, regulation of, 370–371
transcription factors, 29–30
Methionine, 46, 206, 212–213
Methionyl-tRNA, 212–213
3-Methylhistidine, 48f
Met-$tRNA_i^{Met}$, 213, 215, 215f

Mevalonate, 466–467, 467f
Meyerhof, Otto, 64
MHC products, 1095
MHC restriction, 1118
Mice, 272, 274f, 321
Micelle, 444–445
Michaelis, Leonor, 98
Michaelis constant (K_m), 98f, 99
MICOS (mitochondrial contact site and cristae organizing system), 522, 556f
Microexons, 406–407
Microfibrils, 932, 934
Microfilaments, 15, 15f, 754f
    actin filament assembly, 763–769
        Arp2/3 complex, 764–766, 765f, 766f, 768f
        formins, 763–764, 764f
        phagocytosis, 769f
        polymerization, 766–767, 767f
        toxins and, 769
    actin filaments, 752f
        capping proteins, 762–763, 762f
        growth of, 760–761, 760f
        polymerization, 758–759, 759f
        thymosin-β_4, 762
        treadmilling, 761–762, 761f
    actin filaments, dynamics of, 758–763
    actin structures and, 755–758
        F-actin, polarity of, 757–758, 758f

G-actin monomers, 756–757, 757f
overview of, 755–756
actin-based cellular structures, organization of, 770–773
adapter proteins, 770–773, 772f
cross-linking proteins, 770, 771f
Cdc42 protein and, 843–844, 844f
cell migration, 787–793
chemotactic molecules, 793, 793f
coordinate regulation of Cdc42, Rac, and Rho, 791–792
force generation coordinated with cell adhesion and membrane recycling, 788–790
small GTP-binding proteins, 790–791, 791f, 793f
steps in, 788f
definition of, 754
endocytosis, function in, 767–768, 768f
examples of, 756f
melanosomes, transport of, 843
myosin-powered movements, 780–786
actin and myosin II in nonmuscle cells, 783–784
contraction in smooth muscle and nonmuscle cells, 784–785, 785f
myosin thick filaments slide past actin thin filaments, 780–781, 781f
myosin V carries vesicles, 785, 786f
skeletal muscle contraction, regulation of, 782–783, 783f
stabilizing and scaffolding proteins, 781–782, 782f
myosins, 773–779
actin filaments, steps along, 777–779, 778f

ATP hydrolysis and movement, 777, 778f
classes of, 776f
functions of head, neck, and tail domains, 774–775
mechanochemical motor proteins, 775–777
movement, rate of, 779f
superfamily in humans, 776f
neural growth cones, 844, 844f
overview of, 752–755, 797f
Microglia, 1047
MicroRNA (miRNA)
definition of, 274
functions of, 284, 417, 419–421, 419f
gene regulation, 385
in limb development, 421f
overview of, 385t
processing, 420f
tumorigenesis and, 1166–1167
Microsatellites, 256, 286, 287, 287f, 288
Microsome, 579, 580f
Microtubular protein, 802
Microtubule-associated protein (MAP), 802, 806, 806t, 807f
Microtubule-organizing center (MTOC), 799–800, 799f, 800f, 805, 806f
Microtubules, 15, 15f, 147f, 158f, 754, 754f, 796
Cdc42 protein and, 843–844, 844f
cilia and flagella, 820–825
beating, production of, 820
doublet microtubules bridged by dynein motors, 820

intraflagellar transport, 820–823, 822f, 823f
primary cilium, 823–825, 824f
structural organization of, 821f
dynamic instability of, 802–805, 803f, 804f
dynamics in mitosis, 828–829, 830f
growth and shrinkage of, 803f
growth of, 802f
image of, 798f
localized assembly and search and capture, 805
melanosomes, transport of, 843
mitotic spindles, 828, 829f
motor proteins and, 809–819
cooperation of kinesins and dyneins, 817–818, 817f, 818f
dynein motors, 814–816, 815f, 816f
kinesin protein superfamily, 812, 812f
kinesin-1 as processive motor, 812–814, 813f, 814f
kinesin-1 powers anterograde transport of vesicles, 810–811, 811f, 812f
transport of organelles in axons, 810, 810f
tubulin modifications, 818–819, 819f
neural growth cones, 844, 844f
overview of, 797f
plectin cross-links, 843f
regulation of, 806–809
(+) end, and +TIPs, 807, 808f
disassembly, 807–809, 809f
end-binding proteins, 807–809, 809f
severing proteins, 809

side-binding proteins, 806–807
reorganization in plant cells, 836–837, 836f
shortening of, 833
structure and organization of, 797–801, 798f
MTOCs, assembled from, 799–800, 799f, 800f, 806f
polarized structures built from αβ-tubulin dimers, 797–799, 799f
γ-tubulin ring complex (γ-TuRC), 801f
tubulin polymerization, drugs affecting, 805
Microvillus, 752, 773, 1081
Migrating cell, 753f
Miller-Dieker lissencephaly, 816
MINFLUX, 165
Minisatellites, 288, 288f
Minor groove, 181, 182f
Miranda, 1026
Miro, 527
Mismatch excision repair, 192–193, 193f
Missense mutation, 207, 249–250
Mitchell, Peter, 550
Mitochondria, 20f, 576f
in apoptosis regulation, 1034–1035
brown fat and heat generation, 558
contact sites with ER membrane, 527–531
mitochondrial function and dynamics, 529–531, 530f
cryoelectron microscopy and, 131f
electron transport in, 537–538
evolution of, 523–524, 524f

functions of, 20–21, 522t
fusion and fission, 527–529, 528f
genetic code, 526
intracellular movements, 527
membranes and, 446–447, 446f
mitophagy, 529
mutations, 526
oxidation, rate of, 558
protein import
chimeric proteins, 605, 606f
energy inputs, 605–606
structure of, 520–526, 521f
membranes, 520–523
mtDNA, 524–526, 525f
targeting proteins to, 602–611, 603f
amphipathic *N*-terminal targeting sequences, 603–604, 604f
targeting sequences, 578t
Mitochondria-associated membranes (MAMs), 530f, 695, 695f
Mitochondrial calcium importer (MCU), 695
Mitochondrial cristae, 522
Mitochondrial DNA (mtDNA), 523–526, 525f
Mitochondrial inner membrane, 520
Mitochondrial intermembrane space, 522
Mitochondrial matrix, 523, 525, 525f, 549, 603–604, 604f
Mitochondrial membrane, 1035–1037
Mitochondrial outer membrane, 517, 520
Mitochondrial precursor proteins, 603f
Mitochondrial ribosomes, 131f

Mitochondrial-ER MSCs, 529, 530f
Mitochondrion, 515
Mitogen, 52, 849, 869, 1170
Mitogen-activated protein kinase (MAPK) pathway, 849, 867
Mitophagy, 529
Mitosis, 21, 190, 230, 230f
    anaphase A, 833, 834f
    anaphase B, 833
    centrosome duplication, 825–826, 826f
    centrosome duplication and separation, 879, 880f
    chromosomal passenger complex, 832, 833f
    cohesin function during, 897f
    cytokinesis, 835–836
    duplicated chromosomes, alignment of, 831–832
    in epithelial cell, 847f
    in eukaryotic cell cycle, 850–852
    exit from, 886
    five stages of, 826–828, 827f, 828f
    human cells undergoing, 855f
    image of, 796f
    initiation of, 876, 877f
    lamins, dissasembly of, 842
    vs. meiosis, 894–895, 895f
    microtubule dynamics, 828–829, 830f
    microtubule reorganization in plant cells, 836–837, 836f
    mitotic spindles, 828, 829f, 834, 835f
    prometaphase, 829–831, 831f
    stages of, 851f

Mitosis-promoting factor, 855, 860
Mitotic apparatus, 879
Mitotic asters, 850
Mitotic budding, 24f
Mitotic CDKs, 842, 876–882, 878f, 886, 886f
Mitotic checkpoint complex (MCC), 832
Mitotic cyclin-CDK, 857, 859–860
Mitotic exit network, 886
Mitotic spindle
    assembly of, 826, 879–882
    centering and orientation of, 834, 835f
    chromosome, attachment to, 881f
    chromosome, segregation of, 880f
    definition of, 147
    formation of, 826f
    microtubules and, 828, 829f
    poles, separation of, 147f
    segregation of, 883f
    spindle assembly checkpoint pathway, 892–893, 893f
Mitral neurons, 1084
Mobile DNA elements. *See* Interspersed repeats
Model organisms, 22, 852–856
Moderately repeated DNA, 289
Modification enzyme, 238
Molarity, 56
Molecular cell biology, 1176–1177
Molecular chaperones, 90
Molecular complementarity, 35, 35f, 43–44, 43f, 96

Molecular genetic mechanisms
DNA, double-helical structure of, 179–185
DNA repair and recombination, 190–199
DNA replication, 185–190
mRNA, formation of, 199–205
overview of, 177–179, 179f
proteins, synthesis of on ribosomes, 210–219
transcription, 199–205
tRNA, 205–210
viruses, 219–224
Molecular handcuff, 718
Molecular machines, 71, 84, 84f
Molecular selectivity filter, 496–498, 498f
Molecular transformation, 71
Molecules, 6–12
in cells, 7f
macromolecules, quality control of, 12
nucleic acids, 8–11
phospholipids, 11–12
proteins, 7–8
Monastrol, 147, 147f
Mono-astral array, 147
Monocistronic mRNA, 280
Monoclonal antibodies, 144–146, 145f, 673–674, 1095, 1108, 1181
definition of, 153
medical uses, 146
organelle-specific antibodies, 174
Monogenic diseases, 253–255, 255f

Monomeric G protein, 669, 669f, 674
Monomers, 7, 44, 45f
Monopolin complex, 898
Monosaccharides, 49–51
Monotelic attachment, 881
Monoubiquitinylation, 113, 715
Morgan, Thomas Hunt, 31
Morphogen, 743
Morphogenesis, 908f
Morula, 985, 987f, 993
Mosaic development strategy, 991
Motile cells, 943–951
Motor neurons, 1034f, 1043f, 1044, 1046, 1047f
Motor proteins, 71, 754, 809–819
Mouse mammary tumor virus (MMTV), 1174f
Movement, motor proteins and, 71
MreB gene, 756
MRN complex, 890
mRNA surveillance, 426
mRNP exporter, 412, 618
mRNP remodeling, 412–413
mTOR protein kinase, 959
mTORC1, 959, 965–966, 965f, 968f
Muller, H., 232
Muller, Hermann J., 1158
Müllerian-inhibiting hormone (MIH), 733
Multi-adhesive matrix protein, 906t, 927, 932
Multicolor FISH, 313

Multidrug-resistance (MDR) transport protein, 489
Multi-hit model, 1172–1173, 1173f, 1174
Multimeric membrane proteins, 458f
Multipass membrane protein, 588
Multipass transmembrane protein, 457, 457f
Multiple endocrine neoplasia type 2, 1168
Multiple sclerosis (MS), 1062
Multipotent, 992
Multipotent somatic stem cells, 1003–1004
Multiprotein complexes, 534, 542f
    $CoQH_2$-cytochrome *c* reductase (Complex III), 542–543, 543f
    cytochrome *c* oxidase (Complex IV), 544
    electron-transport chain, 539, 540–544
    F_0 and F_1, 551–553, 552f
    NADH-CoQ reductase (Complex I), 540–541, 542f
    succinate-CoQ reductase (Complex II), 541–542, 542f
    supercomplexes, assembly into, 545–547, 546f
Multiubiquitinylation, 113
Multivesicular endosomes, 655, 656–657, 656f
Muscarinic acetylcholine receptors, 679
Muscle contraction, 692–697, 1074–1075, 1074f
Muscle relaxation, 482–483, 483f, 485f
Muscular dystrophies, 32–33, 405, 773, 948–949, 948f
Muscular tissue, 902
MuSK, 1067
Mutagen, 24–25, 228, 229, 1150
Mutations, 2, 190, 228
    causes of, 191

compensating, 392
conditional, 231–232, 233f
driver, 260
frameshift, 207, 249, 271f
functional consequences of, 249–250
gain-of-function, 1166f
genetic analysis of, 228–237
linkage mapping, 255–256
loss-of-function, 1166f
missense, 207, 249–250
in mitochondrial DNA, 526
nonsense, 207, 218, 249
oncogenic driver mutations, 1152–1153
passenger, 260
point, 191, 191f, 229, 249–250, 258
recessive lethal, 232–233
repair, 190–191
segregation of, 230–231
silent, 250
spontaneous, 229
synthetic lethal, 235–236
temperature-sensitive, 25, 232, 233f
transcription factors and, 325f
tRNA mutations, suppression of, 218
Mutually antagonistic complexes, 1016
Myasthenia gravis, 1074
MYC, 371
MYC gene, 1171–1172, 1171f, 1172f

Mycobacterium tuberculosis, 1096
Mycoplasma, 14
MyD88 adapter, 749, 749f
Myddosome, 1141f
Myelin basic protein (MBP), 1062
Myelin sheath, 1043f, 1045, 1047–1048, 1060, 1061–1063, 1061f, 1062f
Myelination, 1059–1060
Myelodysplastic syndromes (MDS), 407–408, 407f
Myeloma cells, 145, 145f
Myofibrils, 780, 781f, 782f
Myosin heavy chains, 774
Myosin I, 776–777, 776f
Myosin II, 773–774, 774f, 776–778, 776f, 779f, 783–784, 784f, 789
Myosin light-chain kinase, 784
Myosin regulatory light chain (LC), 784, 785f
Myosin S1, 758f
Myosin V, 776–777, 776f, 779, 779f, 785, 786f
Myosin VI, 777
Myosin-powered movements, 780–786
    actin and myosin II in nonmuscle cells, 783–784
    contraction in smooth muscle and nonmuscle cells, 784–785, 785f
    myosin thick filaments slide past actin thin filaments, 780–781, 781f
    myosin V carries vesicles, 785, 786f
    skeletal muscle contraction, regulation of, 782–783, 783f
    stabilizing and scaffolding proteins, 781–782, 782f
Myosins, 773–779

actin filaments, steps along, 777–779, 778f
ATP hydrolysis and movement, 777, 778f
classes of, 776f
functions of head, neck, and tail domains, 774–775
head domain, 814f
mechanochemical motor proteins, 775–777
movement, rate of, 779f
myosin II, structure of, 774f
sliding-filament assay and, 775f
superfamily in humans, 776f
Myotonic dystrophy type 1, 287, 296
Myristate anchor, 631

N

N lobe, 713, 714f
N nucleotides, 1132
Na^+ channels, 1052, 1053–1054, 1053f, 1057f, 1078f
Na^+ equilibrium potential (E_{NA}), 1052
Na^+-glucose symporter, 502f
Na^+/H^+ antiporter, 504
$Na^+HCO_3^-/Cl^-$ antiporter, 504
Na^+/K^+ ATPase, 473f, 484–486, 486f, 504
N-acetylgalactosamine, 48, 461, 461f
NAD^+ (nicotinamide adenine dinucleotide), 66–67, 67f, 102, 516, 517, 534
NADH, 534, 537, 544–545, 545f
NADH dehydrogenase-like complex-dependent pathway, 569

NADH-CoQ reductase (Complex I), 540–541, 542f
NADPH, 561–564, 568
Naive B cells, 1106–1107
Nanobodies, 146
Nanotubes, tunneling, 925, 926f
Nascent cell plate, 837
Nascent RNA chain, 202, 203
Native state, 82
Natural killer (NK) cells, 1100, 1101f, 1145f
Natural regulatory T cells, 1136
Neanderthals, 5, 6f
Nebulin, 781, 782
Neck domain, 774, 778f, 779f
Necroptosis, 1028, 1038f, 1039
Necrosis, 1027
Negative elongation factor (NELF), 339–340, 339f, 388
Negative feedback loops, 737–738, 979–981, 1013–1015
Negative selection, 1134–1135
Negative stain, 167–168
Nematodes, 1020–1022, 1020f, 1021f
Neo-antigens, 1146, 1179
Neoblasts, 32, 744, 990f, 1002, 1002f
Nernst equation, 494, 548–549, 1052
Nerve growth factor (NGF), 709, 1033, 1034f
Nervous system, 1042
    action potentials, propagation of, 1052–1064
        glia produce myelin sheaths, 1061–1063, 1061f, 1062f
        magnitude of, 1052

myelinated axons, 1060–1061, 1061f
unidirectionally without diminution, 1055, 1056f
environment, sensing, 1077–1086
mechanoreceptors, 1078–1079, 1078f, 1079f
odor, 1083–1084, 1083f
olfactory receptor neurons, 1084–1086, 1085f
pain receptors, 1079
primary tastes, 1080–1083
gray and white matter, 1060f
memories, forming and storing, 1086–1091
gene expression and, 1090–1091
hippocampus, 1088
long-term memories, 1088f
synapses, number and strength of, 1086–1088
synaptic plasticity, 1088–1090, 1089f
neurons and glia, 1044–1051
action potentials, 1045, 1045f
glial cells form myelin sheaths, 1047–1048
information flows from dendrites to axons, 1044–1045
neural stem cells, 1048–1051, 1050f
signaling circuits, 1046–1047, 1048f
synapses, 1045–1046, 1046f
overview of, 1042–1044, 1043f
synapses, communication at, 1065–1077
action potential, inputs for, 1076, 1076f
Ca^{2+} triggers release of neurotransmitters, 1070
gap junctions, 1076–1077
muscle contraction, 1074–1075, 1074f

neurotransmitters, degradation or reuptake of, 1073–1074
neurotransmitters and H^+-linked antiport proteins, 1068–1070
nicotinic acetylcholine receptor, 1075, 1075f
plasma membrane, fusion with, 1070–1073, 1072f
presynaptic and postsynaptic structure, 1065–1068
presynaptic terminal, 1070
synaptic vesicles, recycling of, 1073
voltage-gated ion channels, 1052–1064, 1052f
channel-inactivating segment, 1059, 1059f
light-activated ion channels, 1063–1064, 1063f
membrane depolarization, 1056–1059
myelination and velocity of impulse conduction, 1059–1060
Na^+ channels, 1052
sequential opening and closing, 1052–1055, 1053f
structures of, 1055–1056, 1057f, 1058f
voltage-sensing S4 α helices, 1056–1059
NETosis, 1102, 1102f
Neural circuits, 1064f, 1072f
Neural growth cones, 844, 844f
Neural stem cells (NSC), 1048
Neural tissue, 902
Neural tube, 1048, 1050f
Neurites, 1033
Neuroblasts, 1026f, 1044, 1050f
Neuroepithelial cells (NECs), 1048, 1050f
Neurofilament, 839, 840
Neurogenesis, 1051f

Neuromodulatory neurons, 1047
Neuromodulatory signals, 1044
Neuromuscular junction (NMJ), 1066, 1067–1068, 1067f, 1074f
Neuromuscular synapse, 1067
Neurons, 998, 1026f, 1032–1034, 1034f, 1044–1051
    definition of, 1042
    glial cells and, 1047–1048
    information flows from dendrites to axons, 1044–1045
    in mammals, 1043f
    signaling circuits, 1046–1047, 1048f
    synapses and, 1045–1046, 1046f
Neurotransmitters, 14, 1045
    cycling of, 1071f
    degradation or reuptake of, 1073–1074
    presynaptic terminal, 1070
    receptors and, 1046f
    structures of, 1069f
    synapses
        communication at, 1065
        diffusion across, 1046
Neurotrophins, 1032–1034, 1034f
Neutrophil, 1101, 1102, 1102f
Nexin, 820
Next generation sequencers, 246
NF1 gene, 250, 251f
NFX1/NXT1 complex, 618, 619f, 620
NF-κB, 741
    activation of, 740–750, 749f

transcription factor, 747–748, 748f
NF-κB signal transduction pathway, 741
Nick, 184, 184f
Nicotinamide adenine dinucleotide (NAD^+), 66–67, 67f, 102, 516, 517, 534
Nicotinamide adenine dinucleotide phosphate ($NADP^+$), 568
Nicotinic acetylcholine receptor, 1075, 1075f
Nidogen/entactin, 906t, 928, 928f
Nitotinic acetylcholine receptor, 1074
NKT cells, 1128
***N*-linked oligosaccharide**, 595–596, 596f, 597f, 601f, 640f, 936
N-lobe, 111, 112f
NMDA receptors, 272
NMR spectroscopy, 130
N-nucleotides, 1112
Nociception, 1079
Nociceptors, 1079
Nocodazole, 856
Node of Ranvier (node), 1043f, 1060
Nodes, 1060–1061, 1061f, 1063
No-go decay, 428
Nomarski interference microscopy, 150
Nonbonding electrons, 37
Noncoding DNA, chromosomal organization, 286–289
Noncovalent binding, protein regulation and, 107–110
Noncovalent interactions, 36–44
hydrogen bonds, 40–41, 41f
ionic bonds, 39–40, 40f

molecular complementarity, 43–44, 43f
strength of, 39, 39f
Nondisjunction, 893
Nongated channels, 473, 496
Nongated ion channels, 494–500
Nonhistone proteins, 311–312
Nonhomologous end joining, 194, 196f, 891
Non-insulin-dependent diabetes. *See* Type II diabetes
Non-ionic detergents, 462–463, 463f
Non-LTR retrotransposons, 295–297
Nonmotile cells, 943–951
Nonmuscle cells, contraction in, 783–785
Nonpermissive temperature, 25, 232, 233f
Non-photochemical quenching, 566
Nonpolar molecules, 37, 42–43
Nonprotein-coding genes, 284, 285f, 285t
Nonsense mutation, 207, 218, 249
Nonsense suppression, 218
Nonsense-mediated decay (NMD), 397, 426–428, 426f, 427f
Non-stop decay, 428
Non-TLR nucleic acid sensors, 1098
Nontranscribed genes, 306f
Nonviral retrotransposons, 295
Noonan syndrome, 109
Noradrenaline, 1065
Norepinephrine, 1065, 1073
Notch/Delta pathway, 738–740, 739f
Notochord, 1050f

Notum gene, 744
NPXY sorting signal, 653
N-region, 1112
NSF protein, 227–228, 636
N-terminal domain, 80
N-terminal globular domains, 1059f
N-terminal signal sequence, 580–581
N-terminal targeting sequence, 603–604
N-terminus, 73, 591–592
Nuclear basket, 412, 614
Nuclear bodies, 438–440, 439f
Nuclear cap-binding complex, 410
Nuclear domains, 438–440
Nuclear envelope, 412–416, 876–879, 878f, 887f
Nuclear exonucleases, 409–410
Nuclear export factor 1 (NSF1), 618, 620
Nuclear export sequence (NES), 162, 163f
Nuclear export transporter 1 (NXT1), 618, 620
Nuclear lamina, 16, 841, 842f, 877, 878f
Nuclear localization sequence (NLS), 162, 163f
Nuclear localization signals, 984
Nuclear lysine acetyl transferases (KATs), 306
Nuclear magnetic resonance (NMR) spectroscopy, 130
Nuclear mRNPs, 387, 412
Nuclear paraspeckles, 439
Nuclear pore complex (NPC), 16, 171f, 412, 415f, 435–436, 614–616, 615f
Nuclear pores, 412, 614

Nuclear receptor, 348, 368–370, 368f
Nuclear speckles, 21, 438–439, 439f
Nuclear SREBP (nSRABP), 971f, 972
Nuclear transcription factors, 1171–1172
Nuclear transport receptor, 616–618, 617f
Nuclear-export signals (NES), 618, 619f
Nuclear-localization signals (NLSs), 614, 616–618, 616f, 736–737, 736f
Nuclear-receptor superfamily, 368, 369f
Nucleation, 765f
Nucleation phase, 759, 759f
Nucleation promoting factor (NPF), 764, 791
Nucleic acids, 8–11, 44, 45f, 48–49, 49f
Nucleocapsid, 221, 222
Nucleoid, 12f, 14
Nucleolar organizers, 308
Nucleolus, 17, 21
Nucleoplasm, 17
Nucleoporins, 614, 879
Nucleoside monophosphates, 49
Nucleoside triphosphates, 49
Nucleosides, 49, 49t
Nucleosomes, 278, 300, 301, 301f, 302f, 303–304, 305f, 354
Nucleotide excision repairs, 193, 195f
Nucleotide-binding pocket, 667
Nucleotides, 8, 48, 48f, 49t
Nucleus, 576f
    functions of, 15–17
    membranes and, 446–447, 446f

proteins, transport into and out of, 614–620
receptors in, 664–665
structure of, 16f
targeting sequences, 578t
Null alleles, 249
Numerical aperture (NA), 150
Nurse, Paul, 25, 860
Nüsslein-Volhard, Christiane, 232

O

Obama, Barack, 1043
Objective lens, 149f, 150
Obligate aerobes, 518
Occipital lobe, 1043f, 1044
Occludin, 920, 921f
Occlusion, 504
Odor, 1083–1084, 1083f
Odorants, 1075, 1083
O-GlcNAc-threonine, 48f
Oil drop model, 77, 78f
Okazaki, Reiji, 187
Okazaki fragments, 187, 187f, 189
Olfaction, anatomy of, 1084f
Olfactory neurons, 997f, 1083f, 1084–1086, 1085f
Olfactory receptor neurons (ORNs), 1083–1084
Olfactory receptors, 1083
Oligodendrocyte intermediate precursor cells (oIPCs), 1050f

Oligodendrocytes, 998, 1043f, 1047, 1061, 1061f, 1062
Oligomerization sites, 86
Oligonucleotide sequences, 262
Oligopeptide, 73
Oligosaccaride side chains, 596–597
Oligosaccarides, 626–627, 936f
Oligosaccharyl transferase, 596
***O*-linked oligosaccharide**, 595, 936, 936f
On the Origin of Species (Darwin), 3f
Oncogenes, 223, 1160–1162, 1165–1166
Oncogenesis, 1149
Oncogenic drivers, 1162–1165, 1165f, 1177–1178
Oncogenic mutations, 1152–1153, 1169–1170, 1169f
Oncogenic transformation, 140
Oncogenic transgenes, 1174f
Onco-metabolites, 1154
Oncoproteins, 737
1,2-diacylglycerol (DAG), 693, 694f, 697, 724
Oocyte expression, 499–500, 499f
Oocytes, 854–855, 854f, 992, 1001f, 1004–1005, 1005f
Op18/stathmin, 808
Open complex, 337f, 338
Open pore, 1059
Open reading frames (ORFs), 207, 249
Opsonization, 768, 769f, 1108
Optical isomers, 37
Optical microscopes, 148f–149f
Optical section, 151

Optogenetics, 162, 1063–1064, 1063f, 1064f
ORC (origin recognition complex), 190
Orexin gene, 682
Orexin-A, 682
Orexin-B, 682
Organelle-like structures, 21
Organelles, 11, 15, 576f
    definition of, 139
    isolation of, 172–175
        antibodies, 174
        cell disruption, 173
        centrifugation, 173–174
        proteomics, 174–175
    transfer between animal cells, 925
Organizing center, 1014
Organoids, 143–144, 144f, 855, 944, 1007
Origin-recognition complex (ORC), 871, 872f
Origins, 186–187
Orphan receptors, 681
Orthologous sequences, 251
Osmosis, 478
Osmotic flow, 478
Osmotic gradient, 508
Osmotic pressure, 478, 478f
Osteoclasts, 509
Osteogenesis imperfecta, 935
Osteopetrosis, 510
Outer mitochondrial membrane, 16, 20, 604–605

Outer nuclear membrane, 16
Outer-arm dyneins, 820
Outer-membrane proteins, 609
Outside-in signaling, 909, 909f, 919
Oxa1, 607
Oxaloacetate, 533
Oxidases, 536
Oxidation, 19–20, 66–67
Oxidation potential, 67
Oxidation reactions, 66–67, 66f
Oxidative phosphorylation, 515, 516, 531
Oxyanion hole, 101, 101f
Oxygen, 107–108, 108f, 972–975, 973f
Oxygen-evolving complex, 567
Oxyntic cells, 509

P

P bodies, 21, 386f, 418, 428
P element, 292
P granules, 1021
P nucleotides, 1112f
P site, 213, 215, 215f
p16 tumor suppressor gene, 1170, 1171f
p38 kinases, 720
p53 gene, 892
p53 protein, 892, 1152, 1174, 1177–1178, 1178f
$p75^{NTR}$, 1033

Pain receptors, 1079
Palade, George, 625
Palindromic sequences, 238
Pancreatic islet cells, 960, 961f
Paneth cells, 1005, 1006f, 1007, 1008
Pannexin, 925
Papillae, 1081
Par complex, 1022–1023, 1022f, 1023f
Par proteins, 1020–1023, 1020f, 1021f, 1022f, 1025–1026, 1025f
Paracellular pathway, 922, 922f
Paracrine signaling, 663
Paralogous sequences, 251
Paranodes, 1063
Parasites, 25–27
Paratope, 1108
Parent (donor) organelles, 629
Parent DNA, 197
Parental types, 255, 256f
Parietal cells, 508–509, 509f
Parietal lobe, 1043f, 1044
Parkinson's disease, 95, 529, 999f
Paroxetine, 1073
Partition defective genes, 1021, 1021f
Passenger mutations, 260
Passive immunization, 1102
Passive spread, 1053
Passive transport, 473
Patch clamping, 498–500, 499f

Patched receptor, 745
Pathogens, 1095–1096
Patterning genes, 30
Pauling, Linus, 98
Pax6 gene, 31, 31f, 327, 328f
PC2 endoproteases, 647
PC3 endoproteases, 647
P-class pump, 480, 481f
PCNA (proliferating cell nuclear antigen), 188, 188f, 193
PCSK9 gene, 259–260
PDZ domain, 921
PEK (pancreatic eIF2 kinase), 424
Pelizaeus–Merzbacher disease, 1062
Pemphigus vulgaris, 918
Penicillin, 146
Pentose, 50
Peptide, 73
Peptide bonds, 44, 72–73, 73f, 87, 101f, 1122f
Peptide groups, 41f, 87f
Peptide mass fingerprint, 129
Peptide-loading complex, 1123
Peptidoglycan, 14, 50–51
Peptidyl-prolyl isomerases, 599
Perforin-mediated cells, 1137f
Perforins, 1108, 1137, 1137f, 1179–1180
Pericentriolar material, 800
Pericytes, 1048, 1049f
Peripheral membrane proteins, 442, 455

Peripheral nervous system, 1043, 1047, 1061–1062, 1061f, 1062f, 1065. *See also* Nervous system
Perlecan, 906t, 928, 928f, 931–932
Permissive temperature, 25, 232, 233f
Peroxisomal matrix, 612–613, 612f, 613f
Peroxisomal proteins, 612–613, 612f, 613f
Peroxisomal-targeting sequence 1 (PTS1), 612–613
Peroxisomes, 19–20, 536, 578t, 602
Perutz, Max, 129
Pervasive transcription, 410
Pex5, 612, 612f, 613
pH, 57–58, 57f, 102, 102f
PH domain, 725
Phage, 219, 221
Phagocyte, 1098
Phagocytosis, 19, 753, 768, 769f
Phalloidin, 769
Pharming, 560
Phase transition, 449
Phase-contrast microscopy, 148f–149f, 150–151, 151f
Phenotype, 229
Phenylalanine, 46
Phenylketonuria (PKU), 254t
Pheromones, 661
Philadelphia chromosome, 1162
Phosphatase, 110, 667–668, 668f
Phosphate ions, 59
Phosphate transporter, 557

Phosphatidylcholine, 52f, 448f
Phosphatidylinositol (PI), 693
Phosphatidylinositol 4,5-biphosphate, hydrolysis of, 693–694, 693f
Phosphatidylinositol-3 kinase (PI-3 kinase), 724–725, 725f
Phosphatidylinositol-3-phosphates, 724–725, 725f
Phosphatidylserine, 453, 1032, 1033f
Phosphoanhydride bonds, 64–65
Phosphodiester, 65
Phosphodiester bonds, 44, 180, 180f, 183f
Phosphodiesterase (PDE), 690f, 694
Phosphofructokinase-1, 517, 517f
Phosphofructokinase-2, 518
Phosphoglucomutase, 685
Phosphoglycerides, 11f, 45, 51–54, 52f, 53t, 447
    in biomembranes, 447–449, 448f
    fatty acids in, 52t
    functions of, 11–12
Phosphoinositide signal transduction pathway, 724–726
    activated protein kinase B, 725–726, 726f
    phosphatidylinositol-3-phosphates, accumulation of, 724–725, 725f
    phospholipase C_γ, 724
    PI-3 kinase pathway, 726
Phosphoinositides, 447, 724
Phospholipase, 453, 453f
Phospholipase A_2, 462f
Phospholipase C (PLC), 693
Phospholipase C_γ, 724

Phospholipid anchors, 592
Phospholipid bilayer, 443f, 444, 444f
formation of, 444–445, 445f
function of, 445–447
gel-like and fluidlike forms of, 450f
permeability of, 471f
Phospholipids, 444–447, 445f
ABC proteins and, 491, 492f
annular, 458, 458f
bHLH transcription factor, 971–972
movement and flippases, 465–466
synthesis and intracellular movement, 464–468
transport of, 467–468, 468f
Phosphoprotein phosphatase (PP), 687
Phosphoproteins, 110
Phosphoproteomics, 135–136
Phosphoric acid, 59, 59f
Phosphorylation, 512, 668f
activating and inhibitory, 861–863, 863f
proteasome regulation, 106
protein kinase mediated, 110–111
protein regulation and, 110, 110f
Phosphoserine, 48f
Phosphothreonine, 48f
Phosphotyrosine, 48f, 713–714, 716f
Phosphotyrosine phosphatases, 731–732
Photoactivated localization microscopy (PALM), 164f, 165
Photoelectron transport, 564–566, 567, 567f

Photoinhibition, 566
Photomotility, 1063
Photons, 563
Photophobia, 1063
Photorespiration, 572–573, 572f
Photosynthesis, 7, 8f, 61, 513
ADP and, 65–66
carbon pathway, 571f
chloroplasts and, 559–566
thylakoid membranes, 559–560
core antenna complexes, 564
light absorption, 561
light-harvesting complexes, 564, 565f
overview of, 513–514, 513f, 562f
photoelectron transport, 567f
rate of, 563f
reactive oxygen species, 564–566
stage 1, 566–567
stage 2, 568–569, 568f
stage 3, 569
stage 4, 570–573
stages 1 and 2, 561–564
thylakoid membranes, 561
Photosystem I (PSI), 561, 564, 568, 568f, 569
Photosystem II (PSII), 561, 562f, 564, 566–567, 569
Photosystems, 561, 565f, 567f, 569
Phototaxis, 1063
Phragmoplast, 837

Physical environment, sensing and responding to, 982–987
Physical maps, 257f
Physiologic state, 86
PI 3-phosphates, 725
PI-3 kinase pathway, 725, 726, 726f
Piezo channels, 1079f
Piezo1, 1078, 1079f
Piezo2, 1078, 1079f
Pigment epithelial cells, 699
Pioneer transcription factors, 359
Pisum sativum, 565f
PKB kinase, 963, 968–969, 968f
Plague, 109–110
Plakin, 843
Plakoglobin, 918f
Plakophilins, 918f
Planar cell polarity pathway, 1023–1025, 1024f
Planar cell polarity (PCP), 1023
Planar peptide bonds, 87, 87f
Planaria, 32
Planarians, 705f, 743f, 1002–1003, 1002f
Plants. *See also* Photosynthesis
    cells, integrating into tissues, 951–956
        cell wall, loosening of, 953
        cell wall, structure of, 952–953, 952f
        compared to animals, 955–956
        plant adhesion molecules, 955–956
        plasmodesmata, 953–955, 954f

mechanosensation in, 956
photosystems in, 565f
Plaques, 94
Plasma cells, 1115
Plasma membrane, 442, 623f, 695–696, 696f, 1052f, 1070–1073, 1072f
Plasmalogens, 448f, 449
Plasmid expression vectors, 265–267
Plasmids, 14, 239, 239f, 240f, 242, 257f, 264f, 299
Plasmodesmata, 27, 913, 913t, 923, 925, 953–955, 954f
Plasmodium falciparum, 25
Plasmodium spp., 26f, 524
Plastoca, 567f
Plastocyanin, 561, 562f, 568
Platelet-activating factor (PAF), 950, 950f
Platelet-derived growth factor (PDGF), 675f, 709, 790
Pleated sheets, 75
Plectin, 843, 843f
Pluripotency, 994–996, 996f
Pluripotent, 991
Pluripotent cells, 144f, 993–994
Pluripotent stem cells, 143–144, 990f, 991, 1001f, 1002–1003
Podosomes, 943
Point mutations, 191, 191f, 229, 249–250, 258, 1164, 1168
Point source fluorescent objects, 162–163
Point spread function, 156
Pointed end, 757
Point-scanning confocal microscope, 156, 157f
Pol I, 329, 330t, 380–382

Pol II, 329, 330f, 330t, 331, 359–361, 380–382
Pol III, 329, 330t, 380–382
Polar bonds, 38
Polar microtubules, 828, 829f
Polar molecules, 42–43
Polarity, 38–39, 757–758, 758f, 799, 992
Polarity regulators, 911
Poliovirus, 1145–1146
Polo-like kinase, 834, 865, 875, 879–882, 888
Poly(A) polymerase, 203, 203f
Poly(A) sites, 279
Poly(A) tail, 203, 217, 242, 398, 419, 423f
Poly(A)-binding protein (PABPC), 217, 399
Polyacrylamide gel electrophoresis (PAGE), 118
Polycistronic mRNA, 279
Polyclonal antibodies, 145, 1106
Polycomb protein complexes, 309, 375–377, 376f, 377f, 996
Polyglutamylation, 818–819, 819f
Polyglycylation, 819, 819f
Poly-K63 ubiquitin-binding domain, 749
Polyketides, 83
Polylinker, 239, 239f, 241f
Polymerase chain reaction (PCR), 238, 242–246, 244f, 245f, 246f, 269f, 288
Polymerization, 185, 200f, 762
    of actin, 758–759, 759f, 766–767, 767f
Polymers, 7, 44, 45f, 178
Polymorphisms, 255–257, 258, 288, 1120

Polyp, 1175f
Polypeptide, 70, 73, 73f
Polyploid cells, 229
Polyribosomes, 217–218
Polysaccharides, 44, 45f, 49–51, 952–953, 952f
Polysomes, 217–218
Polytene chromosome, 316–318
Polytenization, 318
Polyubiquitin chain, 749–750
Polyubiquitin tail, 106, 732
Polyubiquitinylation, 105f, 113, 114f
Polyunsaturated fatty acids, 52
Porins, 458–459, 459f, 523
Positive feedback loops, 1016, 1017f
Positive selection, 1134–1135
Postmitotic cells, 321, 852
Postsynaptic cells, 1046, 1046f
Postsynaptic density (PSD), 1066
Post-transcriptional gene control
    cytoplasmic mechanisms of, 417–430
        mRNA, concentration of, 417
        mRNA, degradation of, 417–422, 418f, 425f
        mRNA, localization of, 428–430
    definition of, 385
    eukaryotic pre-mRNA, processing of, 387–399
    mRNA, transport of across nuclear envelope, 412–416
    nuclear bodies, 438–440
    overview of, 384–385, 386f

pre-mRNA processing, 400–411
rRNA, processing of, 431–438
tRNA, processing of, 431–438
Post-translational modification (PTM), 110, 307f
Post-translational translocation, 585–586, 585f
Potential energy, 60
Power stroke, 777, 778f
Pre-B-cell receptor, 1113–1115, 1114f
Precursor cells, 1003
Precursor mRNA (pre-mRNA), 179f, 202
Preinitiation complex (PIC), 336, 337f, 365
Premature termination codon (PTC), 427f
Pre-mRNA, 384, 389
definition of, 385t
as pre-mRNPs, 387
processing of, 387–399, 387f
3′ cleavage/polyadenylation, 398–399, 398f
5′ cap added, 387–388, 388f
chain elongation, 388–389, 388f
RNA splicing, 391–398, 391f, 392f, 393f, 394f, 395f, 396f
RNA-binding domains and, 389–391, 390f
spliceosomes, 393–398, 395f
splicing cycle, 395–397, 396f
transesterification, 391–393, 392f
regulation of, 400–411
abnormal RNA splicing, 405–408, 407f
Dscam isoforms, expression of, 404–405, 405f
K^+-channel protein isoforms, 402–403

nuclear exonucleases and the exosome, 409–410, 409f
pervasive transcription, 410
RNA editing, 411
RNA splicing, 403–404, 403f
self-splicing group II introns, 408–409, 408f
splice-site selection, 400–402
spliceosomes and, 414–415
splicing, 250
splicing mechanisms, 436f
Pre-mRNPs, 387
Prenylation, 459
Preprophase band, 837
Pre-rRNA, 385t, 388, 432
processing of, 432–436, 433f, 434f
transcription units, structure of, 432f
Presenilin 1 (PS1), 740
Presynaptic cells, 1045, 1046f
Presynaptic terminal, 1070
Pre-tRNA, 385t, 437–438, 437f
Pervasive transcription, 410
Prickle gene, 1025
Primary cell cultures, 140
Primary cell walls, 953
Primary cells, 140
Primary cilium, 746–747, 823, 824
Primary lymphoid organs, 1096
Primary structure, 72–73, 72f, 128–129
Primary tastes, 1080–1083

Primary transcript, 280
Primase, 187
Primates, evolution of chromosomes, 316f
Primer, 185–186, 186f
Primer-binding site (PBS), 294f–295f
Pri-miRNAs, 420, 420f
Primitive endoderm (PE), 994f
Primordial germ cell-like cells (PGCLCs), 1001f
Prion diseases, 95
Proalbumin, 647f
Processed pseudogenes, 298
Procollagen, 933
Prodomain, 733
Product-ion spectrum, 127f
Proenzymes, 647
Professional APCs, 1120
Profilin, 761–762
Progenitor (precursor) cell, 1003
Progeny virions, 221–222, 223f
Progesterone, 854f
Programmed cell death, 31, 992. *See also* Apoptosis
    immune system and, 1179
    oncogenic driver mutations, 1177–1178
    p53 protein, 1178–1179
    tumor microenvironment and immunoediting, 1179–1181
Proinsulin, 647f
Projection domains, 807f
Projection lens, 149f, 150

Projection neurons, 1084
Prokaryotes, 249, 442
Prokaryotic cells, 12–14, 12f
Prokaryotic organisms, 5
Prolactin, 727
Proline, 47, 75f, 974
Proline hydroxylase (PHD), 973
Proline isomerases, 88, 88f
Prometaphase, 827f, 828, 829–831, 831f
Promoter, 201, 279, 326f, 327, 333–336, 341, 730
Promoter mutation, 1164
Promoter region, 201f
Promoter-fusion, 267, 267f
Promoter-proximal element, 340–341
Promoter-proximal region, 339–340
Promoters, 279
Promyelocytic leukemia (PML) nuclear bodies, 439–440
Proofreading, 185–186, 186f, 210
Propeptides, 933f
Prophase, 826, 827f
Pro-proliferation proteins, 106
Proproteins, 647, 647f
Prosthetic group, 102, 539
Pro-survival proteins, 106
Protease, 79
Proteasomes, 104–107, 105f
Protein 0 (P_0), 1062
Protein coat, 622f, 629, 632f, 633

Protein correlation profiling, 135f
Protein disulfide isomerase (PDI), 597, 598f
Protein expression profiling, 133
Protein family, 81, 81f, 282
Protein folding, 87–95, 89f
    amino acid sequences, 88–89
    amyloids and disease, 94–95
    planar peptide bonds, 87
    proline isomerases, 88, 88f
Protein kinase, 718–720
Protein kinase A (PKA), 85f, 683–688, 685f, 686f, 687f, 690f
Protein kinase B (PKB), 724, 725–726, 726f
Protein kinase C (PKC), 697, 724
Protein kinase RNA activated (PKR), 424
Protein kinases (PK), 81f, 667–668, 668f
    mediation, 110–111
    regulation of, 111–113
    structure of, 112f
    study of, 673
    substrate binding, 112f, 113f
Protein phosphatase, 667
Protein self-splicing, 115
Protein sorting, 576–577
Protein synthesis, 423–424, 965–966
Protein targeting, 576–620
    integral membrane proteins, synthesis of, 587–588
    membrane proteins, insertion into ER, 586–594

mitochondria and chloroplasts, targeting proteins to, 602–611, 603f
amphipathic *N*-terminal targeting sequences, 603–604, 604f
chloroplast stromal proteins, import of, 609–610, 610f
energy inputs, 605–606
mitochondrial protein import, features of, 605
outer-membrane receptors and translocons, 604–605
submitochondrial compartments, 606–609, 607f, 608f, 609f
thylakoids, 610–611
modifications, folding, and quality control in ER, 595–602
chaperones and, 598–600, 599f
disulfide bonds (—S—S—), 597–598, 598f
improperly folded proteins, 600–601, 600f
N-linked oligosaccharide, 595–596, 596f, 597f, 601f
oligosaccharide side chains, 596–597
unassembled or misfolded proteins, 601–602
nucleus, transport into and out of, 614–620
mechanism for, 617f
mRNAs and Ran-independent mechanism, 618–620, 619f
nuclear pore complexes, 614–616, 615f
nuclear transport receptors, 616–618, 617f
overview of, 576–579, 577f
peroxisomal proteins, targeting of, 612–613, 612f, 613f
targeting to and across ER membrane, 579–586
ATP hydrolysis powers post-translational translocation, 585–586, 585f
cotranslational translocation, 581–582, 581f, 583f
hydrophobic *N*-terminal signal sequence, 580–581

membrane proteins, topology of, 593–594
passage through translocon driven by translation, 582–585
phospholipid anchors, 592
pulse-chase experiments, 579–580
topology of single-pass proteins, 588–591, 589f, 590f
type IV (multipass) proteins, 591–592
Protein-coding genes, 199
regulation of, 340–353
activation and repression domains, 349–350
activators, 345
distant enhancers, 341–342
DNA-binding domains, 347–349
DNase I footprinting, 343–345
electrophoretic mobility shift assay (EMSA), 343–345
enhancers, 352
promoter-proximal elements, 340–341
repressors, 346–347
transcription factors and, 350–352
transcription-control elements, 342–343
Protein-directed metalloribozyme, 397
Protein-folding catalysts, 600–601, 600f
Protein-fusion, 267, 267f
Proteins
binding, 95–104
enzymes and, 97–99
functions and, 95–97
biomolecular condensates, 84–86
chaperones and, 89–94, 91f

concentrations of, 324f
evolution of, 80–82
functions of, 7–8, 72f, 80–82
    binding and, 95–97
GPI-anchored, 593f
interacting and redundant, 235–236, 236f
introduction to, 70–72
lateral mobility in biomembranes, 449–450, 451f
membranes, interaction with, 455
models of, 9f
movement of, 631t, 637–638, 637f
purification, detection, and characterization of, 116–132
    antibody assays, 122
    centrifugation, 116–117
    chemical methods, 128–129
    conformation, determination of, 129–130
    electrophoresis, 117–119
    enzyme reactions, 121
    gene sequences, 128–129
    green fluorescent proteins, 122
    immunoprecipitation, 122–124
    liquid chromatography, 119–121
    mass spectrometry, 126–128
    radioisotopes and, 124–126
purification of, 121–124
regulation of, 104–116
    allosteric, 107–108
    cooperative, 107–108

covalent modification of, 110
location, 115
noncovalent binding and, 107–110
proteasome, 104–107
proteolytic cleavage, 114–115
synthesis and degradation, 104–107
ubiquitin and, 106–107
ubiquitinylation and deubiquitinylation, 113–114
secretion regulation, 692–697
separation of, 116–121
sequence comparisons, 80–82, 81f
single-pass, 588–591, 589f, 590f
structure of, 72–87, 72f
categories of, 82–83
compared to sequences, 80–82
conformations, 78–79
domains, 79–80
motifs, 76–77
primary structure, 72–73, 128–129
quaternary structure, 80f, 83–87
secondary structure, 74–77, 74f, 75f, 76f
tertiary structure, 77, 80f
supramolecular complexes, 72f, 84
synthesis of, 577f
synthesis of on ribosomes, 210–219
chain elongation, 213–216
eukaryotic translation initiation, 213
methionyl-tRNA, 212–213

mutations, 218
polysomes and rapid ribosome recycling, 217–218
translation, termination of, 216–217
TRPase-superfamily proteins, 218
tagging, 267
tail-anchored, 590–591, 591f
ubiquitin-mediated degradation, 863–865, 864f
Protein-sorting pathways, 577f
Proteoglycans, 460, 735, 906t, 927, 931–932, 934f, 935–937, 935f, 938f
Proteolipid protein (PLP), 1062
Proteolysis, 105f, 1123, 1128
Proteolytic cleavage, 114–115
Proteolytic processing, 647, 647f
Proteome, 71, 132
Proteomics, 107, 132–136, 174–175
Proteostasis, 104
Protocadherins, 916f–917f
Protofibril, 838
Protofilaments, 797, 798f, 799f, 805, 838
Proton pump, 561
Proton pumping, 537–538, 548
Proton transport, 538f
Protonated probes, 153f
Proton-motive force, 514, 514f, 531, 538
ATP synthesis, 550–558
mechanism of, 550–551
ATP-ADP exchange and phosphate transport, 557, 557f
brown fat and heat generation, 558

F_0 c ring rotation, 555
F_1 γ subunit, rotation of, 553–554, 554f
generation of, 537–549
NADH and $FADH_2$, oxidation of, 537
generation of in photosynthesis, 568–569, 568f
mitochondrial oxidation, rate of, 558
mitochondrial protein import and, 606
multiple protons and, 554–555
multiprotein complexes, 551–553, 552f
in photosynthesis, 561–564
voltage gradient and, 548–549
Proton-motive force, 514
Protons, 57, 556f
Proton/sucrose antiporter, 506
Proto-oncogene, 1161, 1169f, 1171, 1171f
Protostomes, 30
Provirus, 223
Pro-α chains, 933–934
Pseudogene, 283, 298
Pseudokinase domain, 730
Pseudosubstrate, 685
PTEN phosphatase, 726
PTS2, 613
Pull-down assays, 674, 675f
Pulmonary ionocyte, 33
Pulse-chase experiments, 125, 125f, 579–580, 651f, 941, 1007f
Pumps. *See* ATP-powered pumps
Purines, 49, 180

P-values, 250
Pyrimidines, 49, 180
Pyroptosis, 1141
Pyruvate, 516–517, 519f, 531–532
Pyruvate dehydrogenase, 531
Pyruvate kinase, 517

Q

Q cycle, 543–544, 543f
Q_i site, 543–544
Quaternary structure, 72f, 80f, 83–87
Query sequence, 250
Quiescent center, 1015
Quinone (Q), 567, 567f

R

R factors, 299–300
R group, 45
Rab effectors, 633
Rab GTPases, 633–635, 634f
Rab proteins, 633, 634, 634f, 635t
Rac protein, 675f, 790–792, 791f, 793f
Rad3-related protein, 889
Radial glial cells (RGCs), 1048, 1050f
Radial spokes, 820
Radiant energy, 60, 61

Radicals, 39
Radioisotopes, 124–126, 124t
Raf gene, 720
RAG proteins, 1110, 1111f, 1112f
RAG1/RAG2, 1111f
Ramón y Cajal, Santiago, 1086–1087, 1087f
Ran, 616
Random coil, 74
Rapamycin, 958f, 965
Ras, 706f–707f, 717–720, 719f
Ras pathway, 1169
Ras protein, 717, 719f, 1152
Ras/MAP kinase pathway, 716–723, 719f
Rate constant, 55
Rate-limiting step, 111
Rate-zonal centrifugation, 116, 117
RB gene, 1163, 1165
Rb protein, 1163
RBR protein, 1014–1015
RDG motif, 939–940
Reaction centers, 561
Reactive oxygen species (ROS), 547–548, 547f, 564–566
Readers, 110
Reading frame, 206, 207f
Receptor hetero-dimerization, 728
Receptor homodimerization, 728
Receptor mutations, 1168–1169, 1169f
Receptor protein, 442

Receptor serine kinases, 733–738
Receptor tyrosine kinases (RTKs), 706, 709–715, 1168
dimerization and activation, 709–711
phospholipase Cγ, activation of, 724
Ras and, 717–718
signal transduction after activation of, 713–714, 716f
signal transduction pathways, 706f–707f
structure of, 709f
types of, 713f
Receptor-ligand complexes, 653–654, 653f, 683
Receptor-mediated endocytosis, 649–654, 768
lipoprotein complexes, formation of, 650–651, 650f, 652f
macromolecular ligands, receptors for, 651–653
receptor-ligand complexes, dissociation of, 653–654, 653f
signaling from, 714–715
signaling from RTKs, 714–715
signaling receptors, down-regulation of, 654
Receptors, 56, 662, 670–671, 671f. *See also* Cell-surface receptors; Signal transduction pathway
Receptor-signal complex, 665f
Recessive, 229, 229f, 230–231, 231f, 234f
RECK (reversion-inducing–cysteine-rich protein with kazal motifs), 942
Recognition helix, 347
Recombinant cells, 477–478
Recombinant chromosomes, 197
Recombinant DNA, 237, 264
Recombinant types, 255

Recombination, 195, 255, 256f, 298f
Recombination nodules, 895
Recombination signal sequences (RSS), 1110, 1111f
Red blood cells, 728f
Redox reactions, 66–67, 67f
Reduction potential (*E*), 67, 544–545, 545f
Reduction reactions, 66–67
Reeves muntjac, 313, 314f
Reflex arc, 1046–1047
Reflexes, 1047, 1047f
Refractive index, 150
Refractory period, 1054
Regulated intramembrane proteolysis, 601, 970, 971
Regulated secretion, 625
Regulators of G protein signaling (RGSs), 669, 680
Regulatory light chain, 775
Regulatory particle (RP), 105f, 106
Regulatory proteins, 71
Rehydration therapy, 508
Release, 221
Release factors (RFs), 216
Remodeling, 906
Remodeling of bone, 509–510
Renaturation, 183–184
Repeats, 906
Repertoire, 1109
Repetitive DNA, 277–278, 286
Replication fork, 186–189, 188f, 195–197

Replication fork collapse, 195–197, 197f
Replication licensing, 871
Replication origin, 186–187, 239, 239f, 318
Replicative helicase, 187
Replicative senescence, 855
Replicon, 871
Repolarization, 1045
Reporter gene, 327
Repressed genes, 373, 374f
Repression domain, 347, 349–350
Repressor, 327, 346–347, 347f, 355–358
Rescue, 803, 804f
Reserve transcriptase, 222
Residues, 45, 73
Resolution, 150, 197–198, 199f
Resonance energy transfer, 564
Resonance hybrids, 39
Respiration, 65–66, 515
Respiratory chain, 537
Respiratory control, 558
Response elements, 369, 370f
Resting K^+ channel, 496, 498f
Resting membrane potential, 494–500, 497f
Restriction elements, 1118, 1120, 1122f
Restriction enzyme, 238f, 243f
Restriction fragments, 238, 239f
Restriction point, 848, 848f, 849, 1170
Restriction sites, 238

Restrictive enzymes, 238
Reticulons, 452
Retinitis pigmentosa, 109, 258
Retinoblastoma, 1163, 1163f
Retinoic acid response element (RARE), 369
Retrograde transport, 625, 638–639, 640f, 810
Retromer, 645
Retroposed RNA, 297–298
Retrotransposons, 223, 290, 290f, 292–295, 292f, 417
Retroviral genomic RNA, 293f, 294f–295f
Retroviral vectors, 266f
Retrovirus, 222, 223–224, 223f, 266, 656–657, 657f, 1161
Retrovirus-like elements, 293
Reverse transcriptase, 222, 223f, 224, 242, 289, 293
Reverse transcriptase-PCR (RT-PCT), 245
Reverse transcription, 293, 294f–295f, 296, 297f
Reversible phosphorylation, 414f
Reversibly arresting, 856
RFC (replication factor C), 188
RGD motif, 919, 940f
Rhabdomyosarcomas, 747
Rhinoviruses (RV), 917–918
Rho kinase, 785
Rho protein, 790–792, 790f, 791f, 793f
RhoA, 887–888, 887f
Rho-binding domain (RBD), 764f
Rhodopsin, 698–701, 698f, 699f, 701f
Rhodopsin kinase, 701

Ribbon diagram, 78f, 79
Ribonucleic acid (RNA), 10–11, 48, 183f, 205f, 386f. *See also* Post-transcriptional gene control
Ribonucleoprotein (RNP) complexes, 387
Ribonucleoside triphosphate monomers (rNTPs), 179f, 200, 200f
Ribose, 48f
Ribosomal RNA (rRNA), 178, 179f
    processing of, 431–438
    roles of in protein synthesis, 205, 205f
    subunit assembly, 435f
    synthesis of, 966
Ribosomal subunit synthesis, 440
Ribosome, 10, 14, 84, 87, 178
    antibiotics and, 211–212
    components of, 212t
    composition of, 179f
    protein synthesis, 210–219
        chain elongation, 213–216
        eukaryotic translation initiation, 213
        methionyl-tRNA, 212–213
        mutations, 218
        polysomes and rapid ribosome recycling, 217–218
        role in, 210–212
        translation, termination of, 216–217
        TRPase-superfamily proteins, 218
    rapid recycling of, 217–218
    structure of, 211f, 212f
Ribozymes, 11, 97, 111, 436

RIM proteins, 1073
RIM-BP proteins, 1073
Ringer, S., 108
RISC, 272, 274
Risk factors, genetic, 258
RNA editing, 411, 411f
RNA export factor (REF), 397
RNA interference (RNAi), 102, 272–274, 275f, 417, 421–422, 744
RNA intermediates, 295f
RNA knockdown, 385
RNA polymerase
    classes of, 330t
    in DNA transcription, 200–202, 324
    in eukaryotic genomes, 329–331
    functions of, 10, 11f, 179f, 200, 200f
    identifying, 329–331, 329f, 330t
    structure of, 202, 202f, 330f, 331, 331f
    in supramolecular complexes, 84, 84f
    in synthesis of mRNA, 323
RNA polymerase I, 329, 330t, 380–381, 380f
RNA polymerase II, 323, 326f, 329, 330t
    carboxy-terminal domain of, 332f
    clamp domain and, 331–332, 332f
    distant enhancers and, 341–342
    general transcription factors, 336–339, 337f
    metazoans and, 370–371
    preinitiation complex, 337f
    in pre-mRNA processing, 388–389, 388f

promoters, TATA box, initiators, and CpG islands, 333–336
structure of, 330f, 331
transcription in eukaryotic genomes, 331–333
transcription initiation, 333
transcription termination, 399
RNA polymerase III, 329, 330t, 381–382, 381f
RNA processing, 179f, 202, 203f
RNA processing factors, 388–389
RNA recognition motif (RRM), 389–391
RNA splicing, 391–398, 391f–396f
definition of, 203
disease and, 405–408, 406f, 407f
in processing of pre-mRNA, 387
regulation of, 403–404, 403f
RNA world hypothesis, 11
RNA-binding domains, 389–391, 390f, 401
RNA-induced silencing complex (RISC), 420
RNA-Seq, 263
Rod cells, 698–699, 698f, 699f, 701f, 702–703, 702f, 1077
Rods, 698
Root apical meristem, 1013
Root meristem, 1015
Roots, 952
Rough endoplasmic reticulum (Rough ER), 18, 18f, 579, 580f, 595–596, 624–625, 624f
Rous, Peyton, 1161
Rous sarcoma virus, 714, 1161
RPA (replication protein A), 188, 189

RPE65 gene, 699
RRM domain, 390f
RS domains, 401
R-Smad, 735, 737
R-spondins, 1007–1009
RTK/Ras/MAP kinase signaling cascade, 1169–1170, 1169f
Rubisco (ribulose 1,5-biphosphate carboxylase), 570–572, 572f
Rubisco activase, 570
Ryanodine receptors, 694

S

S (synthesis) phase, 21, 825, 848f, 849, 850f, 869–871, 869f, 870f
S phase cyclin-CDKs, 857
Saccharomyces cerevisiae, 22, 24, 24f
    alleles, segregation of, 232f
    budding, 853f
    cell cycles in, 852
    centromere sequences, 318–319, 320f
    conditional mutations, 232, 233f
    double mutants in, 236, 237f
    experiments with, 231
    gene interaction network, 237f
    genome, compared to human, 251f
    haploid/diploid states, 229
    homologous recombination, 268–269, 269f
    MAP kinase cascades, 723f
    meiosis in, 894

mRNA, localization of, 428–430, 429f
myosin V, 785, 786f
pre-mRNA splicing, 394f
RNA polymerase and, 330, 330f
S phase, onset of, 870f
Saccharomyces pombe, 22, 24, 24f, 318–319, 853f
SAGA complex, 358
SALL1 gene, 327–328, 329f, 342
Saltatory conduction, 1061
Salty taste, 1082
SAMTOR, 967f, 968
Sanger, Frederick, 245
Sar1 protein, 631, 632, 632f, 637f
Sarcoglycan complex, 32, 949
Sarcolemma, 782
Sarcoma, 1149
Sarcomere, 780, 781f
Sarcoplasmic reticulum (SR), 480, 482–483, 694, 782, 782f
Sarcospan, 949
Saturated, 464
Saturated fatty acids, 52, 464
Saturation curve, 670
Scaffold proteins, 71, 722, 723f, 1068f
Scaffolds, 103
Scanning, 213
Scanning electron microscopy (SEM), 167, 167f, 170–171, 172f
scaRNAs, 438
Scatter factor, 793

Scavenger molecules, 566
Scavenger receptors, 1098
SCF, 864
Schizosaccharomyces pombe, 852, 876
Schleiden, Matthias, 138, 847
Schmidtea mediterranea, 32, 743, 743f
Schwann, Theodore, 138, 847
Schwann cells, 949, 1043f, 1047, 1061, 1061f, 1062, 1065
Scribble complex, 1023, 1026
Scurvy, 934
SC-β cells, 999, 1000f
SDS-polyacrylamide gel electrophoresis (SDS-PAGE), 117–118, 118f, 119f
Seams, 799
Search and capture, 805
sec18 mutant, 227–228
Sec61 complex, 583, 584f, 585
Sec61α, 583, 584, 584f, 588, 589
Sec61β, 583
Sec61γ, 583
Sec63 complex, 585f, 586
Second law of thermodynamics, 62
Second messenger, 665f, 667, 667f, 692–697, 693f, 697f, 724, 923
Secondary active transport, 473
Secondary cell wall, 953
Secondary lymphoid organs, 1096
Secondary structure, 72f, 74–77, 74f, 75f, 76f
***Secretion (sec)* mutants**, 628

Secretory granules, 625
Secretory pathway, 577, 577f, 579, 622, 623, 624f
early stages of, 636–641
cisternal maturation, 639–641, 640f, 641f
COPI vesicles, 638–639
COPII vesicles, 637–638
first stage of, 624–625, 624f
later stages of, 642–649
apical and basolateral regions of polarized cells, 648–649, 648f
dynamin and, 643–644, 644f
lysosomal storage diseases, 645
mannose 6-phosphate residues, 644–645, 645f
protein aggregation in *trans*-Golgi, 645–647, 646f
proteolytic processing, 647, 647f
trans-Golgi, transport from, 642–643, 642f
study techniques, 625–629
cell-free transport assays, 628–629, 629f
protein transport, 625–627, 626f, 627f
yeast mutants and, 627–628, 628f
Secretory proteins, 579–581, 580f, 585–586
Secretory vesicles, 645–647
Securin, 884–886, 886f
SecY complex, 584f
Sediment, 116
Seed sequence, 419–420
Segmental duplication, 283–284
Segregation, 230–231, 231f, 232f

Selectable marker, 239
Selectin, 904, 904f, 950, 1139
Selective serotonin reuptake inhibitors (SSRIs), 1073
Selectivity filter, 497, 498f
Self-splicing group I introns, 436–437, 436f
Self-splicing group II introns, 408–409, 408f
Semiquinone, 540f
Sensitization, 1088
Sensors, 889, 958
Sensory domain, 160, 162
Sensory neurons, 1034f, 1044, 1046
Separase, 884–886, 886f
Septum, 851
Sequence drift, 283
Sequence maps, 257f
Sequence motifs, 77
Sequence-reading helix, 347
Sequence-specific DNA affinity chromatography, 343
Sequence-specific RNA-binding proteins, 424–426
Sequencing technology, 245–246, 246f, 247f
Serial sections, 169
Serine, 47
Serine proteases, 99–102, 100f, 101f
Serotonin, 1073
Sertraline, 1073
Serum response element (SRE), 720
Serum response factor (SRF), 721
Sessile, 902

SET domain, 375
Severe combined immunodeficiency (SCID), 730
Severe myoclonic epilepsy of infancy, 1055
Severing proteins, 809
Sex-lethal genes, 403
SH2 domains, 713–714, 715f, 716f, 718, 730
SH3 domain, 718, 718f
Shaker K^+ channel, 1055, 1057, 1059f
Shaker mutation, 1055
Sheddases, 463
Sheet-forming collagens, 928–931, 930t, 931f, 932–933
Shibire, 1073
Shiverer mutation, 1062
Shmoo formation, 1019, 1019f
Shoot apical meristem, 1013
Shoot apical stem-cells, 1013–1015
Shoot meristem, 1015
Short interfering RNAs (siRNAs), 385t, 417, 421
Short interspersed elements (SINEs), 295
Short tandem repeats (STRs), 256, 257f
Shugoshin family of proteins, 884
Shuttle vectors, 240–241, 241f
Sialyl Lewis-x antigen, 950
Sickle-cell disease, 229, 253, 254t, 1011
Side chain, 45
Side-binding proteins, 806–807
Side-chain-specificity binding pocket, 98f, 100, 100f
σ-factors, 201

Signal amplification, 669–670, 687, 959
Signal peptidase, 585
Signal peptides, 579
Signal recognition particle (SRP), 329, 581, 582, 582f, 583f
extracellular signaling, 663f
Signal sequences, 579, 580–581
Signal transduction pathway, 662, 663–670, 958
activation of by allosteric proteins, 664
activation of by receptor tyrosine kinases and cytokine receptors, 706f–707f
allosteric proteins, as receptors, 664
anabolic, 965–966
Ca^{2+} ions as second messengers, 692–697
gene expression and, 707f
GTP-binding proteins, 668–669
ligands, receptors binding to, 665
ligands, receptors binding to with high affinity, 665–667
mTORC1 kinase activation, 966–969, 967f
on/off switches, 668–669
olfactory GPCRs, 1083f
overview of, 662f
phosphoinositide, 724–726
proteasomal degradation of signaling component, 740–750
Hedgehog signaling, 744, 744f
Hh receptors Patched and Smoothened, 745
Hh signaling, feedback regulation of, 745–746
Hh signaling in vertebrates, 746–747
NF-κB pathway, 749–750

NF-κB transcription factor, 747–748, 748f
Wnt protein, concentration gradient, 743–744, 743f
Wnt signaling, 741–742
protein kinases and phosphatases, 667–668, 668f
proteins, study of, 670–675
binding assays, 670–671, 671f
cellular response to signaling molecule, 671
external signals, sensitivity to, 671–672
immunoprecipitation assays, 673
pull-down assays, 674, 675f
receptors, purification of, 672–673
signaling molecules, chemical analogs of, 672
targeted fluorescent proteins, 674–675
Western blotting, 673–674, 674f
Ras/MAP kinase, 716–723
receptors, location of, 664–665
rhodopsin, 700
second messengers, 667, 667f
signal amplification and feedback repression, 669–670
signaling molecules, 663
site-specific protein cleavage, 738–740
from T-cell receptors and B-cell receptors, 1133f
types of changes produced by, 664
Signal-anchor sequence (SA), 588–591, 589f, 590f
Signal-based targeting, 576
Signaling cascades, 673, 889
Signaling circuits, 1046–1047, 1048f
Signaling kinases, 909f

Signaling pathways, 235, 235f
Signaling proteins, 71
Signaling receptors, down-regulation of, 654
Signalsome, 749–750, 749f
Signal-transducing proteins, 1169–1170
Silencer sequences, 355, 403–404, 403f
Silent mutations, 250
Simple columnar epithelia, 911f, 912, 912f
Simple diffusion, 471–472, 472f
Simple squamous epithelia, 911f
Simple transcription units, 280, 282f
Simple-sequence (satellite) DNA, 286–288, 288f
SINEs (short for short interspersed nuclear elements), 295, 296–297
Single cell RNA-Seq (scRNA-Seq), 264
Single cell sequencing, 33, 33f
Single-nucleotide mutations, 271f
Single-nucleotide polymorphisms (SNPs), 256, 256f, 257f, 259, 259f
Single-particle cryoelectron microscopy, 1079, 1080f
Single-pass membrane protein, 587
Single-pass transmembrane protein, 456, 456f
Single-stranded DNA, 890
Singlet microtubules, 799, 799f
Sinusoids, 1011, 1012f
siRNA knockdown, 422
Sister chromatid resolution, 882
Sister chromatids, 17f, 826, 850f, 884, 885f
Ska complex, 881f
Skeletal muscle

contraction, regulation of, 782–783, 783f
sarcomeres, structure of, 781f
sliding-filament model, 780–781, 781f
stabilizing and scaffolding proteins, 781–782, 782f
SKI complex, 410
Ski protein, 737–738
Skin cancer, 193
SKL sequence, 612–613
Sleep-wake cycle, 979
Sliding clamp, 188, 193
Sliding-filament assay, 775, 775f
SM proteins, 1070
SMAC/DIABLO proteins, 1037
Smads, 733–738, 736f
Small hairpin RNA (shRNA), 274, 275f
Small inhibitory RNAs (siRNAs), 272, 274
Small nuclear RNA (snRNA), 284
Small nuclear RNAs (snRNAs), 284, 385t, 392, 408–409, 408f
Small nucleolar RNA (snoRNA), 284, 298, 385t, 432–436, 433f, 434f
Small rRNA, 211
Small subunit (SSU) processome, 434, 435f
SMC proteins, 303–304, 304f
Smedwi-1 gene, 1002, 1002f
Smo, 745, 746f, 747
Smooth endoplasmic reticulum, 18
Smooth muscles, contraction in, 784–785, 785f
Smoothened receptor, 745
SNARE proteins, 635–636, 659

SNAREs, 1070, 1071, 1071f, 1072f
SNF5 gene, 1167
SnoN protein, 737–738
snRNPs (small ribonuclear particles) U1, U2, U4, U5, and U6, 394, 397f
SOCS box, 732
SOCS proteins, 732
SOD, 548
Sodium, rehydration therapy and, 508
Sodium chloride, model of, 34
Solitary genes, 282–284
Somatic cell, 991, 997–998, 997f, 1003–1004
Somatic cell recombination, 272
Somatic gene rearrangement, 1109, 1110f
Somatic hypermutation, 1113
Somatic mutations, 1159–1160, 1161f, 1163f
Somatic recombination, 1109, 1110f
Somatic-cell nuclear transfer (SCNT), 996, 997f
Sonic hedgehog mRNA, 261f
Sonication, 173
Sorting signals, 633, 638–639, 643, 653
Sos protein, 719f
Sour taste, 1082–1083
Sox11-nucleosome complex, 360f
SP proteins, 413
Special-pair bacteriochlorophylls, 564
Special-pair chlorophylls, 564
Specific activity, 124

Specific transcription factors, 343
Specificity, 95–96, 665
Spectrin, 770, 772
Sperm, 990, 1004–1005
Spherule, 698f
Sphingolipids, 448f, 449, 453–454, 464–468
Sphingomyelins, 449, 451, 452f, 465
Sphingosine, 465
Spinal cord, 1043, 1043f, 1060f
Spinal muscular atrophy, 405–406
Spindle assembly checkpoint pathway, 832, 892–893, 893f
Spindle pole body, 867, 868f
Spindle poles, 800, 800f, 826f, 834f, 850, 879
Spindle position checkpoint, 836
Spinning disk microscope, 156, 157f, 158f
Splice sites, 279, 392
Spliceosome, 393–398, 395f, 408f, 414–415
Splice-site selection, 400–402
Splicing enhancers, 401, 401f, 403–404, 403f
Splicing factors, 396f, 401f
Splicing repressors and activators, 404, 404f
Splicing silencers, 401
Spongiform encephalopathy, 95
Spontaneous mutations, 229
Sporogenous tissue, 952
Spot desmosomes, 913
Squamous cell carcinomas, 193
SR proteins, 401

Src gene, 714
SRE-binding proteins (SREBPs), 970–972, 971f
SRP (signal recognition particle), 217
Stable transfection, 265–266, 265f
Stalk, 815
Standard free-energy change ($\Delta G^{\circ\prime}$), 61–63
Staphylococcal nuclease protein, 70
Starch, 50, 559, 559f, 572
Stargazin, 1090
START, 848, 848f, 849, 853f, 867, 868f
Start (initiation) codon, 206
Start sites, 335
STAT proteins, 730, 731f
Statins, 466, 959
Steady state, 55–56, 56f
Steady-state phase, 759, 759f
Stem, 815
Stem cells, 32, 848, 991. *See also* Embryonic stem (ES) cells
    adult, 143–144
    asymmetric cell division, 1025–1026, 1025f
    in culture, 143–144
    differentiated cell types, formation of, 1003–1004, 1003f, 1004f
    embryonic, 140
    germ-line stem cells, 1004–1005
    hematopoietic, 1009–1013, 1010f, 1011f
    intestinal stem cells generate cells of intestinal epithelium, 1005–1007, 1006f, 1007f
    multicellular organisms and, 1002–1015

multipotent somatic stem cells, 1003–1004
planarians contain pluripotent stem cells, 1002–1003, 1002f
negative feedback loops, 1013–1015
neural, 1048–1051, 1050f
overview of, 991–992
in plants, 1013, 1014f
pluripotent, 143–144, 144f, 990f, 991
transplantation, 1009–1011, 1011f, 1012f
unipotent, 1002
Wnt and R-spondins, 1007–1009
Stem-cell factor (SCF), 1009
Stem-cell niche, 1004
Stems, 952
Step size, 777
Stereoisomers, 37, 37f
Steroid receptor superfamily, 348
Sterol regulatory elements (SREs), 970
Sterols, in biomembranes, 448f, 449
Sterol-sensing domain, 466, 970, 971f
Sticky ends, 238, 238f, 239f, 241, 243f
Stigma/stylar cysteine-rich adhesin (SCA), 955
Stimulated emission depletion (STED) microscopy, 163, 164f
Stimulatory G protein (G_s), 683
Stochastic optical reconstruction microscopy (STORM), 165
Stop (termination) codon, 206, 216–217, 216f
Stop-transfer anchor sequence (STA), 588–591, 589f, 590f
Store-operated channel (SOC), 695–696, 696f
Strand elongation stage of transcription, 201–202, 201f

Stratified squamous epithelia, 911f
Streptomyces lividans, 497f
Stress fibers, 755, 784
Stroma, 21, 559, 560f, 602
Stromal-import sequence, 609
Structural domains, 76, 79, 85f
Structural maintenance of chromosome (SMC) proteins, 303
Structural motifs, 76–77, 76f, 250
Structural nucleoporins, 614
Structural proteins, 71
Structured illumination microscopy (SIM), 163, 164f
SubbaRow, Yellapragada, 64
Subfamilies, 81
Subfornical organ (SFO), 1063, 1064, 1064f
Subgranular zone (SGZ), 1049, 1051f
Subgroups, 81
Submitochondrial compartments, 606–609, 607f, 608f
Substrata, 944
Substrate, 71, 90, 91f, 92, 97
Substrate-binding sites, 97–98, 98f, 100f
Substrate-level phosphorylation, 516
Subunit vaccine, 1146
Subunits, 83, 838
Subventricular zone (SVZ), 1049, 1051f
Succinate, 974
Succinate dehydrogenase, 533–534
Succinate-CoQ reductase (Complex II), 540, 541–542, 542f
Sucrose, 51f, 572

Sugar-phosphate backbone of DNA, 180–181, 181f
Suicide program, 1027
Sulfhydryl group (—SH), 47, 597
Supercoils, 184, 187
Supercomplexes, 84, 545–547, 546f
Super-enhancers, 361, 362f, 363f, 365
Superfamilies, 81
Supernatant, 116, 173
Superoxide anions, 547, 547f
Super-resolution microscopy, 139, 163–165, 164f
Supersecondary structures, 76–77
Suprachiasmatic nucleus (SCN), 981–982, 982f
Supramolecular complexes, 72f, 84
Suppressors of cytokine signaling (SOCS). *See* SOCS proteins
SURF complex, 426
Surveillance mechanisms, in cell cycle regulation, 889–893
SV40 DNA, 187, 188f, 189
Svedberg units (S), 211
Sweet taste, 1081–1082
SWI/SNF chromatin-remodeling complex, 358
Switch I, 668, 669f
Switch II, 668, 669f
Switch proteins, 108
Symmetric cell division, 990
Symporters, 473, 500–507
    Na^+/amino acid, 503–504, 503f
    Na^+-glucose, 502
    Na^+-linked, 501–503, 502f

two-Na^+/one-glucose, 501–503
Synapse elimination, 1068
Synapses
communication at, 1065–1077
action potential, inputs for, 1076, 1076f
Ca^{2+} triggers release of neurotransmitters, 1070
gap junctions, 1076–1077
muscle contraction, 1074–1075, 1074f
neurotransmitters, degradation or reuptake of, 1073–1074
neurotransmitters and H^+-linked antiport proteins, 1068–1070
nicotinic acetylcholine receptor, 1075, 1075f
plasma membrane, fusion with, 1070–1073, 1072f
presynaptic and postsynaptic structure, 1065–1068
presynaptic terminal, 1070
synaptic vesicles, recycling of, 1073
definition of, 28
electrical, 923
in human brain, 385
information flows between neurons, 1045–1046, 1046f
memory formation and, 1086–1088
in translation, 423
tripartite, 1066f
Synapsins, 1070
Synapsis, 230f, 894
Synaptic cleft, 1046f, 1137
Synaptic plasticity, 1042, 1086, 1088–1090, 1089f
Synaptic vesicles, 1045, 1066f, 1070–1073, 1071f, 1072f

Synaptonemal complex (SC), 894
Synaptotagmin, 1072
Synchronous signals, 1044
Syndecans, 937
Synechococcus elongatus, 980f–981f, 981
Synonymous codons, 206
Synonymous mutation, 207
Syntaxin, 635
Syntelic attachment, 881
Synteny, 28, 29f, 314
Synthetic lethal mutations, 235–236, 236f
Szostak, Jack, 318

T

T cells, 142, 142f, 1095, 1096f, 1181–1182
    activation, signals required for, 1136–1137, 1136f
    B cells, collaboration with, 1143–1145, 1144f
    CD8 co-receptor, 1137, 1137f
    cytokines, secretion of, 1138
    cytotoxic, 1137, 1137f
    development of, compared to B cells, 1135f
    MHC molecules, recognition of, 1134–1135
    proliferation and differentiation, 1132–1134, 1133f
Tagging, 267, 267f
Tail domain, 774, 811
Tail-anchored proteins, 590–591, 591f
Talin, 910, 946f–947f

Tamoxifen, 1176
Tandemly repeated arrays, 284
Tangles, 95
Taq polymerase, 244, 244f
Tardigrades, 1f
Target (destination) organelles, 629
Target cell, 662
Target field, 1032
Targeted fluorescent proteins, 674–675
Targeted therapy, 1162–1163
Targeting sequences, 578t, 579, 607f
Taste, sense of, 1080–1083, 1082f
Taste buds, 1080–1083
Taste receptors, 1081
Tat, 339, 340f
Tat (twin-arginine translocation) pathway, 611
TATA box, 334, 342, 342f, 357
TATA box-binding protein (TBP), 182f, 336
Taxol, 270, 805
Tay-Sachs disease, 254t
T-cell help, 1117
T-cell receptor (TCR), 1129, 1130f, 1131f
    similarity to immunoglobulin genes, 1131–1132
    structure of, 1129–1130
    variable residues, encoding of, 1132
T-cell signaling, 1180f
Telocytes, 1007–1008
Telomerase, 319–321, 321f, 1153

Telomere, 318, 354–355, 356f
Telomere (TEL) sequences, 318, 319f
Telomere terminal transferase, 320–321
Telomeric heterochromatin, 356f
Telomeric sequences, 319–321
Telophase, 827f, 828, 887f
Temperature, elevated, 976–978, 977f
Temperature-sensitive mutation, 25, 232, 233f, 852
Template, 185–186, 200
Temporal lobe, 1043f, 1044
Terminal deoxynucleotidyl transferase (TdT), 1112
Termination, 323
Termination codon, 206
Termination stage of transcription, 201f, 202
Termination stage of translation, 216–217, 216f
Ternary complex factor (TCF), 720
Tertiary structure, 72f, 77, 80f
Testants, 1075
Tetrahedral intermediate, 101f
Tetrahymena thermophila, 436
Tetraspanins, 920
TFIIB recognition element (BRE), 334
TGF-β family of growth factors, 733–738
    binding and activating signal transduction, 735
    latency and activation, 734f
    negative feedback loops, 737–738
    Smad transcription factors and, 735–737, 736f
    storage of in extracellular matrix, 733–735

TGF-β receptor family, 733, 735, 735f, 737
TGF-β receptor kinases, 735f
TGF-β signaling, 1173f
Thalassemia, 415
Thermal energy, 60
Thermus aquaticus, 244, 582, 582f
Thick filaments, 774, 780, 781f, 782f
Thick-filament regulation, 785
Thin filaments, 780, 781f, 782f, 783f
Thin sections, 168–169, 169f
Thin-filament regulation, 782
Thioredoxin (Tx), 570
Thompson, D'Arcy, 72
3-Hydroxyproline, 48f
Three Na^+/one-Ca^{2+} antiporter, 504
3′ cleavage/polyadenylation, 387, 398–399, 398f
3′ end, 180, 180f, 185–186
3′ → 5′ direction, 186, 186f
3-D adhesions, 943, 944
Threonine, 47, 668f
Threshold potential, 1076
Thrombopoietin, 727
Thrombospondin (TSP), 1065
Thylakoid lumen, 559
Thylakoid membranes, 559–560, 560f, 561, 562f
Thylakoids, 21, 559, 560f, 602, 610–611, 610f
Thymically derived regulator T cells (Treg), 1136
Thymidine, 124

Thymine, 48, 49, 180
Thymine-thymine dimers, 193, 194f, 195f
Thymocytes, 1125
Thymosin-β_4, 762
Thymus, 1125, 1135–1136
Tic complex, 609
Tight junction, 507, 508, 911, 912, 913t, 920f, 921f, 922f
    in blood-brain barrier, 1048
    composition of, 921f
    functions of, 912f, 913, 920–923, 920f, 922f
Time-of-flight (TOF) instruments, 126, 126f
TIMPs (tissue inhibitors of metalloproteinases), 942
+TIPS (plus-end tracking proteins), 807, 808f
TIRF microscopy, 158, 160f
Tissue culture, 855
Tissues, 901–956
    adhesive interactions in motile and nonmotile cells, 943–951
    basal lamina, 926–932
    cell-cell and cell–extracellular matrix adhesion, 903–910
    cell-cell and cell–extracellular matrix junctions and adhesion molecules, 911–926
    connective tissue, 932–942
    integrating cells into, 901–903
    in plants, 951–956
Titin gene, 280
Titin protein, 782
Titration curve, 59, 59f
TNF receptor-associated death domain protein (TRADD), 1039

Tolerance, 1094, 1095
Toll-like receptors (TLR), 747, 749f, 1098, 1099f, 1140–1143, 1141f
Tomogram, 170
Topogenic sequence, 587, 592f
Topoisomerase I, 184, 184f, 189
Topoisomerase II, 184
Topological domains, 79–80, 309–311
Topologically associated domains (TADs), 874f
Topology, 460, 587, 588–591, 589f, 590f, 593–594
TOR pathway, 424, 964–965
Torsional stress, 184, 184f
Total internal reflection fluorescence (TIRF) microscopy, 158, 160f
Totipotent, 991, 992
Toxins, 487–490, 769
TRAF6, 749
TRAMP complex, 410
Trans (intercellular) interactions, 905
Trans configuration, 87, 88, 88f
Trans fats, 53
Transcellular pathway, 921
Transcellular transport, 507–510, 922f
    bone resorption, 509–510, 509f
    of glucose and amino acids, 507–508, 508f
    rehydration therapy, 508
    stomach contents, acidification of, 508–509, 509f
Transcription, 11, 11f, 178, 179f, 199–205
    activation of, 688, 689f
    activators, 345, 346f

definition of, 87, 199
divergent, 335–336
epigenetic regulation, 372–380
DNA methylation, 372–373
long noncoding RNAs, 377–379
methylation of specific histone lysines, 373–375
by polycomb and trithorax complexes, 375–377
in eukaryotic genomes, 280–282, 327–333
carboxy-terminal domain, 332–333
clamp domain, 331–332
nuclear RNA polymerases, 329–331
regulation of, 342–343
regulatory elements, 327–329
induction of, 721f
initiation of, 361–365
nonhistone proteins and, 311–312
other eukaryotic systems, 380–382
Pol I, 380–381
Pol III, 381–382
regulation of, 324, 324f, 326f
activation and repression domains, 349–350
repression and activation
activators and histone acetylation, 358
bursts, 365
chromatin-remodeling complexes, 358–359
heterochromatin and, 354–355
Mediator, 359–361
molecular mechanisms of, 354–366

pioneer transcription factors, 359
repressors and histone deacetylation, 355–358
transcriptional condensates, 361–365
by RNA polymerase I, 380–381, 380f
by RNA polymerase III, 381–382, 381f
stages of, 201–202, 201f
termination of, 371
Transcription bubble, 201
Transcription factors
activation of, 730
activators and, 346f
activity, regulation of, 366–372
cellular differentiation, 366–368
hormones binding to nuclear receptors, 369–370
metazoans, 370–371
nuclear receptors, 368–370
termination of transcription, 371
AU-rich elements and, 419
definition of, 11
induction of a particular gene, 708f
interphases chromatins and, 311
MAP kinase and, 720–721
master set of, 29–30
mutations, 325f
NF-κB, 747–748, 748f
in Notch/Delta pathway, 738–740
in nuclear-receptor superfamily, 369f
oxygen-sensitive, 974–975, 975f

pioneer, 359
proteins, evaluation of, 345f
regulation of protein-coding genes, 350–352
in secondary structures, 76
Smads, 733–738
template DNA to complement RNA, 200–202
TGF-β/Smad signaling pathway, 735–737, 736f
transcription, regulation of, 178, 324, 326f, 327, 343
Wnt signaling and, 741–742
Transcription initiation complex, 84f
Transcription start sites (TSS), 335–336, 341, 363f
Transcription unit, 280, 282f
Transcriptional (mRNA) profiling, 133
Transcriptional activation, 863–865, 864f
Transcriptional condensates, 361–365
Transcriptional memory, 366
Transcriptional network, 996f
Transcription-control elements, 340, 341f, 342f
Transcription-control region, 324, 327–328, 328f, 333
Transcytosis, 649, 1023, 1105, 1105f
Transducin, 699, 702–703, 702f
Transduction, 664
Transepithelial resistance, 922
Transepithelial transport, 922f
Transesterification, 391–393, 392f
Transfection, 265
Transfer RNA (tRNA), 178, 179f, 205–210
amino acid links, 209–210

folded structure of, 207–208
nonstandard base pairing, 208–209
processing of, 431–438
roles of in protein synthesis, 205, 205f, 207–208, 208f
structure of, 209f
synthesis of, 966
Transferrin, 425
Transformation, 239, 265–266, 1161
Transformed cells, 140
Transformer genes, 403–404
Transforming growth factor (TGF-β), 733
***trans*-Golgi network**
definition of, 625
enzymes, trafficking of, 646f
functions of, 626, 629f, 630, 639, 640–641
N-linked oligosaccharides and, 640f
protein aggregation in, 645–647
in secretory and endocytic pathways, 624f
vesicle-mediated protein trafficking, 642f, 643
Transient dipoles, 41–42
Transient receptor potential (TRP) channels, 909
Transient transfection, 265, 265f
Transit amplifying cells, 1003, 1003f
Transition state, 63, 97, 100f
Transitional epithelia, 911f
Transition-state intermediate, 63
Translation, 10, 87, 178
cytoplasmic polyadenylation and, 422–423, 423f

efficiency of, 217–218, 217f
initiation of, 213, 214f, 582
polypeptide passage through translocon, 582–585
regulation of
sequence-specific RNA-binding proteins and, 424–426
surveillance mechanisms, 426–428, 426f, 427f
simultaneous translocation, 581f
termination of, 216–217
Translesion polymerase, 193
Translocation, 215, 579–580, 581–582, 581f, 583f, 610–611
Translocation channel, 579
Translocon, 582–585, 584f
Transmembrane AMPA receptor regulatory proteins (TARPs), 1090
Transmembrane channels, 555
Transmembrane collagens, 930t, 933
Transmembrane domain (TMD), 711f
Transmembrane electric gradient, 494–495, 495f
Transmembrane proteins. *See also* Integral membrane proteins
alpha (α) helix, 455–458, 456f, 457f
orientation in bilayer, 460–461
Transmembrane transport, 470–510
antiporters, cotransport by, 500–507
ATP-powered pumps, 480–484
Ca^{2+} pump mechanism of action, 483–484
classes of, 480–482, 481f
ionic gradients, 482
muscle relaxation and, 482–483, 483f, 485f
cotransport

anion antiporter, 504–505, 505f, 506f
cytosolic pH, regulation of, 504
metabolites and ions, accumulation of, 505–507, 506f
Na^+ entry into mammalian cells, 500–501, 501f
Na^+/amino acid symporter, 503–504, 503f
Na^+-linked Ca^{2+} antiporter, 504
Na^+-linked symporters, 501–503, 502f
facilitated transport of glucose and water, 474–480
aquaporins, 478–480, 479f
GLUT proteins, 477
GLUT1 uniporter, 475–477, 475f, 476f
osmotic pressure, 478, 478f
transport proteins, study of, 477–478
uniport, 475
intracellular ionic environment, 484–493, 490f
ABC proteins, 487–491, 489t, 490f, 492f
cystic fibrosis transmembrane regulator (CFTR), 491–493, 493f
Na^+/K^+ ATPase, 484–486, 486f
V-class ATPases, 486–487, 487f, 488f
nongated ion channels, 494–500
molecular selectivity filter, 496–498, 498f
oocyte expression and patch clamping, 499–500, 499f
patch clamps and, 498–500, 499f
transmembrane electric gradient, 494–495, 495f
overview of, 471–474, 474t
membrane proteins, classes of, 472–474, 472f, 473f
simple diffusion, 471–472, 472f

resting membrane potential, 494–500
K^+ ions and, 495–496, 497f
symporters, cotransport by, 500–507
transcellular transport, 507–510
bone resorption, 509–510, 509f
of glucose and amino acids, 507–508, 508f
rehydration therapy, 508
stomach contents, acidification of, 508–509, 509f
Transmigration, 951
Transmission electron microscope, 167, 167f, 168f, 169f, 170f
Transmutation of Species (Darwin), 3f
Transplant rejection, 1120
Transport proteins, 477–478, 507–508
Transport vesicles, 622f, 623, 623f, 627–628, 633t
Transporters, 473
Transposable (mobile) DNA element, 289–300, 290f
in bacteria, 290–291
in eukaryotes, 291–292
evolution and, 298–300
LTR retrotransposons, 292–295
movement of, 289–290
non-LTR retrotransposons, 295–297
retroposed RNAs, 297–298
Transposase, 291, 291f, 292, 299f
Transposition, 289, 291f, 299f
Transposons, 289–290, 290f, 292, 422
Trans-splicing, 397–398
Transverse tubules, 782

TRBP (Tar binding protein), 420
Treadmilling, 760
Triacylglycerols (triglycerides), 52–53, 534
Tricarboxylic acid (TCA) cycle, 533, 1153–1154
Tricellulin, 921, 921f
Trichomonas vaginalis, 25
Triglycerides, 52–53, 534
Trimeric G protein, 424
Triplet code, 206
Triplet microtubules, 799, 799f, 821f
Triplet oxygen ($^3O_2$), 564–565
Triskelions, 642, 643f
Tri-snRNP complex, 395
Trisomy, 898
Trithorax complexes, 375–377, 377f
$tRNA^{Tyr}$ gene, 218
Trophectoderm (TE), 993, 994f, 995
Trophic factor, 1027, 1037
Tropoelastin, 941–942
Tropomodulin, 763, 782, 782f
Tropomyosin, 782, 783, 783f
Troponin, 782, 783, 783f
TRPase-superfamily proteins, 218
TRPV1 channel, 1079, 1080f
True-breeding strains, 231
Trypanosoma brucei, 25
Trypsin, 98f, 100f
Tryptophan, 46

t-SNAREs, 630, 634f, 635–636, 637f, 639
Tuberous sclerosis complex (TSC), 969
Tubulin, 797
Tubulin proteins, 250–251, 251f, 754
    modifications of, 818–819, 819f
    polymerization, drugs affecting, 805
TUG protein, 961–962
Tumor cells
    cell-cell interactions, 1154
    cellular housekeeping functions, 1153–1154
    formation of new blood vessels, 1155
    genetic makeup of, 1152, 1152f, 1153f
    invasion and metastasis, 1155–1156, 1155f, 1156f
    vs. normal cells, 1151–1156
    uncontrolled proliferation, 1152–1153
Tumor microenvironment, 1149f, 1154, 1179–1181
Tumor necrosis factor, 1039
Tumor necrosis factor alpha (TNFα), 747, 1028, 1039
Tumor suppressor gene, 981, 1165–1166
Tumor viruses, 1160–1161
Tumorigenesis, 1149, 1155–1156
    animal models, 1174–1176, 1176f
    colon cancers, 1173–1174, 1175f
    growth control pathways, 1170–1171
    molecular cell biology, 1176–1177
    multi-hit model, experimental reconstruction of, 1172–1173
    nuclear transcription factors, 1171–1172
    receptor mutations, 1168–1169, 1169f

signaling pathways, aberrations in, 1172
signal-transducing proteins, 1169–1170
Tunicamycin, 596f
Tunneling nanotubules, 527, 925, 926f
Turgor pressure, 478, 938
Turnover number, 99
20S proteasome, 105–106
26S complex, 105f, 106
2′-hydroxyl group, 182, 183f
Two-dimensional electrophoresis, 118–119, 119f
Two-Na^+/one-glucose symporter, 501–503
2-Oxoglutarate, 974
Two-photon excitation microscopy, 157–158, 159f
Ty transcription, 295f
Type I collagen, 935
Type I diabetes, 964, 998–999
Type I interferons, 1100
Type I proteins, 588, 589f
Type II collagen, 935
Type II diabetes, 259, 960, 964, 998–999
Type II proteins, 588–590, 590f
Type III proteins, 588–590, 590f
Type IV collagen, 928, 928f, 931
Type IV (multipass) proteins, 591–592
Type IX collagen, 935
Type VI collagen, 934, 934f
Tyrosine, 46, 180f
Tyrosine kinase domain (TKD), 711f

Tyrosine tRNA (tRNATyr) gene, 436, 436f

U

U2 associated factor (U2AF), 395
Ubiquinone [coenzyme Q (CoQ)], 540, 540f
Ubiquitin, 105f, 106–107, 114f, 715, 863–865, 864f
Ubiquitin-activating enzyme (E1), 106, 107
Ubiquitin-conjugating enzyme (E2), 106, 107
Ubiquitin-protein ligase (E3), 106, 107
Ubiquitinylation, 113–114
Umami taste, 1081–1082
Unclassified intergenic DNA, 288–289
Uncoupler, 558
Uncoupling protein 1 (UCP1), 558
Unfolded-protein response, 600, 600f
Uniport, 475
Uniporters, 473, 475, 502. *See also* GLUT1
Universal code, 207
Unsaturated, 464
Unsaturated fatty acids, 52, 464, 970–972
Unstimulated state, 666
Untranslated regions (UTRs), 203, 203f, 422
Upstream, 200f, 201, 327
Upstream activating sequences (UASs), 343
Upstream control element, 380
Uracil, 48, 49
UV light, 1158

V

V gene segments, 1109–1111, 1132
Vaccines, 1102, 1145–1146
Vacuoles, 20f
    acidity of, 486–487, 487f
    functions of, 19
    metabolites and ions, accumulation of, 505–507, 506f
Valence, 494
Valine, 46
VAMP (vesicle-associated membrane protein), 635, 1070–1071, 1071f
Van der Waals interactions, 41–42, 42f, 444f
Van Leeuwenhoek, Antoni, 138
Vang gene, 1025
Variable region of the heavy chain (V_H), 1106
Variable region of the light chain (V_L), 1106
Vascular CAM-1 (VCAM-1), 948
Vascular endothelial growth factor (VEGF), 973
Vascular tissue, 952
Vassopressin, 479
V-class ATPases, 486–487, 487f, 488f
V-class pumps, 480–481, 481f, 487, 487f, 509–510, 509f
VDAC (voltage-dependent anion channel), 523
V(D)J recombination, 1111–1113, 1111f, 1112f
Vector plasmid, 266, 266f
Vectors, 237, 238, 239f, 240f, 242, 264f
Velocity (*V*), 63
Ventricles, 1048, 1050f

Ventricular zone (VZ), 1048, 1050f
Venus flytrap, 956
Vertebrates, 746–747, 1037–1039
Vesicle budding and fusion
    molecular mechanisms of, 629–636, 630f, 632f
        cargo proteins, targeting sequences on, 633
        GTPase switch proteins, 631–633
        protein coat, assembly of, 629, 632f
        Rab GTPases control docking, 633–635, 634f
        Rab proteins, 635t
        SNARE proteins, 635–636
        sorting signals, 633t
    visualization of, 631f
Vesicle buds, 630, 631f
Vesicle trafficking, 623f
Vesicle-based trafficking, 576
Vesicles, 175f
    coated, 631–633, 631t, 632f
    formation of, 629
    Rab proteins involved in tethering and movement of, 635t
    synaptic, 1066f
Vesicular GABA transporter (VGATs), 1070
Vesicular glutamate transporters (VGLUTs), 1069–1070
Vesicular stomatitis virus (VSV), 626, 626f, 627, 627f
Vestibule, 497
Vibrio cholerae, 923
Vibrio parahaemolyticus, 502f
Villi, 1005

Viral capsids, 220–221
Viral envelope, 221
Viral infection, time course of, 1145f
Virchow, Rudolf, 847
Virion, 219, 220f, 221, 222, 223f
Virus, 219–224
    DNA of in host-cell genome, 222–224
    lytic growth cycles, 221–222
    range of, 219
    viral capsids, 220–221
Virus-associated (VA) RNA, 424
Vision, 698–703
    ambient light, adaptation to, 702–703, 702f
    cGMP-gated cation channels, closing of, 700, 701f
    GTP hydrolysis, 702
    rhodopsin in rod cells, activation of, 698–699, 699f, 701f
    rhodopsin phosphorylation and arrestin binding, 701
    signal amplification, 700
    temporal resolution of, 700
Visual cortex, 698, 1078
Vitamin D_3 response element (VDRE), 369
V_{max} (maximal velocity), 98, 98f
Voltage-gated Ca^{2+} channels, 1070
Voltage-gated channels, 1052–1064, 1052f
    channel-inactivating segment, 1059, 1059f
    K^+ channels, 1054–1055, 1054f
    myelination and velocity of impulse conduction, 1059–1060
    Na^+ channels, 1052, 1053–1054, 1053f

sequential opening and closing, 1052–1055, 1053f
structures of, 1055–1056, 1057f, 1058f
Voltage-gated K^+ channels, 1053, 1054–1055, 1056f, 1057, 1057f, 1059
Voltage-gated Na^+ channels, 1053–1054, 1053f, 1056–1057, 1056f, 1059
Voltage-sensing α helices, 1054
Voltage-sensitive K^+ channel, 1058f
Von Hippel-Lindau (VHL) protein, 974
Von Willabrand factor, 948
v-SNAREs, 630, 634f, 635–636, 637f, 639, 644

W

Warburg, Otto, 1153
Warburg effect, 1153
WASp, 765
Water
dipole nature of, 38, 38f
facilitate transport of, 474–480
hydrogen bonds and, 41f
Water-accessible surface diagram, 78f
Watson, James D., 8, 180
Watson-Crick base pairs, 181
Web-footed chicken, 1027f
Wee1, 864, 876
Western blotting, 122, 123f, 142, 673–674, 674f, 1095. *See also* Immunoblotting (IB)
White-fat tissue, 558

Whole genomic shotgun sequencing, 246
Whole transcriptome shotgun sequencing (RNA-Seq), 263
Wieschaus, Eric, 232
Wild type alleles, 228, 232f
Wilkins, Maurice, 180
Wilson, H. V., 906
Winged helix, 359
Wnt, 740
Wnt protein, 743–744, 743f
Wnt signaling, 705f, 706f–707f, 708, 741–742, 742f, 1007–1009
Wobble position, 208–209, 210f
Wounded-cell monolayer assay, 792f
Writers, 110
WUSCHEL (WUS) gene, 1013–1014

X

X-chromosome inactivation, 308–309, 377–378, 378f
Xenopus cells, 390f
Xenopus laevis, 850, 852, 854–855, 854f, 878f
Xeroderma pigmentosum, 193, 195f, 1159t
X-inactivation center, 377, 378f
Xist gene, 377–378, 378f
X-linked mental retardation, 109
X-linked recessive diseases, 254–255, 254t, 255f
X-linked Severe Combined Immune Deficiency (SCID), 1011
XMAP215 protein, 807, 809f, 829, 830f
X-ray crystallography, 129–130, 129f

X-rays, 1158
Xyloglucans, 953

Y

Yamanaka, Shinya, 997
Yeasts
- asymmetric growth during mating, 1018–1019, 1019f
- budding and fission, 852–853, 853f
- conditional mutations, 231–232
- genomic libraries, 240–241, 241f, 242f
- post-translational translocation, 585–586, 585f
- in research, 24f
- cell cycle proteins, 24–25
- cell structure and function, 22–24
- ribosomes, structure of, 212f
- shmoo formation, 1019f

Yellow fluorescent protein (YFP), 160, 161f, 162, 162f
YXXφ sorting signal, 643

Z

Z disk, 780, 781f, 782, 782f
Zebrafish, 32
Zellweger syndrome, 613
Zinc finger, 76f, 77, 348–349
Zones, 117
Zonula occludens (ZO), 921

Zwitterion, 58
Zygote, 854, 992
Zymogens, 115

Glossary

Boldface terms within a definition are also defined in this glossary.

Figures and tables that illustrate defined terms are noted in parentheses.

AAA ATPase family
A group of proteins that couple hydrolysis of ATP with large molecular movements usually associated with unfolding of protein substrates or the disassembly of multisubunit protein complexes.

ABC superfamily
A large group of integral membrane proteins that often function as ATP-powered **membrane transport proteins** to move diverse molecules (e.g., phospholipids, cholesterol, sugars, ions, peptides) across cellular membranes. (Figure 11-15)

acetyl CoA
Small, water-soluble metabolite comprising an acetyl group linked to coenzyme A (CoA). The acetyl group is transferred to citrate in the **citric acid cycle** and is used as a carbon source in the synthesis of fatty acids, steroids, and other molecules. (Figure 12-12)

acetylcholine (ACh)
Neurotransmitter that functions at vertebrate neuromuscular junctions and at various neuron-neuron synapses in the brain and peripheral nervous system. (Figure 22-26)

acid
Any compound that can donate a proton (H^+). The carboxyl and phosphate groups are the primary acidic groups in biological macromolecules.

actin
Abundant structural protein in eukaryotic cells that interacts with many other proteins. The monomeric globular form (*G-actin*) polymerizes to form actin filaments (*F-actin*). In muscle cells, F-actin interacts with **myosin** during contraction. See also **microfilament**. (Figure 17-5)

actin cross-linking proteins
Proteins in a vertebrate cell that assemble actin filaments into organized structures.

action potential
Rapid, transient, all-or-none electrical activity propagated in the plasma membrane of excitable cells (e.g., neurons and muscle cells) as the result of the selective opening and closing of voltage-gated Na^+ and K^+ channels. (Figures 23-2 and 23-9)

activation domain
A region of an activator transcription factor that will stimulate transcription when fused to a DNA-binding domain.

activation energy
The input of energy required to (overcome the barrier to) initiate a chemical reaction. By reducing the activation energy, an **enzyme** increases the rate of a reaction. (Figure 2-30)

activation loop
A region in most protein-tyrosine kinases, containing a tyrosine residue that, when phosphorylated, increases kinase activity.

activator
Specific **transcription factor** that stimulates transcription.

active site
Specific region of an enzyme that binds a **substrate** molecule(s) and promotes a chemical change in the bound substrate. (Figure 3-24)

active transport
Protein-mediated movement of an ion or small molecule across a membrane against its concentration gradient or electrochemical gradient driven by the coupled hydrolysis of ATP. (Figure 11-3, **[1]**; Table 11-1)

adapter proteins
Adapter proteins physically link one protein to another protein by binding to both of them. Adapter proteins directly or indirectly (via additional adapters) connect cell-adhesion molecules or adhesion receptors to elements of the cytoskeleton or to intracellular signaling proteins.

adapter protein (AP) complexes
Heterotetrameric protein complexes mediate intracellular membrane trafficking along endocytic and secretory transport pathways. Three different AP complexes are known (AP1, AP2, and AP3), each with four subunits of different, though related, proteins.

adenosine triphosphate (ATP)

A nucleotide that is the most important molecule for capturing and transferring free energy in cells. Hydrolysis of each of the two phosphoanhydride bonds in ATP releases a large amount of free energy that can be used to drive energy-requiring cellular processes. (Figure 2-31)

adenylyl cyclase
One of several enzymes that is activated by binding of certain ligands to their cell-surface receptors and catalyzes formation of cyclic AMP (cAMP) from ATP; also called *adenylate cyclase*. (Figure 15-20)

adherens junction
A type of anchoring junction that connects the lateral membranes of adjacent epithelial cells and is usually located near the apical surface, just below the tight junction. (Figure 20-11)

adhesion receptor
Protein in the plasma membrane of animal cells that binds components of the **extracellular matrix**, thereby mediating cell-matrix adhesion. The **integrins** are major adhesion receptors. (Figure 20-1, **[5]**)

ADP (adenosine diphosphate)
The product, along with inorganic phosphate, of ATP hydrolysis by ATPases.

aerobic
Referring to a cell, organism, or metabolic process that utilizes gaseous oxygen (O_2) or that can grow in the presence of O_2.

aerobic oxidation
Oxygen-requiring metabolism of sugars and fatty acids to CO_2 and H_2O coupled to the synthesis of ATP.

aerobic respiration
See **aerobic oxidation**.

afferent neurons
Nerves that transmit signals from peripheral tissues to the central nervous system.

agonist
A molecule, often synthetic, that mimics the biological function of a natural molecule (e.g., a hormone).

Agrin

A glycoprotein synthesized by developing motor neurons that increases MuSK kinase activity in a muscle cell, facilitating development of a neuromuscular junction. (Figure 23-24)

Akt
A cytosolic serine/threonine kinase that is activated following binding to PI 3,4-bisphosphate and PI 3,4,5-trisphosphate; also called *protein kinase B*.

allele
One of two or more alternative forms of a gene. Diploid cells contain two alleles of each gene, located at the corresponding site (locus) on **homologous chromosomes**.

allostery
Change in the tertiary and/or quaternary structure of a protein induced by binding of a small molecule to a specific regulatory site, causing a change in the protein's activity.

alpha carbon atom (C*α*)
In amino acids, the central carbon atom that is bonded to four different chemical groups (except in glycine) including the side chain, or R group. (Figure 2-4)

alpha α helix
Common protein **secondary structure** in which the linear sequence of amino acids is folded into a right-handed spiral stabilized by hydrogen bonds between carboxyl and amide groups in the backbone. (Figure 3-4)

alternative splice site
The specific site at which splicing takes place during the processing of pre-mRNA into mature mRNA. ~95 percent of mammalian mRNAs are processed from primary transcripts that are alternatively spliced into more than one type of mRNA, often in a cell-type-specific manner.

alternative splicing (alternative RNA splicing)
Process by which the exons of one pre-mRNA are spliced together in different combinations, generating two or more different, mature mRNAs from a single pre-mRNA. (Figure 5-28)

amino acid
An organic compound containing at least one amino group and one carboxyl group. In the amino acids that are the **monomers** for building proteins, an amino group and carboxyl group are linked to a central carbon atom, the α carbon, to which a variable side chain is attached. (Figures 2-4 and 2-14)

aminoacyl-tRNA
Activated form of an amino acid, used in protein synthesis, consisting of an amino acid linked via a high-energy ester bond to the 3′-hydroxyl group of a **tRNA** molecule. (Figure 5-31)

amphipathic (or amphiphilic)
Referring to a molecule or structure that has both a **hydrophobic** and a **hydrophilic** part.

amphitelic attachment
Describes the correct attachment of chromosomes to the mitotic spindle, where sister kinetochores attach to microtubules emanating from opposite poles. (Figure 19-25)

anaphase
Mitotic stage during which the **sister chromatids** (or duplicated homologs in meiosis I) separate and move apart (segregate) toward the spindle poles. (Figures 18-38 and 18-39)

anchoring junctions
Specialized regions on the cell surface containing **cell-adhesion molecules** or **adhesion receptors**; includes *adherens junctions* and **desmosomes**, which mediate cell-cell adhesion, and *hemidesmosomes*, which mediate cell-matrix adhesion. (Figures 20-14 and 20-16)

anaerobic respiration
Respiration in which molecules other than oxygen, such as sulfate or nitrate, are used as the final recipient of the electrons transported via the electron-transport chain.

anaphase-promoting complex (cyclosome; APC/C)
A ubiquitin ligase that targets securin, mitotic cyclins, and other proteins for proteasomal degradation from the onset of anaphase until entry into the subsequent cell cycle.

aneuploidy
Any deviation from the normal diploid number of chromosomes in which extra copies of one or more chromosomes are present or one of the normal copies is missing.

angiogenesis
The developmental process for the formation of blood vessels that are needed to supply nutrients to a growing tumor.

anion
A negatively charged ion.

antagonist

A molecule, often synthetic, that blocks the biological function of a natural molecule (e.g., hormone).

anterograde
Vesicular trafficking in the direction of ER to plasma membrane.

antibody
A protein (immunoglobulin), normally produced in response to an **antigen**, that interacts with a particular site (**epitope**) on the same antigen and facilitates its clearance from the body. (Figure 3-22)

anticodon
Sequence of three nucleotides in a tRNA that is complementary to a **codon** in an mRNA. During protein synthesis, base pairing between a codon and anticodon aligns the tRNA carrying the corresponding amino acid for addition to the growing polypeptide chain. (Figure 5-32)

antigen
Any material (usually foreign) that elicits an immune response. For B cells, an antigen elicits formation of antibody that specifically binds the same antigen; for T cells, an antigen elicits a proliferative response, followed by production of **cytokines** or the activation of cytotoxic activity.

antigen-presenting cell (APC)
Any cell that can digest an antigen into small peptides and display the peptides in association with class II MHC molecules on the cell surface where they can be recognized by T cells. *Professional* APCs (dendritic cells, macrophages, and B cells) constitutively express class II MHC molecules. (Figures 24-26 and 24-28)

antigen processing and presentation
The sequence of events by which a protein capable of eliciting an immune response is converted into peptides, presented by products of the major histocompatibility complex for display at the cell surface.

antiport
A type of **cotransport** in which a membrane protein (*antiporter*) transports two different molecules or ions across a cell membrane in opposite directions. See also **symport**. (Figure 11-3, **[2C]**)

antiporter

An integral membrane protein that functions as a cotransporter to transport two molecules in opposite directions across a membrane.

apical
Referring to the tip (apex) of a cell, an organ, or other body structure. In the case of epithelial cells, the apical surface is exposed to the exterior of the body or to an internal open space (e.g., intestinal lumen, duct). (Figure 20-10)

Apicomplexa
A large family of unicellular parasitic organisms; many cause human diseases such as malaria and toxoplasmosis.

apoptosis
A genetically regulated process, occurring in specific tissues during development and disease, by which a cell destroys itself; marked by the breakdown of most cell components and a series of well-defined morphological changes; also called *programmed cell death.* See also **caspases**. (Figures 22-34 and 22-42)

aquaporins
A family of **membrane transport proteins** that allow water and a few other small uncharged molecules, such as glycerol, to cross biomembranes. (Figure 11-8)

ARF protein
A family of small GTPases that trigger the assembly of vesicles coated with COPI or clathrin-AP complexes.

Argonaute protein
A multidomain member of a protein family with a recognizable conserved sequence. Several Argonaute proteins are expressed in some organisms, especially plants, and are found in distinct **RISCs** with different functions. Humans express four Argonaute proteins.

Arp2/3 complex
A complex of seven proteins, including two actin-related proteins, able to nucleate the assembly of actin filaments when stimulated by F-actin and a nucleation promoting factor (NPF).

aster
Structure composed of microtubules (astral fibers) that radiate outward from a **centrosome** during mitosis. (Figures 18-38 and 18-39)

astrocytes

Star-shaped glial cells in the brain and spinal cord that perform many functions, including support of endothelial cells that form the blood-brain barrier, maintain extracellular ion composition, and provide nutrients to neurons.

asymmetric cell division
Any cell division in which the two daughter cells receive the same genes but otherwise inherit different components (e.g., mRNAs, proteins) from the parental cell. (Figure 22-24b)

ataxia telangiectasia and Rad3-related protein (ATR)
One of a pair of homologous protein kinases, together with **ataxia telangiectasia mutated (ATM)**, that is central for DNA damage signaling and repair. A related protein kinase called **DNA-PK** is also critical.

ataxia telangiectasia mutated (ATM)
One of a pair of homologous protein kinases, together with **ataxia telangiectasia and Rad3-related protein (ATR)**, that is central to DNA damage signaling and repair. A related protein kinase called **DNA-PK** is also critical.

ATP-powered pump
Any transmembrane protein that has ATPase activity and couples hydrolysis of ATP to the active transport of an ion or small molecule across a biomembrane against its electrochemical gradient; often simply called *pump*. (Figure 11-9)

ATP synthase
Multimeric protein complex, bound to inner mitochondrial membranes, thylakoid membranes of chloroplasts, and the bacterial plasma membrane, that catalyzes synthesis of ATP during oxidative phosphorylation and photosynthesis; also called F_0F_1 *complex*. (Figure 12-23a)

Aurora kinases
Serine/threonine kinases that play a crucial role in cell division by controlling chromatid segregation. Aurora B kinase destabilizes faulty microtubule-kinetochore interactions by phosphorylating microtubule-binding components within the kinetochore.

autoimmunity
A state in which a host's immune system can mistake the host's own cells and tissues for pathogens and mount an attack against them, characterized by the presence of antibodies and/or T cells specific for antigens in that same individual.

autophagy

Literally, "eating oneself"; the process by which cytosolic proteins and organelles are delivered to the lysosome, degraded, and recycled. Autophagy involves the formation of a double-membrane vesicle called an *autophagosome* or *autophagic* vesicle. (Figure 14-34)

autoradiography
Technique for visualizing radioactive molecules in a sample (e.g., a tissue section or electrophoretic gel) by exposing a photographic film (emulsion) or two-dimensional electronic detector to the sample. The exposed film is called an *autoradiogram* or *autoradiograph*.

autosome
Any chromosome other than a sex chromosome.

axon
Long process extending from the cell body of a neuron that is capable of conducting an electric impulse (**action potential**), generated at the junction with the cell body, toward its distal, branching end (the axon terminus). (Figure 23-1)

axonal transport
Motor protein–mediated transport of organelles and vesicles along microtubules in axons of nerve cells. *Anterograde* transport occurs from cell body toward axon terminal); *retrograde* transport, from axon terminal toward cell body. (Figures 18-16 and 18-17)

axoneme
Bundle of **microtubules** and associated proteins present in **cilia** and **flagella** and responsible for their structure and movement. (Figure 18-31)

bacteria
Class of prokaryotes that constitutes one of the three distinct evolutionary lineages of modern-day organisms; also called *eubacteria*. Phylogenetically distinct from archaea and eukaryotes. (Figure 1-1)

bacteriophage (phage)
Any virus that infects bacterial cells. Some phages are widely used as **vectors** in **DNA cloning**.

basal
See **basolateral**.

basal body

Structure at the base of a **cilium** or **flagellum** from which microtubules forming the **axoneme** assemble; structurally similar to a **centriole**. (Figure 18-31)

basal lamina (pl. basal laminae)
A thin sheet-like network of **extracellular-matrix** components that underlies most animal epithelia and other organized groups of cells (e.g., muscle), separating them from connective tissue or other cells. (Figures 20-22 and 20-23)

base
Any compound, often containing nitrogen, that can accept a proton (H^+) from an acid. Also, commonly used to denote the **purines** and **pyrimidines** in DNA and RNA.

base pair
Association of two complementary **nucleotides** in a DNA or RNA molecule stabilized by hydrogen bonding between their base components. Adenine pairs with thymine or uracil ($A \cdot T$, $A \cdot U$) and guanine pairs with cytosine ($G \cdot C$).

basic helix-loop-helix (bHLH)
See **helix-loop-helix**.

basolateral
Referring to the base (basal) and side (lateral) of a polarized cell, organ, or other body structure. In the case of epithelial cells, the basolateral surface abuts adjacent cells and the underlying **basal lamina**. (Figure 20-10)

B cell
A lymphocyte that matures in the bone marrow and expresses antigen-specific receptors (membrane-bound **immunoglobulin**). After interacting with antigen, a B cell proliferates and differentiates into **antibody**-secreting *plasma cells*.

B-cell receptor
Complex composed of an antigen-specific membrane-bound immunoglobulin molecule and associated signal-transducing Igα and Igβ chains. (Figure 24-19)

Bcl-2 family
A group of evolutionarily conserved proteins that contain BH domains and that either promote or inhibit apoptosis.

benign
Referring to a tumor containing cells that closely resemble normal cells. Benign tumors stay in the tissue where they originate but can be harmful due to continued growth. See also **malignant**.

beta (β)-adrenergic receptors
Seven spanning G protein–coupled receptors that bind adrenaline and related molecules, leading to activation of adenylyl cyclase.

beta (β) sheet
A flat **secondary structure** in proteins that is created by hydrogen bonding between the backbone atoms in two different polypeptide chains or segments of a single folded chain. (Figure 3-5)

beta (β) turn
A short U-shaped **secondary structure** in proteins. (Figure 3-6)

BH3-only protein
A member of the Bcl-2 family that contains only a BH3 domain and that activates apoptosis.

binding-change mechanism
The mechanism by which the β subunits of the F_1 ATP synthase convert ADP and P_i to ATP.

biomolecular condensate
A specialized subunit of a cell, not bounded by a membrane. Often compared to a liquid droplet, it is chemically and physically distinct from its surroundings and is formed as a consequence of a liquid-liquid phase separation or aggregation of proteins and RNAs.

bi-oriented
Indicates that the kinetochores of **sister chromatids** have attached to microtubules emanating from opposite spindle poles.

blastocyst
Stage of mammalian embryo composed of ≈64 cells that have separated into two cell types — **trophectoderm**, which will form extra-embryonic tissues; and the **inner cell mass**, which gives rise to the embryo proper; stage that implants in the uterine wall and corresponds to the *blastula* of other animal embryos. (Figure 22-2)

blastopore
The first opening that forms during embryogenesis of bilaterally symmetric animals, which later becomes the gut. This opening may become either the mouth or the anus.

bromodomain
A protein domain of ∼120 amino acids that binds acetylated lysine; found in chromosome-associated proteins involved in transcriptional activation.

buffer

A solution of the acid (HA) and base (A^-) form of a compound that undergoes little change in pH when small quantities of strong acid or base are added at pH values near the compound's pK_a.

Burkitt's lymphoma
A cancer of B cells caused either by infection with the Epstein-Barr virus or by a chromosomal translocation that causes constitutive activation of the *myc* oncogene.

C_4 pathway
The phosphoenolpyruvate carboxylase-mediated, two-step pathway of CO_2 fixation in mesophyll and bundle sheath cells in plants.

cadherins
A family of dimeric **cell-adhesion molecules** that aggregate in adherens junctions and desmosomes and mediate Ca^{2+}-dependent cell-cell homophilic interactions. (Figure 20-2)

Calvin cycle
See **carbon fixation**.

cancer stem cell
A subpopulation of cancer cells within a tumor that is capable of seeding new tumors.

capsid
The outer proteinaceous coat of a **virus**, formed by multiple copies of one or more protein subunits and enclosing the viral nucleic acid.

CapZ protein
A protein that can bind the (+) end of an actin filament to cap it from actin subunit addition or loss.

CAR T cells
Cytotoxic T cells genetically engineered to contain a chimeric antigen receptor (CAR) that targets an antigen specific to tumor cells.

carbohydrate
General term for certain polyhydroxyaldehydes, polyhydroxyketones, or compounds derived from these usually having the formula $(CH_2O)_n$. Primary type of compound used for storing and supplying energy in animal cells. (Figure 2-18)

carbon fixation
The major metabolic pathway that fixes CO_2 into carbohydrates during photosynthesis; also called the *Calvin cycle*. It is indirectly dependent on light but can occur both in the dark

and light. (Figure 12-40)

carboxy-terminal domain (CTD)
A domain in RNA polymerase II at the carboxyl end of its RPB1 subunit. The domain is made of seven amino acids repeated multiple times, with a consensus sequence of Tyr-Ser-Pro-Thr-Ser-Pro-Ser. (Figure 8-8, red squiggle)

carcinogen
Any chemical or physical agent that can cause cancer when cells or organisms are exposed to it.

carcinoma
A tumor originating from epithelial cells. Carcinomas are the most common type of cancer.

caspases
A class of vertebrate protein-degrading enzymes (proteases) that function in **apoptosis** and work in a cascade with each type activating the next. (Figures 22-36 and 22-42)

catabolism
Cellular degradation of complex molecules to simpler ones usually accompanied by the release of energy. *Anabolism* is the reverse process in which energy is used to synthesize complex molecules from simpler ones.

catalyst
A substance that increases the rate of a chemical reaction without undergoing a permanent change in its structure. Enzymes are proteins with catalytic activity, and ribozymes are RNAs that can function as catalysts. (Figure 3-23)

cation
A positively charged ion.

Cdc14 phosphatase
A dual-specificity protein phosphatase that triggers mitotic CDK inactivation at the end of mitosis.

Cdc25B phosphatase
One of a pair of dual-specificity phosphatases, along with Cdc25C, that de-phosphorylates CDKs on threonine 14 and tyrosine 15, thereby activating CDKs.

Cdc25C phosphatase
One of a pair of dual-specificity phosphatases, along with Cdc25B, that de-phosphorylates CDKs on threonine 14 and tyrosine 15, thereby activating CDKs.

Cdc42 protein
A small signaling GTPase protein that in its active state can bind the nucleation promoting factor WASp as part of the activation of the actin-nucleating Arp2/3 complex.

CDK-activating kinase (CAK)
Phosphorylates CDKs on a threonine residue near the active site. This phosphorylation is essential for CDK activity.

cDNA
DNA molecule copied from an mRNA molecule by reverse transcriptase and therefore lacking the introns present in the DNA of the genome.

cell-adhesion molecules (CAMs)
Proteins in the plasma membrane of cells that bind similar proteins on other cells, thereby mediating cell-cell adhesion. Four major classes of CAMs include the **cadherins, IgCAMs, integrins**, and **selectins**. (Figures 20-1 and 20-2)

cell cycle
Ordered sequence of events in which a eukaryotic cell duplicates its chromosomes and divides into two. The cell cycle normally consists of four phases: G_1 before DNA synthesis occurs; S when DNA replication occurs; G_2 after DNA synthesis; and M when **cell division** occurs, yielding two daughter cells. Under certain conditions, cells exit the cell cycle during G_1 and remain in the G_0 state as nondividing cells. (Figures 1-17 and 19-1)

cell division
See **symmetric cell division**.

cell junctions
Specialized regions on the cell surface through which cells are joined to each other or to the **extracellular matrix**. (Figure 20-11; Table 20-3)

cell line
A population of cultured cells, of plant or animal origin, that has undergone a genetic change allowing the cells to grow indefinitely.

cell lineage
A pathway of cell division and differentiation whereby an ancestral cell gives rise to multiple cell progeny; also refers to the developmental history of a particular cell traced back to the ancestral cell from which it arose.

cell migration
The process whereby cells can move across a substrate.

cell polarity
The ability of cells to organize their internal structure, resulting in changes of cell shape and generating regions of the plasma membrane with different protein and lipid compositions.

cell strain
A population of cultured cells, of plant or animal origin, that has a finite life span and eventually dies, commonly after 25–50 generations.

cell wall
A specialized, rigid **extracellular matrix** that lies next to the plasma membrane, protecting a cell and maintaining its shape; prominent in most fungi, plants, and prokaryotes but absent in most multicellular animals. (Figure 20-43)

cell-matrix adhesion receptor
A cell-surface receptor that binds one or more components of the **extracellular matrix** (ECM), often serving as a link between the ECM and intracellular signaling pathways and the cytoskeleton.

cell-surface receptor
Protein embedded in the plasma membrane that has an extracellular domain that binds an extracellular molecule(s), called a *ligand*. Many cell-surface receptors bind to extracellular signaling molecules; such binding induces conformational changes in the receptor, altering the activity of the receptor's intracellular domain, which transmits the signal to the interior of the cell.

cellular communication
The transfer of information via signaling molecules or ions from one cell to another.

cellular respiration
See **respiration**.

cellulose
A structural polysaccharide made of glucose units linked together by $\beta(1\rightarrow4)$ **glycosidic bonds**. It forms long microfibrils, which are the major component of the **cell wall** in plants.

centriole
Either of two cylindrical structures within the **centrosome** of animal cells and containing nine sets of triplet microtubules; structurally similar to a **basal body**. (Figure 18-6)

centromere

DNA sequence required for proper **segregation** of chromosomes during mitosis and meiosis; the region of mitotic chromosomes where the **kinetochore** forms and that appears constricted. (Figure 6-40)

centrosome (cell center)
Structure located near the nucleus of animal cells that is the primary **microtubule-organizing center (MTOC)**; it contains a pair of **centrioles** embedded in a protein matrix and duplicates before mitosis, with each centrosome becoming a spindle pole. (Figures 18-6 and 18-35a)

centrosome (spindle pole body)
An organelle that serves as the main microtubule organizing center for animal cells. Centrosomes are made from arrangements of two barrel-shaped clusters of microtubules, called **centrioles**, and a complex of proteins that help additional microtubules to form.

centrosome disjunction
During mitotic spindle formation, at the G_2/M transition the protein linker connecting the two original centrioles is dissolved and two separate centrosomes are created.

channels
Membrane proteins that transport water, ions, or small hydrophilic molecules across membranes down their concentration or electric potential gradients.

chaperone
Collective term for two types of proteins — *molecular chaperones* and *chaperonins* — that prevent misfolding of a target protein or actively facilitate proper folding of an incompletely folded target protein, respectively. (Figures 3-19 and 3-20)

checkpoint pathway
Surveillance mechanism that prevents initiation of each step in cell division until earlier steps on which it depends have been completed and mistakes that occurred during the process have been corrected.

chemical equilibrium
The state of a chemical reaction in which the concentration of all products and reactants is constant because the rates of the forward and reverse reactions are equal.

chemical potential energy
The energy stored in the bonds connecting atoms in molecules.

chemiosmosis

Process whereby an electrochemical proton gradient (pH plus electric potential) across a membrane is used to drive an energy-requiring process such as ATP synthesis; also called *chemiosmotic coupling*. See **proton-motive force**. (Figure 12-2)

chemokine
Any of numerous small, secreted proteins that function as chemotatic cues for leukocytes.

chemotaxis
Movement of a cell or organism toward or away from certain chemicals.

chlorophylls
A group of light-absorbing porphyrin pigments that are critical in **photosynthesis**. (Figure 12-35)

chloroplast
A specialized organelle in plant cells that is surrounded by a double membrane and contains internal chlorophyll-containing membranes (**thylakoids**) where the light-absorbing reactions of photosynthesis occur. (Figure 12-33)

cholesterol
A lipid containing the four-ring steroid structure with a hydroxyl group on one ring; a component of many eukaryotic membranes and the precursor of steroid hormones, bile acids, and vitamin D. (Figure 10-8c)

chromatid
One copy of a replicated chromosome, formed during the S phase of the cell cycle, that is joined to the other copy; also called **sister chromatid**. During mitosis, the two chromatids separate, each becoming a chromosome of one of the two daughter cells. (Figure 7-31)

chromatin
Complex of DNA, histones, and nonhistone proteins from which eukaryotic chromosomes are formed. Condensation of chromatin during mitosis yields the visible **metaphase** chromosomes. (Figure 7-22)

chromatin immunoprecipitation
A technique that localizes where a protein of interest associates with the genome.

chromatin-mediated repression
A condensed chromatin structure that sterically blocks transcription factors from interacting with the DNA.

chromosome

In eukaryotes, the structural unit of the genetic material consisting of a single, linear double-stranded DNA molecule and associated proteins. In most prokaryotes, a single, circular double-stranded DNA molecule constitutes the bulk of the genetic material. See also **chromatin** and **karyotype**.

cilium (pl. cilia)
Short, membrane-enclosed structure extending from the surface of eukaryotic cells and containing a core bundle of **microtubules**. Cilia usually occur in groups and beat rhythmically to move a cell (e.g., single-celled organism) or to move small particles or fluid along a surface (e.g., trachea cells). See also **axoneme** and **flagellum**.

circadian clock (circadian oscillator)
A group of proteins and mRNAs whose levels change in a regular pattern that is synchronized with the 24-hour solar clock; allows cells and organisms to anticipate the daily light/dark cycle.

cis (lateral) interactions
Binding interactions between molecules on the plasma membrane in one cell.

cisterna (pl. cisternae)
Flattened membrane-bounded compartment, as found in the Golgi complex and endoplasmic reticulum.

cisternal maturation
The process by which most proteins traverse the Golgi by sequential maturation of cisternal compartments.

citric acid cycle
A set of nine coupled reactions occurring in the matrix of the **mitochondrion** in which acetyl groups are oxidized, generating CO_2, and reduced intermediates used to produce ATP; also called *Krebs cycle* and *tricarboxylic acid (TCA) cycle.* (Figure 12-13)

classical cadherin
A classical cadherin is a member of the calcium-dependent cadherin family of cell-adhesion molecules (CAMs), also called *cell-cell adhesion receptors*, that bind directly to molecules expressed on an adjacent cell. Classical cadherins contain a short C-terminal cytosolic domain, a single transmembrane domain, and five extracellular cadherin domains that mediate cis and trans homophilic binding.

clathrin

A fibrous protein that with the aid of assembly proteins polymerizes into a lattice-like network at specific regions on the cytosolic side of a membrane, thereby forming a clathrin-coated pit that buds off to form a vesicle. (Figure 14-19; Table 14-1)

clone
(1) A population of genetically identical cells, viruses, or organisms descended from a common ancestor. (2) Multiple identical copies of a gene or DNA fragment generated and maintained via **DNA cloning**.

clustered protocadherin
A clustered protocadherin is a member of a subfamily of the cadherins that mediate cell-cell binding and play a role in establishing the complex wiring diagram of neurons in mammalian brains.

co-activator
A protein or protein complex required for transcription activation that does not bind directly to DNA. In contrast, an activator has a DNA-binding domain and binds directly to a DNA transcription control sequence.

codon
Sequence of three nucleotides in DNA or mRNA that specifies a particular amino acid during protein synthesis; also called *triplet*. Of the 64 possible codons, three are stop codons, which do not specify amino acids and cause termination of synthesis. (Table 5-1)

cofilin
An actin regulatory factor that can bind to and disassemble older regions of actin filaments.

cohesin
Protein complex that holds the replicated sister chromatids together.

collagen
A triple-helical glycoprotein rich in glycine and praline that is a major component of the **extracellular matrix** and connective tissues. The numerous subtypes differ in their tissue distribution and the extracellular components and cell-surface proteins with which they associate. (Figure 20-25; Table 20-4)

complement
A group of constitutive serum proteins that bind directly to microbial or fungal surfaces, thereby activating a proteolytic cascade that culminates in formation of the cytolytic *membrane attack complex*. (Figure 24-5)

complementary
(1) Referring to two nucleic acid sequences or strands that can form perfect **base pairs** with each other. (2) Describing regions on two interacting molecules (e.g., an enzyme and its substrate) that fit together in a lock-and-key fashion.

complementary DNA (cDNA)
See **cDNA**.

concentration gradient
In cell biology, a difference in the concentration of a substance in different regions of a cell or embryo or on different sides of a cellular membrane.

condensin
Protein complex that promotes chromosome condensation.

conformation
The precise shape of a protein or other macromolecule in three dimensions resulting from the spatial location of the atoms in the molecule. (Figure 3-13)

connective tissue
Tissue that contains mostly **extracellular matrix** with relatively sparsely distributed cells (i.e., the tissue underlying epithelial cells and their associated basal lamina).

connexins
A family of transmembrane proteins that form **gap junctions** in vertebrates. (Figure 20-22)

conservation
A DNA sequence is considered to be conserved if it has remained substantially unchanged in organisms that are evolutionarily so far apart that most sequences have lost recognizable similarity due to random mutations (the sequence is constrained to perform some important function and so mutational variations have been systematically weeded out by natural selection).

constitutive
Referring to the continuous production or activity of a cellular molecule or the continuous operation of a cellular process (e.g., constitutive secretion) that is not regulated by internal or external signals.

constitutive secretion
The process in which in all cell types at least some proteins are secreted continuously.

contractile bundles

Bundles of **actin** and **myosin** in nonmuscle cells that function in cell adhesion (e.g., *stress fibers*) or cell movement (e.g., *contractile ring* in dividing cells).

co-oriented
Indicates that sister kinetochores attach to microtubules emanating from the *same* spindle pole rather than from opposite spindle poles.

COPI
A class of proteins that coat transport vesicles in the **secretory pathway**. COPI-coated vesicles move proteins from the Golgi to the endoplasmic reticulum and from later to earlier Golgi cisternae. (Table 14-1)

COPII
A class of proteins that coat transport vesicles in the **secretory pathway**. COPII-coated vesicles move proteins from the endoplasmic reticulum to the Golgi. (Table 14-1)

co-repressor
A protein or protein complex required for transcription repression that does not bind directly to DNA. In contrast, a repressor has a DNA-binding domain and binds directly to a DNA transcription control sequence.

cotranslational translocation
Simultaneous transport of a secretory protein into the endoplasmic reticulum as the nascent protein is still bound to the ribosome and being elongated. (Figure 13-6)

cotransport
Protein-mediated movement of an ion or small molecule across a membrane against a concentration gradient driven by coupling to movement of a second molecule down its concentration gradient in the same (**symport**) or opposite (**antiport**) direction. (Figure 11-3 and Table 11-1)

cotransporter
An integral membrane protein that transports two molecules across a membrane. The transport of one molecule along its electrochemical gradient powers the transport of the other against its electrochemical gradient.

covalent bond
Stable chemical force that holds the atoms in molecules together by sharing of one or more pairs of electrons. See also **noncovalent interaction**. (Figures 2-2 and 2-6)

CpG islands

Regions in vertebrate DNA ~100 to ~1000 bp that have an unusually high occurrence of the sequence CG. Many CpG islands function as promoters for transcription initiation, usually in both directions.

CRISPR
Named for *c*lustered *r*egularly *i*nterspaced *s*hort *p*alindromic *r*epeats, a mechanism used by many bacterial cells to protect against invasion by foreign DNA. The mechanism has been utilized in the lab to edit the genomic DNA of metazoan organisms.

cristae
Sheet-like and tubelike invaginations that extend from the boundary membrane into the center of the mitochondrion.

critical cell size
The cell size at which cells can enter the cell cycle. The critical cell size varies with nutrient availability.

critical concentration (C_c)
The concentration of actin monomers above which actin filaments can be formed and below which they disassemble.

cross-exon recognition complex
Large assembly including RNA-binding SR proteins and other components that helps delineate exons in the **pre-mRNAs** of higher eukaryotes and ensure correct **RNA splicing**. (Figure 9-16)

cryoelectron microscopy
A technique whereby a sample is rapidly frozen in water and then imaged at very low temperature in a transmission electron microscope.

cyclic guanosine monophosphate (cGMP)
A **second messenger** that opens cation channels in rod cells and activates protein kinase G in vascular smooth muscle and other cells. (Figures 15-5 and 15-18)

cytochromes
A group of colored, heme-containing proteins, some of which function as **electron carriers** during cellular respiration and photosynthesis. (Figure 12-17a)

cytokine
Any of numerous small, secreted proteins (e.g., erythropoietin, G-CSF, interferons, interleukins) that bind to cell-surface receptors on blood and immune-system cells to trigger their differentiation or proliferation.

cytokine receptor
Member of major class of cell-surface signaling receptors, including those for erythropoietin, growth hormone, interleukins, and interferons. Ligand binding leads to activation of cytosolic JAK kinases associated with the receptor, thereby initiating intracellular signaling pathways. (Figure 16-21)

cytokinesis
The division of the cytoplasm following **mitosis** to generate two daughter cells, each with a nucleus and cytoplasmic organelles. (Figure 17-34)

cytoplasm
Viscous contents of a cell that are contained within the plasma membrane but, in eukaryotic cells, outside the nucleus.

cytoskeleton
Network of fibrous elements, consisting primarily of **microtubules, microfilaments**, and **intermediate filaments**, found in the cytoplasm of eukaryotic cells. The cytoskeleton provides organization and structural support for the cell and permits directed movement of organelles, chromosomes, and the cell itself. (Figures 17-1, 17-2, and 18-1)

cytosol
Unstructured aqueous phase of the cytoplasm excluding organelles, membranes, and insoluble cytoskeletal components.

cytosolic face
The face of a cell membrane directed toward the cytosol. (Figure 10-5)

cytosolic stress granules
Aggregates of proteins and RNA in the cell cytosol that are formed when the cells are under certain types of stresses.

dalton
Unit of molecular mass approximately equal to the mass of a hydrogen atom (1.66×10^{-24} g).

denaturation (melting)
Drastic alteration in the **conformation** of a protein or nucleic acid due to disruption of various noncovalent interactions caused by heating or exposure to certain chemicals; usually results in loss of biological function.

dendrite

Process extending from the cell body of a neuron that is relatively short and typically branched and receives signals from **axons** of other neurons. (Figure 23-1)

dendritic cells
Phagocytic professional **antigen-presenting cells** that reside in various tissues and can detect broad patterns of pathogen markers via their **Toll-like receptors**. After internalizing antigen at a site of tissue injury or infection, they migrate to lymph nodes and initiate activation of **T cells**. (Figure 24-7)

deoxyribonucleic acid
See **DNA**.

depolarization
Decrease in the cytosolic-face negative electric potential that normally exists across the plasma membrane of a cell at rest, resulting in a less inside-negative or an inside-positive **membrane potential**.

depolarized state
When the potential across the plasma membrane is less negative than that of the normal resting state of the cell.

desmin
A subunit protein of the intermediate filaments found in smooth and skeletal muscle.

desmosomal cadherin
Specialized cadherins, desmoglein and desmocollin, that mediate cell-cell binding in anchoring junctions called **desmosomes**.

desmosome
An anchoring junction that functions as a snap-like point of contact between cells.

deuterostomes
A group of bilaterally symmetric animals whose anus develops close to the blastopore and has a dorsal nerve cord. This group includes all chordates (fish, amphibians, reptiles, birds, and mammals) and echinoderms (sea stars, sea urchins).

Dicer
A cytoplasmic double-stranded RNA-specific RNase. Part of the RNase III family, the Dicer enzyme cleaves double-stranded RNA and pre-miRNA into **siRNA** and **miRNA**, respectively. (Figure 9-32)

dipole

A positive charge separated in space from an equal but opposite negative charge.

dipole moment
A quantitative measure of the extent of charge separation, or strength, of a dipole, which for a chemical bond is the product of the partial charge on each atom and the distance between the two atoms.

disaccharide
A small carbohydrate (sugar) composed of two monosaccharides covalently joined by a **glycosidic bond**. (Figure 2-19)

dislocation
Transport of a defective secretory protein from the ER lumen to the cytosol, where it can be degraded by the proteasome.

dissociation constant (K_d)
See **equilibrium constant**.

disulfide bond (—S—S—)
A common covalent linkage between the sulfur atoms on two cysteine residues in different polypeptides or in different parts of the same polypeptide.

diversity
The entire set of antigen-specific receptors encoded by an immune system.

DNA (deoxyribonucleic acid)
Long linear polymer, composed of four kinds of deoxyribose **nucleotides**, that is the carrier of genetic information. See also **double helix**. (Figure 5-3)

DNA-binding domain
The domain of a transcription factor that binds specific, closely related DNA sequences.

DNA cloning
Recombinant DNA technique in which specific cDNAs or fragments of genomic DNA are inserted into a cloning **vector**, which then is incorporated into cultured host cells and maintained during growth of the host cells; also called *gene cloning*. (Figure 6-14)

DNA library
Collection of cloned DNA molecules consisting of fragments of the entire genome (*genomic library*) or of DNA copies of all the mRNAs produced by a cell type (*cDNA library*) inserted into a suitable cloning **vector**.

DNA microarray
An ordered set of thousands of different nucleotide sequences arrayed on a microscope slide or other solid surface; can be used to determine patterns of gene expression in different cell types or in a particular cell type at different developmental stages or under different conditions. (Figures 5-11 and 5-12)

DNA-PK (DNA-dependent protein kinase)
A protein kinase critical for DNA damage signaling and repair, along with the pair of homologous protein kinases **ATM** and **ATR**.

DNA polymerase
An enzyme that copies one strand of DNA (the template strand) to make the complementary strand, forming a new double-stranded DNA molecule. All DNA polymerases add deoxyribonucleotides one at a time in the $5' \rightarrow 3'$ direction to the $3'$ end of a short preexisting primer strand of DNA or RNA.

DNA transposon
A mobile element that transposes directly as DNA (also referred to as just a *transposon*). See also **retrotransposon**. (Figure 7-8)

DNase I footprinting
A technique that reveals the region of a DNA sequence where a transcription factor binds. Samples of a DNA fragment that has been labeled with a radioactive molecule at one end of one strand are digested under carefully controlled conditions in the presence and absence of a DNA-binding protein. Then the DNA is purified away from all of the protein, denatured and electrophoresed, and the resulting gel is subjected to autoradiography. The region protected by the bound protein appears as a gap, or footprint, in the array of bands resulting from digestion in the absence of the protein.

dolichol phosphate
An isoprenoid-based lipid in the ER membrane onto which the precursor for *N*-linked oligosaccharide modification is assembled.

domain
A region of protein that has a distinct, and often independent, function or structure, or that has a distinct topology relative to the rest of the protein.

dominant
In genetics, referring to that allele of a gene expressed in the **phenotype** of a heterozygote; the nonexpressed allele is **recessive**; also refers to the phenotype associated with a

dominant allele. Mutations that produce dominant alleles generally result in a gain of function. (Figure 6-2)

double helix
The most common three-dimensional structure for cellular DNA in which the two polynucleotide strands are antiparallel and wound around each other with complementary bases hydrogen-bonded. (Figure 5-3)

downstream
(1) For a gene, the direction RNA polymerase moves during transcription, which is toward the end of the template DNA strand with a 5′-hydroxyl group. Nucleotides downstream from the +1 position (the first transcribed a nucleotide) are designated +2, +3, and so on. (2) Events that occur later in a cascade of steps (e.g., signaling pathway). See also **upstream**.

Drosha
This nuclear RNase specific for double-stranded RNA acts in concert with a nuclear double-stranded RNA-binding protein called *DGCR8* in humans and cleaves the hairpin region out of the long precursor RNA, generating a pre-miRNA.

duty ratio
The total cycle time of a motor can be divided into time spent in strong states (tending to remain attached, able to exert and support force), and time spent in weak states (tending to detach); the ratio of the two is the duty ratio.

dynamin
A GTPase that can drive the final step of pinching off of a clathrin-coated vesicle from a donor membrane.

dyneins
A class of **motor proteins** that use the energy released by ATP hydrolysis to move toward the (−) end of **microtubules**. Dyneins can transport vesicles and organelles, are responsible for the movement of cilia and flagella, and play a role in chromosome movement during mitosis. (Figures 18-24 and 18-25)

effector
Ultimate component of a signal transduction pathway that elicits a response to the transmitted signal.

effector protein
An enzyme, transcription factor, or cytoskeletal protein whose activation leads directly to changes in cellular activities. (Figure 15-1)

EF hand
A type of helix-loop-helix **structural motif** that occurs in many Ca^{2+}-binding proteins such as **calmodulin**. (Figure 3-7b)

efferent neurons
Nerves that transmit signals from the central nervous system to peripheral tissues such as muscles and endocrine cells.

elastin
A protein that forms the rubberlike amorphous core of elastic fibers.

electric potential
The energy associated with the separation of positive and negative charges. An electric potential is maintained across the plasma membrane of nearly all cells.

electrochemical gradient
The driving force that determines the energetically favorable direction of transport of an ion (or charged molecule) across a membrane. It represents the combined influence of the ion's **concentration gradient** across the membrane and the **membrane potential**.

electron-transport chain
Set of four large multiprotein complexes in the inner mitochondrial membrane plus diffusible cytochrome *c* and coenzyme Q through which electrons flow from reduced electron donors (e.g., NADH) to O_2. Each member of the chain contains one or more bound electron carriers. (Figure 12-19)

electrophoresis
Any of several techniques for separating macromolecules based on their migration in a gel or other medium subjected to a strong electric field. (Figure 3-41)

electrophoretic mobility shift assay (EMSA)
A common affinity electrophoresis technique used to study protein–DNA or protein–RNA interactions.

elongation
Addition of nucleotides to a growing polynucleotide chain, as templated by a complementary DNA coding strand. (Figure 5-24)

elongation factor (EF)
One of a group of nonribosomal proteins required for continued **translation** of mRNA (protein synthesis) following initiation. (Figure 5-37)

embryonic stem (ES) cells
A line of cultured cells derived from very early embryos that can differentiate into a wide range of cell types either in vitro or after reinsertion into a host embryo. (Figure 22-4)

endergonic
Referring to reactions and processes that have a positive *G* and thus require an input of **free energy** in order to proceed; opposite of **exergonic**.

endocytic pathway
Cellular pathway involving **receptor-mediated endocytocis** that internalizes extracellular materials too large to be imported by membrane transport proteins and to remove receptor proteins from the cell surface as a way to down-regulate their activity. (Figure 14-29)

endocytosis
General term for uptake of extracellular material by invagination of the plasma membrane; includes **receptor-mediated endocytosis, phagocytosis**, and pinocytosis.

endoderm
Innermost of the three primary cell layers of the animal embryo; gives rise to the gut and most of the respiratory tract.

endoplasmic reticulum (ER)
Network of interconnected membranous structures within the cytoplasm of eukaryotic cells contiguous with the outer nuclear envelope. The *rough ER,* which is associated with **ribosomes**, functions in the synthesis and processing of secreted and membrane proteins; the *smooth ER*, which lacks ribosomes, functions in lipid synthesis. (Figure 1-13)

endosome
One of two types of membrane-bounded compartments: *early* endosomes (or endocytic vesicles), which bud off from the plasma membrane during receptor-mediated endocytosis, and late endosomes, which have an acidic internal pH and function in sorting of proteins to **lysosomes**. (Figures 14-2 and 14-29)

endosymbiont/endosymbiont hypothesis
Bacterium that resides inside a eukaryotic cell in a mutually beneficial partnership. According to the endosymbiont hypothesis, both mitochondria and chloroplasts evolved from endosymbionts. (Figure 12-7)

endothelium
The thin layer of cells that lines the interior surface of blood and lymphatic vessels.

endothermic
Referring to reactions and processes that have a positive change in **enthalpy**, ΔH, and thus must absorb heat in order to proceed; opposite of **exothermic**.

energy charge
A measure of the fraction of total adenosine phosphates that have "high-energy" phosphoanhydride bonds, which is equal to $([\mathrm{ATP}] + 0.5\,[\mathrm{ADP}])/([\mathrm{ATP}] + [\mathrm{ADP}] + [\mathrm{AMP}])$.

enhanceosome
Large nucleoprotein complex that assembles from transcription factors (activators and repressors) as they bind cooperatively to their binding sites in an **enhancer** with the assistance of DNA-bending proteins. (Figure 8-29)

enhancer
A regulatory sequence in eukaryotic DNA that may be located at a great distance from the gene it controls or even within the coding sequence. Binding of specific proteins to an enhancer modulates the rate of transcription of the associated gene. (Figure 8-18)

enthalpy (ΔH)
Heat; in a chemical reaction, the enthalpy of the reactants or products is equal to their total bond energies.

entropy (S)
A measure of the degree of disorder or randomness in a system; the higher the entropy, the greater the disorder.

enzyme
A protein that catalyzes a particular chemical reaction involving a specific **substrate** or small number of related substrates.

epigenetic modification
Referring to a process that affects the expression of specific genes and is inherited by daughter cells but does not involve a change in DNA sequence.

epigenetics
The study of inherited changes in the phenotype of a cell that do not result from changes in DNA sequence.

epinephrine
A catecholamine secreted by the adrenal gland and some neurons in response to stress; also called *adrenaline*. It functions as both a hormone and neurotransmitter, mediating

fight-or-flight responses (e.g., increased blood glucose levels and heart rate).

epithelia
Tightly packed, sheet-like layers of cells that act as a regulatable, selectively permeable barrier and permit the generation of chemically and functionally distinct compartments in an organism.

epithelial-to-mesenchymal transition (EMT)
Describes a developmental program during which epithelial cells acquire the characteristics of mesenchymal cells. Cells lose adhesive properties and acquire motility.

epithelium (pl. epithelia)
Sheet-like covering, composed of one or more layers of tightly adhering cells, on external and internal body surfaces. (Figure 20-10)

epitope
The part of an antigen molecule that binds to an antigen-specific receptor on B or T cells or to antibody. Large protein antigens usually possess multiple epitopes that bind to antibodies of different specificity.

equilibrium constant (K_{eq})
Ratio of forward and reverse rate constants for a reaction. For a binding reaction, $\mathrm{A} + \mathrm{B} \rightleftarrows \mathrm{AB}$, the association constant (K_{a}) equals K, and the dissociation constant (K_{d}) equals $1/K$.

ERAD (ER-associated protein degradation)
The process by which defective proteins that are unable to fold properly in the ER are degraded. ERAD usually involves dislocation to the cytosol and degradation by the proteasome.

erythropoietin (Epo)
A **cytokine** that triggers production of red blood cells by inducing the proliferation and differentiation of erythroid progenitor cells in the bone marrow. (Figure 16-18b)

ESCRT proteins
A set of proteins responsible for the inward budding of membrane-bounded vesicles into the interior of a multivesicular endosome.

E2F transcription factor (E2F)
A small family of transcription factors related to G_1 CDKs.

E2F transcription factor complex

Transcription factor that promotes the transcription of G_1/S phase cyclins and many other genes whose function is required for S phase.

euchromatin
Less condensed portions of **chromatin** present in interphase chromosomes; includes most transcriptionally active regions. See also **heterochromatin**. (Figure 7-28a)

eukaryotic cells
Cells with a nucleus bounded by a nuclear membrane.

eukaryotic translation initiation factors (eIFs)
Proteins required to initiate protein synthesis in eukaryotic cells. (Figure 5-36)

excision-repair system, DNA
One of several mechanisms for repairing DNA damage due to spontaneous depurination or deamination or exposure to **carcinogens**. These repair systems normally operate with a high degree of fidelity and their loss is associated with increased risk for certain cancers.

excitatory receptor
A membrane receptor that, upon neurotransmitter binding, triggers depolarization and increases the likelihood that the target neuron will fire an action potential.

excitatory synapse
A synapse in which the neurotransmitter induces a depolarization of the postsynaptic cell, favoring generation of an action potential.

exergonic
Referring to reactions and processes that have a negative ΔG and thus release free energy as they proceed; opposite of **endergonic**.

exocytosis
Release of intracellular molecules (e.g., hormones, matrix proteins) contained within a membrane-bounded vesicle by fusion of the vesicle with the plasma membrane of a cell.

exon
Segment of a eukaryotic gene (or of its primary transcript) that reaches the cytoplasm as part of a mature mRNA, rRNA, or tRNA molecule. See also **intron**.

exon shuffling
Evolutionary process for creating new genes (i.e., new combinations of exons) from preexisting ones by recombination between introns of two separate genes or by transposition of mobile DNA elements. (Figures 7-18 and 7-19)

exoplasmic face
The face of a cell membrane directed away from the cytosol. (Figure 10-5)

exosome
Large exonuclease-containing complex that degrades spliced-out introns and improperly processed pre-mRNAs in the nucleus or mRNAs with shortened poly(A) tails in the cytoplasm. (Figure 9-1)

exothermic
Referring to reactions and processes that have a negative change in **enthalpy**, ΔH, and thus release heat as they proceed; opposite of **endothermic**.

extracellular matrix
A complex interdigitating meshwork of proteins and polysaccharides secreted by cells into the spaces between them.

F_0F_1 complex
See **ATP synthase**.

facilitated transport
Protein-aided transport of an ion or small molecule across a cell membrane down its concentration gradient at a rate greater than that obtained by **simple diffusion**; also called *facilitated diffusion*. (Table 11-1)

FAD (flavin adenine dinucleotide)
A small organic molecule that functions as an electron carrier by accepting two electrons from a donor molecule and two H^+ from the solution. (Figure 2-33b)

fatty acid
Any long hydrocarbon chain that has a carboxyl group at one end; a major source of energy during metabolism and a precursor for synthesis of phospholipids, triglycerides, and cholesteryl esters. (Figure 2-21; Table 2-4)

Fc receptor (FcR)
A receptor that recognizes the Fc portion of immunoglobulins.

feedback
A regulatory process in a signaling pathway where one of the later components of the pathway, when activated, inhibits (negative feedback) or activates (positive feedback) an earlier step in the pathway.

feedback repression

An essential feature of most signal transduction pathways, in which an end product of a signaling pathway blocks an early step in that pathway.

fermentation
The conversion of some of the energy in organic molecule nutrients such as glucose into ATP via their oxidation into organic molecule "waste" products such as lactic acid or ethanol, typically involving the simultaneous cyclical reduction and oxidation of $NAD^+/NADH$.

FG-nucleoporins
Proteins on the inner surface of the nuclear pore complex with a globular domain that forms part of the pore structure and a random-coil domain of hydrophilic amino acids punctuated by short repeats rich in phenylalanine and glycine.

fibrillar collagen
These molecules, characterized by long, thin, triple helices composed of Gly-X-Y amino acid triplet repeats, bundle together into microfibrils that associate into higher order fibrils and fibers in the **extracellular matrix** of connective tissues.

fibroblast
A common type of connective tissue cell that secretes **collagen** and other components of the **extracellular matrix**; migrates and proliferates during wound healing and in tissue culture.

fibronectin
An abundant **multi-adhesive matrix protein** that occurs in numerous isoforms, generated by alternative splicing, in various cell types. Binds many other components of the **extracellular matrix** and to integrin adhesion receptors. (Figure 20-34)

filopodia
Slender surface protrusions containing a bundle of actin filaments.

fission (mitochondrial)
The process by which one mitochondrion breaks apart to form two distinct mitochondria.

5′ capping
The initial step in mRNA processing, shortly after RNA polymerase II initiates transcription at the first nucleotide of the first exon of a gene, a protective cap composed of 7-methylguanosine and methylated ribose is added to the 5′ end of eukaryotic mRNAs. (Figure 9-2)

flagellum (pl. flagella)

Long locomotory structure (usually one per cell) extending from the surface of some eukaryotic cells (e.g., **sperm**), whose whiplike bending propels the cell through a fluid medium. Bacterial flagella are smaller and much simpler structures. See also **axoneme** and **cilium**. (Figure 18-31)

flavin adenine dinucleotide
See **FAD**.

flippase
Protein that facilitates the movement of membrane lipids from one leaflet to the other leaflet of a phospholipid bilayer. (Figure 11-15)

fluorescence-activated cell sorter (FACS)
A machine, based on flow cytometry, that separates cells having different levels of fluorescence. (Figure 4-1)

fluorescence in situ hybridization (FISH)
Any of several related techniques for detecting specific DNA or RNA sequences in cells and tissues by treating samples with fluorescent probes that hybridize to the sequence of interest and observing the samples by fluorescence microscopy.

fluorescent staining
General technique for visualizing cellular components by treating cells or tissues with a fluorescent dye–labeled agent (e.g., antibody) that binds specifically to a component of interest and observing the sample by fluorescence microscopy.

focal adhesion/focal contact
Sites of integrin-based cell-matrix adhesion that anchor cells to the **extracellular matrix** and contribute to cellular shape, signaling, force transmission, and movement.

formin
A regulated actin-nucleating protein responsible for generating long, unbranched actin filaments.

Förster resonance energy transfer (FRET)
A technique whereby the distance between two fluorescent proteins — fused to separate proteins or within a single protein — can be determined by measuring the efficiency of energy transfer from one fluorescent protein to the other.

free-energy change (ΔG)
The difference in the Gibbs free energy (G) between the products and reactants for a reaction ($G_{products} - G_{reactants}$), where the free energy is a measure of the potential energy

of a system, which is a function of the enthalpy (H) and entropy (S).

functional complementation
The ability of a functional gene introduced into a mutant cell or organism to restore the activity missing in the mutant.

fusion (mitochondrial)
The process in which two distinct mitochondria merge together.

G_1, G_2 phase
See **cell cycle**.

G_1 cyclin-CDKs
Cyclin-CDK complexes that promote entry into the cell cycle.

G_1/S phase cyclin CDKs
Cyclin-CDK complexes that promote entry into the cell cycle together with G_1 CDKs.

gap junction
Protein-lined channel connecting the cytoplasms of adjacent animal cells that allows passage of ions and small molecules between the cells. See also **plasmodesmata**. (Figure 20-22)

gene
Physical and functional unit of heredity, which carries information from one generation to the next. In molecular terms, it is the entire DNA sequence — including **exons, introns**, and **transcription-control regions** — necessary for production of a functional polypeptide or RNA. See also **transcription unit**.

gene control
All of the mechanisms involved in regulating **gene expression**. Most common is regulation of transcription, although mechanisms influencing the processing, stabilization, and translation of mRNAs help control expression of some genes.

gene conversion
A type of DNA recombination in which one DNA sequence is converted to the sequence of a second, homologous DNA sequence in the same cell.

gene family
Set of genes that arose by duplication of a common ancestral gene and subsequent divergence due to small changes in the nucleotide sequence.

gene knockout
Selective inactivation of a specific gene by replacing it with a nonfunctional (disrupted) allele in an otherwise normal organism.

general import pore
The pore in the outer mitochondrial membrane through which proteins pass as they are imported into mitochondria.

genetic code
The set of rules whereby nucleotide triplets (**codons**) in DNA or RNA specify amino acids in proteins. (Table 5-1)

genetic complementation
Restoration of a wild-type function in diploid heterozygous cells generated from haploid cells, each of which carries a mutation in a different gene whose encoded protein is required for the same biochemical pathway. Complementation analysis can determine if recessive mutations in two mutants with the same mutant phenotype are in the same or different genes. (Figure 6-7)

genome
Total genetic information carried by a cell or organism.

genome-wide association study (GWAS)
A statistical method based on linkage disequilibrium of identifying genes for human diseases or other traits that exhibit genetic heterogeneity or may be polygenic.

genomics
Comparative analyses of the complete genomic sequences from different organisms and determination of global patterns of gene expression; used to assess evolutionary relations among species and to predict the number and general types of RNAs produced by an organism.

genotype
Entire genetic constitution of an individual cell or organism, usually with emphasis on the particular alleles at one or more specific loci.

germ line
Cells in a multicellular organism that are set aside early in development to ultimately give rise to the cells that form the sperm and eggs.

glial cells

Supporting cells of nervous tissue that, unlike neurons, do not conduct electrical impulses. Of the four types, *Schwann cells* and *oligodendrocytes* produce **myelin sheaths**, *astrocytes* function in **synapse** formation, and *microglia* make **trophic factors** and serve in immune responses. (Figure 23-18)

glomeruli
In the brain, glomeruli are spherical structures consisting of a mass of presynaptic axonal terminals forming on postsynaptic dendrites. In the kidney, glomeruli are spherical structures consisting of capillary blood vessels through which blood is filtered into urine.

glucagon
A peptide hormone produced in the cells of pancreatic islets that triggers the conversion of glycogen to glucose by the liver; acts with **insulin** to control blood glucose levels.

gluconeogenesis
A biochemical pathway whereby glucose is synthesized from smaller molecules, often lactate, pyruvate, or certain amino acids.

GLUT4 storage vesicle (GSV)
An intracellular vesicle with GLUT4 transporters in its membrane. Upon insulin stimulation, GSVs fuse with the cell membrane, exposing GLUT4s to the extracellular space, from which they can transport glucose into the cytosol.

glycogen
A very long, branched polysaccharide, composed exclusively of glucose units, that is the primary storage carbohydrate in animals; found primarily in liver and muscle cells.

glycolipid
Any lipid to which a short carbohydrate chain is covalently linked; commonly found in the plasma membrane.

glycolysis
Metabolic pathway in which sugars are degraded anaerobically to lactate or pyruvate in the cytosol with the production of ATP. (Figure 12-3)

glycoprotein
Any protein to which one or more oligosaccharide chains are covalently linked. Most secreted proteins and many membrane proteins are glycoproteins.

glycosaminoglycan (GAG)
A long, linear, highly charged polymer of repeating disaccharides in which many residues often are sulfated. GAGs are major components of the **extracellular matrix**, usually as

components of **proteoglycans**. (Figure 20-30)

glycosidic bond
The covalent linkage between two monosaccharide residues formed when a carbon atom in one sugar reacts with a hydroxyl group on a second sugar with the net release of a water molecule (dehydration). (Figure 2-13)

G protein–coupled receptor (GPCR)
Member of a large class of cell-surface signaling receptors, including those for epinephrine, glucagon, and yeast mating factors. All GPCRs contain seven transmembrane α helices. Ligand binding leads to activation of a coupled trimeric G protein, thereby initiating intracellular signaling pathways. (Figures 15-16 and 15-17)

granulocyte colony–stimulating factor (G-CSF)
A cytokine that induces a granulocyte progenitor cell in the bone marrow to divide and differentiate into granulocytes.

group I introns
A type of self-splicing introns present in nuclear rRNA genes of protozoans.

group II introns
A type of self-splicing introns present in protein-coding genes and some rRNA and tRNA genes in mitochondria and chloroplasts of plants and fungi.

growth factor
An extracellular polypeptide molecule that binds to a cell-surface receptor, triggering an intracellular signaling pathway generally leading to cell proliferation.

growth hormone (GH)
A cytokine secreted by the anterior pituitary gland that stimulates proliferation of a variety of cells.

GTPase superfamily
Group of intracellular switch proteins that cycle between an inactive state with bound GDP and an active state with bound GTP. Includes the G_α subunit of trimeric (large) G proteins, monomeric (small) G proteins (e.g., **Ras**, Rab, Ran, and Rac), and certain **elongation factors** used in protein synthesis. (Figure 3-35)

heat-shock gene
A gene that encodes molecular chaperones that refold denatured proteins and other proteins that help the cell deal with the effects of heat shock.

heat-shock proteins (HSPs)
Proteins that perform several processes that protect cells from the consequences of unfolding a large fraction of cellular proteins.

heat-shock response
A cellular response to an elevated temperature or other stress, whereby synthesis of chaperones is induced to protect the cell from accumulation of unfolded or misfolded proteins.

Hedgehog (Hh)
A family of secreted signaling proteins that are important regulators of the development of most tissues and organs in diverse animal species. Mutations in Hh signal transduction components are implicated in human cancer and birth defects. The receptor is the Patched transmembrane protein. (Figures 16-28 and 16-29)

helicase
(1) Any enzyme that moves along a DNA duplex using the energy released by ATP hydrolysis to separate (unwind) the two strands; required for DNA replication. (2) Activity of certain initiation factors that can unwind the secondary structures in mRNA during initiation of translation.

helix-loop-helix
A conserved DNA-binding **structural motif**, consisting of two α helices connected by a short loop, that is found in many dimeric eukaryotic transcription factors. (Figure 8-25d)

helix-turn-helix
A structural motif in which two alpha helices are connected by a short stretch of connecting residues, sometimes also called a *loop*. Helix-turn-helix/helix-loop-helix structural motifs can perform various functions, including binding calcium and binding DNA.

HER family
Group of receptors, belonging to the receptor tyrosine kinase (RTK) class, that bind to members of the epidermal growth factor (EGF) family of signaling molecules in humans. Overexpression of HER2 protein is associated with some breast cancers. (Figure 16-7)

heredity
The transfer of genetically determined characteristics from one generation to the next.

heterochromatin

Regions of **chromatin** that remain highly condensed and transcriptionally inactive during interphase. (Figure 7-28a)

heterophilic binding
The binding of one cell-adhesion molecule (CAM or cell-cell adhesion receptor) to a different class of CAM on an adjacent cell.

heterotrimeric G proteins
A class of GTPase switch proteins, composed of alpha, beta, and gamma polypeptides, that bind to and are activated by certain cell-surface receptors. When activated, heterotrimeric G proteins release GDP and bind GTP.

heterotypic adhesion
The CAM-mediated binding of one type of cell to a different type of adjacent cell.

heterozygous
Referring to a diploid cell or organism having two different **alleles** of a particular gene.

hexose
A six-carbon **monosaccharide**.

Hif-1α
A transcription factor in metazoans whose level increases in response to a reduction in oxygen.

Hippo pathway
A signal transduction pathway that controls cell growth in the context of tissues.

Hippocampus
A region of the brain, located in the temporal lobe, that is involved in memory.

histone
One of several small, highly conserved basic proteins, found in the **chromatin** of all eukaryotic cells, that associate with DNA in the **nucleosome**. (Figure 7-20)

histone deacetylation
A process by which the lysine residues within the N-terminal tail protruding from the histone core of the nucleosome are deacetylated (an acetyl group is removed from a molecule) as part of gene regulation.

histone octamer
A symmetrical, roughly globular protein complex at the center of the nucleosome.

Holliday structure
An intermediate in DNA recombination with four DNA strands. (Figure 5-22)

homeodomain
Conserved DNA-binding **structural motif** (a helix-turn-helix) found in many developmentally important transcription factors.

homeostasis
The state of an organism or cell in which its internal state is maintained within certain functional limits.

homolog
A protein that shares a common ancestor, and therefore is similar in sequence and/or structure, with another protein.

homologs
Maternal and paternal copies of each morphologic type of chromosome present in a diploid cell; also called *homologues*.

homologous
See **homologs**.

homologous chromosome
One of the two copies of each morphologic type of chromosome present in a diploid cell; also called **homolog**. Each homolog is derived from a different parent.

homologous recombination
See **recombination**.

homology
Similarity in characteristics (e.g., protein and nucleic acid sequences or the structure of an organ) that reflects a common evolutionary origin. Proteins or genes that exhibit homology are said to be homologous and sometimes are called *homologs*. In contrast, *analogy* is a similarity in structure or function that does not reflect a common evolutionary origin.

homophilic binding
The binding of one cell-adhesion molecule (CAM or cell-cell adhesion receptor) to the same kind of CAM on an adjacent cell.

homotypic adhesion
The CAM-mediated binding of one type of cell to the same type of adjacent cell.

homozygous
Referring to a diploid cell or organism having two identical **alleles** of a particular gene.

hormone
Generally any extracellular substance that induces specific responses in target cells; specifically, those signaling molecules that circulate in the blood and mediate endocrine signaling.

hyaluronan
A large, highly hydrated **glycosaminoglycan (GAG)** that is a major component of the **extracellular matrix**; also called *hyaluronic acid* and *hyaluronate*. It imparts stiffness and resilience as well as a lubricating quality to many types of connective tissue. (Figure 20-30a)

hybridization
Association of two **complementary** nucleic acid strands to form double-stranded molecules, which can contain two DNA strands, two RNA strands, or one DNA and one RNA strand. Used experimentally in various ways to detect specific DNA or RNA sequences.

hybridoma
A **clone** of hybrid cells that are immortal and produce **monoclonal antibody**; formed by fusion of a normal antibody-producing B cell with a myeloma cell. (Figure 4-6)

hydrocarbon
Any compound containing only carbon and hydrogen atoms.

hydrogen bond
A **noncovalent interaction** between an atom (commonly oxygen or nitrogen) carrying a partial negative charge and a hydrogen atom carrying a partial positive charge. Important in stabilizing the conformation of proteins and in formation of **base pairs** between nucleic acid strands. (Figure 2-8)

hydropathy profile
A graphical plot of the relative hydrophobicity of the amino acids along the linear sequence of a protein, often used to identify hydrophobic transmembrane-spanning segments.

hydrophilic
Interacting effectively with water. See also **polar**.

hydrophobic
Not interacting effectively with water; in general, poorly soluble or insoluble in water. See also **nonpolar**.

hydrophobic effect
The tendency of nonpolar molecules or parts of molecules to associate with each other in aqueous solution so as to minimize their direct interactions with water; commonly called a *hydrophobic interaction* or *bond*. (Figure 2-11)

hyperpolarization
Increase in the magnitude of the cytosolic-face negative electric potential that normally exists across the plasma membrane of a cell at rest, resulting in a more negative **membrane potential**.

immune checkpoint pathways
Genetically engineered cytotoxic T cells that contain a chimeric antigen receptor (CAR) targeting a tumor-specific antigen.

immunoediting
The process in which tumor cells are initially eliminated by the immune system, reach a steady state plateau, and then eventually escape from immune control by a variety of molecular mechanisms.

immunity
A state of protection against pathogens and their products (toxins), afforded by the presence of immune cells (B cells, T cells) and their products (immunoglobulins, cytokines).

immunoblotting
Technique in which proteins separated by electrophoresis are attached to a nitrocellulose or other membrane, and specific proteins then are detected by use of labeled antibodies; also called *Western blotting*.

immunoglobulin
Any of the serum proteins, produced by fully differentiated **B cells**, that can function as antibodies; also occur in membrane-bound form as part of the **B-cell receptor**. Immunoglobulins are divided into five main classes (*isotypes*) that exhibit distinct functional properties. See also **antibody**. (Figures 24-9 and 24-10)

immunoglobulin fold
Evolutionarily ancient structural motif found in antibodies, the T-cell receptor, and numerous other eukaryotic proteins not directly involved in antigen-specific recognition; also called *Ig domain*. (Figure 24-13b)

immunoprecipitation (IP)

A technique that uses antibodies to separate a target molecule of interest from other molecules in a complex mixture in solution by cross-linking the target molecule into a large aggregate, resulting in the formation of an insoluble solid (precipitate) that can be easily separated and analyzed.

in situ hybridization
Any technique for detecting specific DNA or RNA sequences in cells and tissues by treating samples with single-stranded RNA or DNA probes that hybridize to the sequence of interest. (Figure 6-30)

induced pluripotent stem (iPS) cells
A mammalian cell with properties of an embryonic stem cell that is formed from a differentiated cell type by expression of one or more transcription factors or other genes that confer pluripotency.

inflammasome
A multisubunit protein complex involved in the production of pro-inflammatory cytokines in response to a noxious substance, which may either be bacterially derived or is composed of a (foreign) material such as asbestos or uric acid crystals.

inflammation
Localized response to injury or infection that leads to the activation of immune-system cells and their recruitment to the affected site; marked by the four classical signs of redness, swelling, heat, and pain. (Figure 24-7)

inhibitory receptor
A membrane receptor that, upon neurotransmitter binding, triggers hyperpolarization and decreases the likelihood that the target neuron will fire an action potential.

inhibitory synapse
A synapse in which the neurotransmitter induces a hyperpolarization of the postsynaptic cell, inhibiting generation of an action potential.

initiation
The process by which an RNA polymerase separates DNA strands and synthesizes the first phosphodiester bond of an RNA chain as templated by the DNA strand that enters the RNA polymerase active site. (Figure 5-24)

initiation factor (IF)
One of a group of nonribosomal proteins that promote the proper association of ribosomes and mRNA and are required for initiation of **translation** (protein synthesis). (Figure 5-36)

initiator
A DNA sequence that specifies transcription initiation within the sequence.

inner cell mass (ICM)
The part of an early embryo that will form the embryo proper but not the extra-embryonic tissues, including the placenta.

inner mitochondrial membrane
The highly invaginated membrane that lies immediately underneath the outer mitochondrial membrane and that comprises the boundary membrane, cristae, and crista junctions.

inside-out signaling
Transfer of information from the extracellular environment (e.g., via hormones, growth factors, and **extracellular matrix**) to inside the cell, usually mediated via transmembrane receptors and adapter proteins.

insulin
A protein hormone produced in the β cells of the pancreatic islets that stimulates uptake of glucose into muscle and fat cells; acts with **glucagon** to help regulate blood glucose levels. Insulin also functions as a growth factor for many cells.

integral membrane protein
Any protein that contains one or more hydrophobic segments embedded within the core of the **phospholipids bilayer**; also called *transmembrane protein*. (Figure 13-10)

integrins
A large family of heterodimeric transmembrane proteins that function as adhesion receptors, promoting cell-matrix adhesion, or as cell-adhesion molecules, promoting cell-cell adhesion. (Table 20-4)

interleukins
Large group of cytokines, some released in response to inflammation, that promote proliferation and functioning of T cells and antibody-producing B cells of the immune system.

intermediate filament
Cytoskeletal fiber (10 nm in diameter) formed by polymerization of related, but tissue-specific, subunit proteins, including **keratins, lamins**, and **neurofilaments**. (Figure 18-49; Table 18-2)

intermembrane space

The mitochondrial compartment between the inner and outer mitochondrial membranes, which is continuous with the spaces inside the cristae.

interneurons
Nerves that receive signals from other nerve cells and that in turn transmit signals to other nerve cells.

interspersed repeats
Sequences from transposons that occur at multiple sites throughout the genomes of multicellular animals and plants. (Figure 7-8; Table 7-1)

intrinsically disordered region (IDR)
A segment (region) of a protein that is very flexible, with no fixed, well-ordered conformation. If the entire protein has no fixed conformation, it is called an *intrinsically disordered protein (IDP)*.

intron
Part of a primary transcript (or the DNA encoding it) that is removed by splicing during RNA processing and is not included in the mature, functional mRNA, rRNA, or tRNA.

ionic bond (ionic interaction)
A **noncovalent interaction** between a positively charged ion (cation) and negatively charged ion (anion).

iron-response element–binding protein (IRE-BP)
4Fe-4S cluster-containing proteins that act as cellular iron sensors. When the cytosolic $Fe2^{+}$ concentration is too low, the iron in the Fe-S clusters dissociates from the protein and the IRE-BP becomes an active RNA-binding protein that binds to a specific nucleotide sequence within stem-loop secondary structures in the untranslated regions of the ferritin and transferrin receptor mRNAs. (Figure 9-35).

isoelectric point (pI)
The pH of a solution at which a dissolved protein or other potentially charged molecule has a net charge of zero and therefore does not move in an electric field. (Figure 3-42)

isoform
One of several forms of the same protein whose amino acid sequences differ slightly and whose general activities are similar. Isoforms may be encoded by different genes or by a single gene whose primary transcript undergoes **alternative splicing**.

isotype
The term for a class of immunoglobulins, based on their distinct biochemical properties.

karyotype
Number, sizes, and shapes of the entire set of **metaphase** chromosomes of a eukaryotic cell. (Chapter 7 opening figure)

keratins
A group of **intermediate filament** proteins found in epithelial cells that assemble into heteropolymeric filaments. (Figure 18-51)

kinase
An enzyme that transfers the terminal (γ) phosphate group from ATP to a substrate. Protein kinases, which phosphorylate specific serine, threonine, or tyrosine residues, play a critical role in regulating the activity of many cellular proteins. See also **phosphatases**. (Figure 3-34)

kinesins
A class of **motor proteins** that use energy released by ATP hydrolysis to move toward the (+) end of a **microtubule**. Kinesins can transport vesicles and organelles and play a role in chromosome movement during mitosis. (Figures 18-18 through 18-20)

kinetic energy
Energy of movement, such as the motion of molecules.

kinetochore
A multilayer protein structure located at or near the **centromere** of each mitotic chromosome from which microtubules extend toward the spindle poles of the cell; plays an active role in movement of chromosomes toward the poles during anaphase. (Figure 18-42)

lagging strand
One of the two daughter DNA strands formed at a **replication fork** as short, discontinuous segments (Okazaki fragments), which are synthesized in the $5' \rightarrow 3'$ direction and later joined. See also **leading strand**. (Figure 5-11)

lamellipodium
A two-dimensional network of actin filaments under the plasma membrane found at the leading edge of a motile cell.

laminin
Large heterotrimeric **multi-adhesive matrix protein** that is found in all **basal lamina**. (Figure 20-24)

lamins

A group of **intermediate filament** proteins that form a fibrous network, the **nuclear lamina**, on the inner surface of the nuclear envelope.

late endosome
See **endosome**.

lateral
See **basolateral**.

leading edge
The actin-rich region at the front of a migrating cell.

leading strand
One of the two daughter DNA strands formed at a **replication fork** by continuous synthesis in the $5' \rightarrow 3'$ direction. The direction of leading-strand synthesis is the same as movement of the replication fork. See also **lagging strand**. (Figure 5-11)

lectin
Any protein that binds tightly to specific sugars. Lectins assist in the proper folding of some glycoproteins in the endoplasmic reticulum and can be used in affinity chromatography to purify glycoproteins or as reagents to detect them in situ.

leucine zipper
A type of coiled-coil **structural motif** composed of two α helices that form specific homo- or heterodimers; common motif in many eukaryotic transcription factors. (Figures 3-7 and 8-25c)

leukemia
Cancer originating from one of the cell lineages responsible for the formation of red or white blood cells.

ligand
Any molecule, other than an enzyme **substrate**, that binds tightly and specifically to a macromolecule, usually a protein, forming a macromolecule-ligand complex.

light-harvesting complexes (LHCs)
Specialized proteins that use pigments such as chlorophylls to efficiently capture light energy and transmit it to reaction centers to generate high-energy electrons.

linkage
In genetics, the tendency of two different loci on the same chromosome to be inherited together. The closer two loci are, the lower the frequency of **recombination** between them

and the greater their linkage.

linker DNA
A "beads-on-a-string" organization of nucleosomes separated by a variable length (–50–150 bp) of DNA.

lipid-anchored membrane protein
Any protein that is tethered to a cellular membrane by one or more covalently attached lipid groups, which are embedded in the phospholipids bilayer. (Figure 9-19)

lipid raft
Microdomain in the plasma membrane that is enriched in cholesterol, sphingomyelin, and certain proteins.

liposome
Artificial spherical **phospholipid bilayer** structure with an aqueous interior that forms in vitro from phospholipids and may contain membrane proteins. (Figure 10-3c)

liquid chromatography
An analytical technique used to separate components in a liquid, usually by flowing the liquid across a surface whose properties result in the components of the liquid differentially interacting with the surface.

long interspersed elements (LINEs)
Abundant mobile elements in mammals generated by retrotransposons lacking long terminal repeats. (Figure 7-17)

long noncoding RNA (lncRNA)
RNAs that can be many kilobases in length and lack long open reading frames.

long terminal repeats (LTRs)
Direct repeat sequences, containing up to 600 base pairs, that flank the coding region of integrated retroviral DNA and viral **retrotransposons**.

loss of heterozygosity (LOH)
The process by which the heterozygous, wild-type copy of a gene is lost to uncover a loss-of-function mutation. This is particularly important in the expression of tumor-suppressor mutations.

low-density lipoprotein (LDL)
A class of lipoprotein, containing apolipoprotein B-100, that is a primary transporter of cholesterol in the form of cholesteryl esters between tissues, especially to the liver. (Figure

14-27)

lumen
The aqueous interior of an organelle.

lymphocytes
Two classes of white blood cells that can recognize foreign molecules (**antigens**) and mediate immune responses. B lymphocytes (B cells) are responsible for production of antibodies; T lymphocytes (T cells) are responsible for destroying virus and bacteria-infected cells, foreign cells, and cancer cells.

lymphomas
Tumors that originate in lymphocytes.

lysosome
Small organelle that has an internal pH of 4–5, contains hydrolytic enzymes, and functions in degradation of materials internalized by endocytosis and of cellular components in autophagy. (Figures 1-13 and 4-12)

M (mitotic) phase
See **cell cycle**.

macromolecule
A molecule containing a very large number of atoms; the three main types of biological macromolecules are proteins, nucleic acids, and polysaccharides.

macrophages
Phagocytic leukocytes that can detect broad patterns of pathogen markers via **Toll-like receptors**. They function as professional **antigen-presenting cells** and are a major source of **cytokines**.

major histocompatibility complex (MHC)
Set of adjacent genes that encode class I and class II MHC molecules and other proteins required for antigen presentation, as well as some complement proteins; called the *H-2 complex* in mice and the *HLA complex* in humans. (Figure 24-22)

malignant
Referring to a tumor or tumor cells that can invade surrounding normal tissue and/or undergo **metastasis**. See also **benign**.

mannose 6-phosphate (M6P)

A specific modification of *N*-linked oligosaccharides responsible for sorting of proteins in the trans-Golgi to the lysosome.

MAP kinase
Any of a family of protein kinases that are activated in response to cell stimulation by many different growth factors and that mediate cellular responses by phosphorylating specific transcription factors and other target proteins. (Figure 16-14)

matriglycan
A polymer of the Xyl-GlcUA disaccharide that is attached to proteins such as dystroglycan via a phosphorylated *O*-linked mannose attached to the protein and intervening short phosphorylated oligosaccharide. Matriglycan provides a binding site for ECM molecules, such as laminin and perlecan.

matrix
The lumen of the innermost compartment of the mitochondrion; also the fibrous proteins and carbohydrates external to a cell (called the ***extracellular matrix***).

matrix metalloproteases (MMPs)
Proteolytic enzymes that employ the metal zinc in their active sites. They operate in the extracellular space, where they cut proteins in the **extracellular matrix** and sometimes other proteins (e.g., some cell-surface receptors).

mechanoreceptors
A sensory cell that responds to a mechanical stress or movement; also the intracellular molecule that senses this stress or movement.

mechanosensor
(1) Any of several types of sensory structures that are embedded in various tissues and respond to touch, the positions and movements of the limbs and head, pain, and temperature. (2) An individual molecule or complex that responds to a mechanical stimulus by changing its shape and activity.

Mediator
A very large multiprotein complex that forms a molecular bridge between transcriptional activators bound to an **enhancer** and to RNA polymerase II bound at a **promoter**; functions as a coactivator in stimulating transcription. (Figures 8-36 and 8-37)

meiosis
In eukaryotes, a special type of cell division that occurs during maturation of germ cells; comprises two successive nuclear and cellular divisions with only one round of DNA

replication. Results in production of four genetically nonequivalent haploid cells (gametes) from an initial diploid cell. (Figure 6-3)

membrane contact sites (MCSs)
Sites at which membranes from two organelles bind to one another via their protein constituents.

membrane potential
Electric potential difference, expressed in volts, across a membrane due to the slight excess of positive ions (cations) on one side and negative ions (anions) on the other. (Figures 11-18 and 11-19)

membrane transport protein
Collective term for any integral membrane protein that mediates movement of one or more specific ions or small molecules across a cellular membrane regardless of the transport mechanism. (Figure 11-2)

memory
The ability of an antigen-experienced immune system to respond more rapidly to a re-exposure to that same antigenic stimulus.

meristem
Organized group of undifferentiated, dividing cells that are maintained at the tips of growing shoots and roots in plants. All the adult structures arise from meristems.

merotelic attachment
Indicates that a single kinetochore attaches to microtubules emanating from two opposite spindle poles.

messenger RNA
See **mRNA**.

metaphase
Stage of mitosis at which condensed chromosomes are aligned equidistant between the poles of the mitotic spindle but have not yet started to segregate toward the spindle poles. (Figures 18-38 and 18-39)

metastasis
Spread of cancer cells from their site of origin and establishment of areas of secondary growth.

micelle

A water-soluble spherical aggregate of phospholipids or other amphipathic molecules that form spontaneously in aqueous solution.

Michaelis constant (Km)
A parameter that describes the affinity of an enzyme for its substrate and equals the substrate concentration that yields the half-maximal reaction rate. A similar parameter describes the affinity of a transport protein for the transported molecule or the affinity of a receptor for its ligand. (Figure 3-25)

microexons
The most common and extensively conserved type of alternatively spliced exons in the central nervous system. They are unusually short: 3 to 27 bases long in multiples of three so that they maintain the reading frames of up- and downstream exons of more typical longer lengths (average ~150 bases in humans).

microfilament
Cytoskeletal fiber (≈7 nm in diameter) that is formed by polymerization of monomeric globular (G) actin; also called actin filament. Microfilaments play an important role in muscle contraction, cytokinesis, cell movement, and other cellular functions and structures. (Figure 17-4)

micro-RNA
See **miRNA**.

microsatellites
Simple-sequence repeated DNA sequences 1–13 bases long (most are <5 bp) that occur in the human genome in tandem repeats of 150 repeats or fewer.

microsome
The cellular fraction that contains ribosomes bound to the ER membrane. This fraction corresponds to the rough ER as characterized by electron microscopy.

microtubule
Cytoskeletal fiber (≈25 nm in diameter) that is formed by polymerization of **α, β-tubulin** monomers and exhibits structural and functional polarity. Microtubules are important components of cilia, flagella, the mitotic spindle, and other cellular structures. (Figures 18-2 and 18-3)

microtubule-associated protein (MAP)
Any protein that binds to microtubules and regulates their stability. (Figures 18-13 through 18-15)

microtubule-organizing center
See **MTOC**.

microvilli
Small, membrane-covered projections on the surface of an animal cell, containing cores of actin filaments. Numerous microvilli are present on the absorptive surface of intestinal epithelial cells, increasing the surface area for transport of nutrients. (Figures 17-4 and 20-10)

miRNA (micro-RNA)
Any of numerous small, endogenous cellular RNAs, 20–30 nucleotides long, that are processed from double-stranded regions of hairpin secondary structures in long precursor RNAs. A single strand of the mature miRNA associates with several proteins to form an **RNA-induced silencing complex (RISC)** that inhibits translation of a target mRNA to which the miRNA hybridizes imperfectly. Several miRNAs must hybridize to a single mRNA to inhibit its translation. See also **siRNA**. (9-32)

mitochondria-associated membranes (MAMs)
Portions of the endoplasmic reticulum that closely contact regions of mitochondria, influencing mitochondrial shape, function, and fission.

mitochondrial DNA
DNA present in the mitochondrial matrix that encodes a limited number of mitochondrial genes.

mitochondrial inner membrane
A single continuous membrane comprising the boundary membrane, mitochondrial cristae, and crista junctions that separates the intermembrane space and the mitochondrial matrix. The mitochondrial electron-transport chain and the ATP synthase reside in this membrane. Protons in the matrix are pumped across the inner membrane to generate a proton-motive force.

mitochondrial intermembrane space
The fluid-filled space between the mitochondrial outer and inner membranes.

mitochondrial matrix
The fluid-filled space surrounded by the mitochondrial inner membrane. Protons in the matrix are pumped across the inner membrane to generate a proton-motive force.

mitochondrial outer membrane

The continuous membrane surrounding mitochondria that permits transfer of molecules from the cytosol into the mitochondria and mediates direct contact with other organelles via membrane contact sites.

mitochondrion (pl. mitochondria)
Large organelle that is surrounded by two phospholipid bilayer membranes, contains DNA, and carries out **oxidative phosphorylation**, thereby producing most of the ATP in eukaryotic cells. (Figures 1-21 and 12-6)

mitogen
Any extracellular molecule, such as a growth factor, that promotes cell proliferation.

mitophagy
The process by which mitochondria are surrounded by a membrane and then delivered to lysosomes for degradation.

mitosis
In eukaryotic cells, the process whereby the nucleus divides, producing two genetically equivalent daughter nuclei with the diploid number of chromosomes. See also **cytokinesis** and **meiosis**. (Figure 18-36)

mitotic cyclin-CDKs
Cyclin-CDK complexes that promote entry into and progression through mitosis.

mitotic spindle
A specialized temporary structure, present in eukaryotic cells during mitosis, that captures the chromosomes and then pushes and pulls them to opposite sides of the dividing cell; also called *mitotic apparatus*. (Figures 18-38 and 18-39)

molecular complementarity
Lock-and-key kind of fit between the shapes, charges, hydrophobicity, and/or other physical properties of two molecules or portions thereof that allow formation of multiple **noncovalent interactions** between them at close range. (Figure 2-12)

monoclonal antibody
Antibody produced by the progeny of a single B cell and thus a homogeneous protein that recognizes a single antigen (epitope). It can be produced experimentally by use of a **hybridoma**. (Figure 4-6)

monomer
Any small molecule that can be linked chemically with others of the same type to form a **polymer**. Examples include amino acids, nucleotides, and monosaccharides.

monomeric G protein
A monomeric GTPase with a structure similar to that of the Ras protein that changes conformation when a bound GTP is hydrolyzed to GDP and phosphate. (Figure 15-7)

monopolin complex
Protein complex that promotes the co-orientation of **sister chromatids** during meiosis I in budding yeast.

monosaccharide
Any simple sugar with the formula $(CH_2O)_n$ where $n = 3–7$.

monotelic attachment
Indicates that only one of the sister kinetochore pair attached to microtubules.

morphogen
A signaling molecule whose concentration determines the fate of target-differentiating cells.

motor neurons
Nerve cells that form synapses on muscle cells.

motor protein
Any member of a special class of mechanochemical enzymes that use energy from ATP hydrolysis to generate either linear or rotary motion; also called *molecular motor*. See also **dyneins**, **kinesins**, and **myosins**.

mRNA (messenger RNA)
Any RNA that specifies the order of amino acids in a protein (i.e., the primary structure). It is produced by **transcription** of DNA by RNA polymerase. In eukaryotes, the initial RNA product (primary transcript) undergoes processing to yield functional mRNA. See also **translation**. (Figure 5-27)

mRNA surveillance
The processes that lead to the degradation of a pre-mRNA or mRNA that has been improperly processed.

mRNP exporter
A heterodimeric protein that binds to mRNA-containing ribonucleoprotein particles (mRNPs) and directs their export from the nucleus to cytoplasm by interacting transiently with **nucleoporins** in the nuclear pore complex. (Figure 9-27)

MTOC (microtubule-organizing center)

General term for any structure (e.g., centrosome, spindle pole, basal body) that organizes microtubules in cells. (Figure 18-5)

mTOR protein kinase
The protein kinase that is the signaling component of the mTORC1 complex.

mTORC1
A large complex of proteins formed on the cytosolic surface of a lysosome that integrates signals from the level of amino acids, ATP, and extracellular hormones and that regulates many aspects of cell growth and division via the mTOR protein kinase.

multi-adhesive matrix proteins
Group of long flexible proteins that bind to other components of the **extracellular matrix** and to cell-surface receptors, thereby cross-linking matrix components to the cell membrane. Examples include **laminin**, a major component of the basal lamina, and **fibronectin**, present in many tissues.

multi-hit model
The fundamental concept that most cancers progress slowly by cells sequentially acquiring multiple mutations activating different oncogenic processes.

multimeric
For proteins, containing several polypeptide chains (or subunits).

multipass membrane protein
A membrane protein with more than one transmembrane-spanning domain.

multipotent
Describes a stem cell that is able to give rise to several, but not all, types of differentiated cells in an organism.

multiprotein complexes
An assembly of multiple, distinct proteins that performs a distinct function, such as mediating one or more steps in the electron-transport pathway in mitochondria.

multiubiquitinylation
The covalent addition of several single ubiquitin molecules, each at a distinct site, on a single target protein.

multivesicular endosomes
Endosomes that segregate membrane proteins destined for degradation in the lysosome on vesicles that bud into the interior of the endosome.

MuSK
A receptor tyrosine kinase localized in the myotube plasma membrane that both induces clustering of acetylcholine receptors and serves to attract the termini of growing motor neuron axons.

mutagen
A chemical or physical agent that induces mutations.

mutation
In genetics, a permanent, heritable change in the nucleotide sequence of a chromosome, usually in a single gene; commonly causes an alteration in the function of the gene product.

myelin sheath
Stacked specialized cell membrane that forms an insulating layer around vertebrate **axons** and increases the speed of impulse conduction. (Figure 23-19)

myofibril
Long, slender structures within cytoplasm of muscle cells consisting of a regular repeating array of **sarcomeres** composed of thick (**myosin**) filaments and thin (**actin**) filaments. (Figure 17-29)

myosin light-chain kinase
(Ch. 17) A protein kinase that can phosphorylate a specific light chain of myosin-II to activate it.

myosins
A class of **motor proteins** that have actin-stimulated ATPase activity. Myosins move along actin **microfilaments** during muscle contraction and cytokinesis and also mediate vesicle translocation. (Figure 17-22)

NAD^+ (nicotinamide adenine dinucleotide)
A small organic molecule that functions as an electron carrier by accepting two electrons from a donor molecule and one H^+ from the solution. (Figure 2-33a)

$NADP^+$ (nicotinamide adenine dinucleotide phosphate)
Phosphorylated form of NAD1 that is used extensively as an electron carrier in biosynthetic pathways and during photosynthesis.

Na^+/K^+ ATPase
A P-class ATP-powered pump that couples hydrolysis of one ATP molecule to export of Na^+ ions and import of K^+ ions; largely responsible for maintaining the normal intracellular

concentrations of Na^+ (low) and K^+ (high) in animal cells; commonly called Na^+/K^+ *pump*. (Figure 11-12)

natural killer (NK) cells
Components of the innate immune system that nonspecifically detect and kill virus-infected cells and tumor cells. (Figure 24-6)

necroptosis
A type of necrosis triggered by extracellular hormones that causes inflammation.

necrosis
Cell death resulting from tissue damage or other pathology; usually marked by swelling and bursting of cells with release of their contents. Contrast with **apoptosis**.

neoblasts
Stem cells in planaria that can proliferate and regenerate many or all body cells.

neurofilaments (NFs)
A group of **intermediate filament** proteins, found only in neurons, that contribute to axonal structure and rate of transmission of action potentials down axons. (Figure 18-2b)

nervous system
Consists of the central (brain and spinal cord) and peripheral (nerves outside of the brain and spinal cord) nervous systems, and regulates all aspects of bodily function.

neuromuscular junction (NMJ)
The chemical synapse formed between a presynaptic motor neuron and a postsynaptic muscle cell.

neuron (nerve cell)
Any of the impulse-conducting cells of the nervous system. A typical neuron contains a cell body; multiple short, branched processes (**dendrites**); and one long process (**axon**). (Figure 23-1)

neurotransmitter
Extracellular signaling molecule that is released by the presynaptic neuron at a chemical **synapse** and relays the signal to the postsynaptic cell. The response elicited by a neurotransmitter, either excitatory or inhibitory, is determined by its receptor on the postsynaptic cell. (Figures 23-26 and 23-27)

neurotrophins

Family of structurally and functionally related **trophic factors** that bind to receptors called *Trks* and are required for survival of neurons; include nerve growth factor (NGF) and brain-derived neurotrophic factor (BDNF).

neutrophils
Phagocytic leukocytes that are attracted to sites of tissue damage and migrate into the tissue. Once activated, neutrophils secrete various chemokines, cytokines, bacteria-destroying enzymes (e.g., lysozyme), and other products that contribute to **inflammation** and help clear invading pathogens.

NF-κB
A transcription factor that is sequestered in the cytosol until its inhibitory protein is degraded.

nicotinamide adenine dinucleotide
See $\mathbf{NAD^+}$.

nicotinamide adenine dinucleotide phosphate
See $\mathbf{NADP^+}$.

***N*-linked oligosaccharide**
A branched oligosaccharide chain attached to the side-chain amino group of an asparagine residue in a glycoprotein. See also ***O*-linked oligosaccharide**.

nociceptor
Mechanosensor that responds to pain associated with injury to body tissues caused by mechanical trauma, heat, electricity, or toxic chemicals.

node of Ranvier (node)
An unmyelinated area of axonal membrane about 1 μm in length, which separates each region of myelin formed by an individual glial cell from the next region.

noncovalent interaction
Any relatively weak chemical interaction that does not involve an intimate sharing of electrons. (Figures 2-6 and 2-12)

nondisjunction
Mis-segregation of chromosomes during mitosis, resulting in aneuploidy or trisomy.

nonhomologous end joining
A pathway that repairs double-strand breaks in DNA in which the break ends are directly ligated without the need for a homologous template.

nonpolar
Referring to a molecule or structure that lacks any net electric charge or asymmetric distribution of positive and negative charges. Nonpolar molecules generally are less soluble in water than polar molecules and are often water insoluble.

nuclear bodies
Roughly spherical, functionally specialized regions in the nucleus, containing specific proteins and RNAs; many function in the assembly of ribonucleoprotein (RNP) complexes. The most prominent type is the nucleolus.

nuclear envelope
Double-membrane structure surrounding the nucleus; the outer membrane is continuous with the endoplasmic reticulum and the two membranes are perforated by **nuclear pore complexes**. (Figure 1-13a)

nuclear lamina
Fibrous network on the inner surface of the nuclear envelope composed of lamin intermediate filaments. (Figure 19-22b)

nuclear pore complex (NPC)
Large, multiprotein structure, composed largely of nucleoporins, that extends across the nuclear envelope. Ions and small molecules freely diffuse through NPCs; large proteins and ribonucleoprotein particles are selectively transported through NPCs with the aid of soluble proteins. (Figure 13-32a, b)

nuclear receptor
Member of a class of intracellular receptors that bind lipid-soluble molecules (e.g., steroid hormones), forming ligand-receptor complexes that activate transcription; also called *steroid receptor superfamily*. (Figure 8-45d)

nuclear speckles
A type of structure within the cell nucleus enriched in pre-mRNA splicing factors.

nuclear transport receptor
A general class of proteins that facilitate protein transport into or out of the nucleus by binding to a nuclear import sequence or a nuclear export sequence.

nucleation promoting factor (NPF)
A regulatory protein involved in activating the Arp2/3 complex in the nucleation of actin filaments.

nucleic acid

A polymer of **nucleotides** linked by **phosphodiester bonds.** DNA and RNA are the primary nucleic acids in cells.

nucleocapsid
A viral **capsid** plus the enclosed nucleic acid.

nucleoporin
Large group of proteins that make up the **nuclear pore complex**. One class (**FG-nucleoporins**) participates in nuclear import and export.

nucleoside
A small molecule composed of a **purine** or **pyrimidine** base linked to a pentose (either ribose or deoxyribose). (Table 2-3)

nucleosome
Structural unit of **chromatin** consisting of a disk-shaped core of **histone** proteins around which a 147-bp segment of DNA is wrapped. (Figure 7-20)

nucleotide
A **nucleoside** with one or more phosphate groups linked via an ester bond to the sugar moiety, generally to the 5′ carbon atom. DNA and RNA are polymers of nucleotides containing deoxyribose and ribose, respectively. (Figure 2-16; Table 2-3)

nucleus
Large membrane-bounded organelle in eukaryotic cells that contains DNA organized into chromosomes; synthesis and processing of RNA and ribosome assembly occur in the nucleus.

odorants
Small molecules that bind olfactory receptors to activate the sense of smell.

Okazaki fragments
Short (<1000 bases), single-stranded DNA fragments that are formed during synthesis of the **lagging strand** in DNA replication and are rapidly joined by DNA ligase to form a continuous DNA strand. (Figure 5-11)

olfactory receptors
A class of G protein–coupled receptors that bind odorants to activate the sense of smell.

oligodendrocytes
Glial cells in the central nervous system that create myelin sheaths that wrap around and insulate central neuronal axons.

oligopeptide
A small to medium-size linear polymer composed of amino acids connected by peptide bonds. The terms *peptide* and *oligopeptide* are often used interchangeably.

***O*-linked oligosaccharide**
Oligosaccharide chain that is attached to the side-chain hydroxyl group in a serine or threonine residue in a glycoprotein. See also ***N*-linked oligosaccharides**.

oncogene
A gene whose product is involved either in transforming cells in culture or in inducing cancer in animals. Generally is a mutant form of a normal gene (**proto-oncogene**) for a protein involved in the control of cell growth or division. (Figure 25-20)

oncogenesis
The genetic process by which cancer forms as a result of many genetic changes to a normal somatic cell.

open reading frame (ORF)
Region of sequenced DNA that is not interrupted by stop codons in one of the triplet reading frames. An ORF that begins with a start codon (usually AUG) and extends for 100 or more codons is called a *long open reading frame* and has a high probability of encoding a protein.

opsonization
The decoration of bacterial or mammalian cells with immunoglobulins and/or fragments of complement proteins to facilitate activation of phagocytic and antigen-presenting cells.

optogenetics
A technique in which channelrhodopsins are expressed in electrically excitable cells, allowing their membrane potential to be manipulated using light.

organelle
Any membrane-limited subcellular structure found in eukaryotic cells. (Figures 1-13, 1-18, 1-20, and 1-21)

osmosis
Net movement of water across a semipermeable membrane (permeable to water but not to solute) from a solution of lesser to one of greater solute concentration. (Figure 11-7)

outer mitochondrial membrane
The smooth outside boundary of the mitochondrion.

outside-in signaling
Transfer of information from inside the cell to outside the cell, usually mediated via transmembrane receptors and adapter proteins.

oxidation
Loss of electrons from an atom or molecule as occurs when a hydrogen atom is removed from a molecule or oxygen is added; opposite of **reduction**.

oxidation potential
The voltage change when an atom or molecule loses an electron; a measure of the tendency of a molecule to lose an electron. For a given oxidation reaction, the oxidation potential has the same magnitude but opposite sign as the **reduction potential** for the reverse (reduction) reaction.

oxidative phosphorylation
The phosphorylation of ADP to form ATP driven by the transfer of electrons to oxygen (O_2) in bacteria and mitochondria. Involves generation of a **proton-motive force** during electron transport and its subsequent use to power ATP synthesis.

p53
The product of a **tumor-suppressor gene** that plays a critical role in the arrest of cells with damaged DNA. Inactivating mutations in the *p53* gene are found in many human cancers. (Figure 25-15)

paracellular pathway
The extracellular pathway between cells by which molecules and ions move from one side of an epithelium to the other without crossing the plasma membranes of the cells.

paracrine
Referring to signaling mechanism in which a target cell responds to a signaling molecule (e.g., growth factor, neurotransmitter) that is produced by a nearby cell(s) and reaches the target by diffusion.

paratope
The portion of an immunoglobulin molecule that contacts the antigen for which it is specific.

patch clamping
Technique for determining ion flow through a single ion channel or across the membrane of an entire cell by use of a micropipette whose tip is applied to a small patch of the cell membrane. (Figure 11-22)

patterning genes
Genes involved in metazoan development that determine the general organization of the animal, including the major body axes and segmentation.

P body
Dense cytoplasmic domain, containing no ribosomes or translation factors, that functions in repression of translation and degradation of associated mRNAs; also called *cytoplasmic RNA-processing body*.

P-class pump
A pump that possess two identical catalytic α subunits, each of which contains an ATP-binding site. Most also have two smaller α subunits, which usually have regulatory functions. During transport, at least one of the α subunits becomes phosphorylated (hence the name "P" class), and the transported ions move through the phosphorylated subunit.

PDZ domain
A protein domain ~80–90 amino acids long that usually binds to a specific, short, C-terminal sequence motif in a "target" protein.

pentose
A five-carbon **monosaccharide**. The pentoses ribose and deoxyribose are present in RNA and DNA, respectively. (Figure 2-16)

peptide
A small linear polymer composed of amino acids connected by peptide bonds. The terms *peptide* and *oligopeptide* are often used interchangeably. See also **polypeptide**.

peptide bond
The covalent amide linkage between amino acids formed between the amino group of one amino acid and the carboxyl group of another with the net release of a water molecule (dehydration). (Figure 2-13)

peptidoglycan
A polysaccharide chain cross-linked by peptide cross-bridges in a bacterial cell wall that confers rigidity and helps to determine the cell's shape.

pericentriolar material
Amorphous material seen by thin-section electron microscopy surrounding the centrioles of animal cells. Pericentriolar material contains many components, including the gamma-tubulin ring complex (gamma-TuRC), which promotes nucleation of microtubule assembly. (Figure 18-6)

peripheral membrane protein
Any protein that associates with the cytosolic or exoplasmic face of a membrane but does not enter the hydrophobic core of the phospholipid bilayer. See also **integral membrane protein**. (Figure 10-1)

perlecan
A large multidomain **proteoglycan** component of the **extracellular matrix** (ECM) that binds to many ECM components, cell-surface molecules, and growth factors; a major component of the **basal lamina**.

peroxisome
Small organelle that contains enzymes for degrading fatty acids and amino acids by reactions that generate hydrogen peroxide, which is converted to water and oxygen by catalase.

pervasive transcription
The phenomenon where the vast majority of the genome is transcribed, well beyond the boundaries of known genes.

pH
A measure of the acidity or alkalinity of a solution defined as the negative logarithm of the hydrogen ion concentration in moles per liter: $pH = -\log [H^+]$. Neutrality is equivalent to a pH of 7; values below this are acidic and those above are alkaline.

phagocyte
Any cell that can ingest and destroy pathogens and other particulate antigens. The primary phagocytes are neutrophils, macrophages, and dendritic cells.

phagocytosis
Process by which relatively large particles (e.g., bacterial cells) are internalized by certain eukaryotic cells in a process that involves extensive remodeling of the actin cytoskeleton; distinct from **receptor-mediated endocytosis**. (Figure 17-19)

phenotype
The detectable physical and physiological characteristics of a cell or organism determined by its **genotype**; also, the specific trait associated with a particular **allele**.

Philadelphia chromosome
A translocation between chromosome 9 and chromosome 22 causing constitutive activation of ABL kinase and giving rise to chronic myeloid leukemia.

phosphatase

An enzyme that removes a phosphate group from a substrate by hydrolysis. Phosphoprotein phosphatases act with protein kinases to control the activity of many cellular proteins. (Figure 3-36)

phosphoanhydride bond
A type of high-energy bond formed between two phosphate groups, such as the γ and β phosphates and the β and α phosphates in ATP. (Figure 2-31)

phosphodiester bond
Chemical linkage between adjacent nucleotides in DNA and RNA; consists of two phosphoester bonds, one on the 5′ side of the phosphate and another on the 3′ side. (Figure 5-2)

phosphoglycerides
Amphipathic derivatives of glycerol 3-phosphate that generally consist of two hydrophobic fatty acyl chains esterified to the hydroxyl groups in glycerol and a polar head group attached to the phosphate; the most abundant lipids in biomembranes. (Figures 2-20 and 9-8a)

phosphoinositides
A group of membrane-bound lipids containing phosphorylated inositol derivatives; some function as **second messengers** in several signal transduction pathways. (Figures 15-27 and 16-16)

phospholipase
One of several enzymes that cleave various bonds in the hydrophilic end of **phospholipids**. (Figure 10-12)

phospholipid
The major class of lipids present in biomembranes, including **phosphoglycerides** and **sphingolipids**. (Figures 2-20 and 10-8a, b)

phospholipid bilayer
A two-layer, sheet-like structure in which the polar head groups of phospholipids are exposed to the aqueous media on either side and the nonpolar fatty acyl chains are in the center; the foundation for all biomembranes. (Figure 10-3a, b)

phosphorylation
The covalent addition of a phosphate group to a molecule such as a sugar or a protein. The hydrolysis of ATP often accompanies phosphorylation, providing energy to drive the

reaction and the phosphate group that is covalently added to the target molecule. Enzymes that catalyze phosphorylation are called *kinases*.

photoelectron transport
Light-driven electron transport that generates a charge separation across the **thylakoid membrane** that drives subsequent events in photosynthesis. (Figure 12-31)

photorespiration
A reaction pathway that competes with CO_2 fixation (**Calvin cycle**) by consuming ATP and generating CO_2, thus reducing the efficiency of photosynthesis. (Figure 12-42)

photosynthesis
Complex series of reactions occurring in some bacteria and in plant **chloroplasts** in which light energy is used to generate carbohydrates from CO_2, usually with the consumption of H_2O and evolution of O_2.

photosystems
Multiprotein complexes, present in all photosynthetic organisms, that consist of light-harvesting complexes containing **chlorophylls** and a reaction center where **photoelectron transport** occurs. (Figure 12-37)

photosystem I (PSI)
One of two multiprotein complex photosystems that converts light energy to high-energy electrons, which in turn drive the formation of NADPH. Primarily in unstacked regions of the thylakoid membranes.

photosystem II (PSII)
One of two multiprotein complex photosystems that converts light energy to high-energy electrons, which in turn drive the formation of NADPH. Primarily in stacked regions of the thylakoid membranes.

phragmoplast
In plants, a temporary structure, formed during telophase, whose membranes become the plasma membranes of the daughter cells and whose contents develop into the new cell wall between them. (Figure 18-48)

plakins
A family of proteins that help attach **intermediate filaments** to other structures.

planar cell polarity (PCP)
A signaling pathway that organizes the orientation of individual cells in a cell layer.

plaque assay
Technique for determining the number of infectious viral particles in a sample by culturing a diluted sample on a layer of susceptible host cells and then counting the clear areas of lysed cells (plaques) that develop.

plasma membrane
The membrane surrounding a cell that separates the cell from its external environment; consists of a **phospholipid bilayer** and associated membrane lipids and proteins. (Figures 1-13a, 10-1, and 10-5)

plasmid
Small, circular extrachromosomal DNA molecule capable of autonomous replication in a cell; commonly used as a **vector** in **DNA cloning**.

plasmodesmata (sing. plasmodesma)
Tubelike cell junctions that interconnect the cytoplasms of adjacent plant cells and are functionally analogous to **gap junctions** in animal cells. (Figure 20-43)

pluripotent
Describes a type of stem cell that can differentiate into a wide range of cell types from the three primary germ layers, but not necessarily cells that form the placenta.

point mutation
Change of a single nucleotide in DNA, especially in a region coding for protein; can result in formation of a codon specifying a different amino acid or a stop codon. Addition or deletion of a single nucleotide will cause a shift in the **reading frame**.

polar
Referring to a molecule or structure with a net electric charge or asymmetric distribution of positive and negative charges. Polar molecules are usually soluble in water.

polarity
In cell biology, the presence of functional and/or structural differences in distinct regions of a cell or cellular component. See also **cell polarity**.

poly(A) tail
A long chain of adenine nucleotides that is added to an mRNA molecule during RNA processing to increase the stability of the molecule.

polyclonal antibody
Antibodies raised in an experimental animal by injecting a specific protein.

polymer
Any large molecule composed of multiple identical or similar units (**monomers**) linked by covalent bonds. (Figure 2-13)

polymerase chain reaction (PCR)
Cycles of synthesis of a specific segment of DNA carried out in vitro. The reaction is carried out by sequential cycles of heating to dissociate DNA strands, annealing to specific oligonucleotide primers, and DNA synthesis with a heat-stable DNA polymerase. Each reaction cycle leads to a doubling in the amount of the target DNA.

polypeptide
Linear polymer of amino acids connected by peptide bonds, usually containing 20 or more residues. See also **protein**.

polyribosome
A complex containing several **ribosomes**, all translating a single messenger RNA; also called *polysome*. (Figure 5-39)

polysaccharide
Linear or branched polymer of **monosaccharides** linked by glycosidic bonds and usually containing more than 15 residues. Those with fewer than 15 residues are often called *oligosaccharides*.

polytene chromosome
Enlarged chromosome composed of many parallel copies of itself formed by multiple cycles of DNA replication without chromosomal separation; found in the salivary glands and some other tissues of *Drosophila* and other dipteran insects. (Figure 7-36)

polyubiquitination
The covalent addition of a chain of covalently linked ubiquitin molecules to a site on a target protein.

polyunsaturated
Referring to a compound (e.g., fatty acid) in which two or more of the carbon-carbon bonds are double or triple bonds.

porins
Class of trimeric transmembrane proteins through which small water-soluble molecules can cross the membrane; present in outer mitochondrial and chloroplast membranes and in the outer membrane of gram-negative bacteria. (Figure 10-18)

post-translational modification (PTM)

The covalent and generally enzymatic modification of proteins following protein biosynthesis.

post-translational translocation
Some secretory proteins enter the ER lumen after translation has been completed.

potential energy
Stored energy. In biological systems, the primary forms of potential energy are chemical bonds, concentration gradients, and electric potentials across cellular membranes.

precursor mRNA (pre-mRNA)
The primary RNA product formed by transcription of a gene. Splicing of introns and modifications to the 5′ and 3′ ends of the pre-mRNA are necessary to form a mature mRNA that can be transported out of the nucleus and translated on ribosomes. (Figures 5-27 and 9-2)

pre-rRNA
Large precursor ribosomal RNA that is synthesized in the nucleolus of eukaryotic cells and processed to yield three of the four RNAs present in ribosomes. (Figures 9-43 and 9-44)

primary cilium
The single nonmotile cilium present on almost all vertebrate cells that serves as a sensory organelle to detect extracellular signals.

primary structure
In proteins, the linear arrangement (sequence) of amino acids within a polypeptide chain.

primase
A specialized RNA polymerase that synthesizes short stretches of RNA used as primers for DNA synthesis. (Figure 5-12)

primer
A short nucleic acid sequence containing a free 3′-hydroxyl group that forms **base pairs** with a complementary template strand and functions as the starting point for addition of nucleotides to copy the template strand.

progenitor (precursor) cells
A type of undifferentiated cell that, when provided with the appropriate signals, will divide and differentiate into one or a few cell types.

programmed cell death
See **apoptosis**.

prokaryotic organisms
Class of organisms, including the **bacteria** (eubacteria) and archaea, that lack a true membrane-limited nucleus and other organelles. (Figure 1-1)

prolactin
A cytokine released during pregnancy of mammals that induces the development of mammary glands to produce and secrete milk proteins.

prometaphase
Second stage in mitosis, during which the nuclear envelope and nuclear lamina break down and microtubules assembled from the spindle poles "capture" chromosome pairs at specialized structures called **kinetochores**. (Figures 18-38 and 18-39)

promoter
DNA sequence that determines the site of **transcription** initiation for an RNA polymerase. (Figure 5-24)

promoter-proximal element
Any regulatory sequence in eukaryotic DNA that is located within ~200 base pairs of the transcription start site. Transcription of many genes is controlled by multiple promoter-proximal elements. (Figure 8-18)

prophase
Earliest stage in mitosis, during which the chromosomes condense, the duplicated centrosomes separate to become the spindle poles, and the mitotic spindle begins to form. (Figures 18-38 and 18-39)

protease
Any enzyme that cleaves one or more peptide bonds in target proteins.

proteasome
Large multifunctional protease complex in the cytosol that degrades intracellular proteins marked for destruction by attachment of multiple **ubiquitin** molecules. (Figure 3-32)

protein
A macromolecule composed of one or more linear **polypeptide** chains and folded into a characteristic three-dimensional shape (**conformation**) in its native, biologically active state.

protein disulfide isomerase (PDI)
An enzyme, present in all eukaryotic cells, and especially abundant in the ER of secretory cells in organs such as the liver and pancreas, where large quantities of proteins that

contain disulfide bonds are produced. (Figure 13-19)

protein domain
Distinct regions of a protein's three-dimensional structure. A *functional* domain exhibits a particular activity characteristic of the protein; a *structural* domain is ≈40 or more amino acids in length, arranged in a distinct secondary or tertiary structure; a *topological* domain has a distinctive spatial relationship to the rest of the protein.

protein family
Set of homologous proteins encoded by a **gene family**.

protein kinase (PK)
An enzyme that chemically adds a phosphate group, usually the gamma phosphate from ATP, to a serine, threonine, or tyrosine residue of another protein.

protein kinase B (PKB)
Cytosolic enzyme that is recruited to the plasma membrane by signal-induced **phosphoinositides** and subsequently activated; also called **Akt**. (Figure 16-17)

protein phosphatase
An enzyme that chemically removes a phosphate group from another protein by hydrolysis.

proteoglycans
A group of glycoproteins (e.g., perlecan and aggrecan) that contain a core protein to which is attached one or more **glycosaminoglycan (GAG)** chains. They are found in nearly all animal extracellular matrices, and some are integral membrane proteins. (Figure 20-33)

proteome
All the proteins in a cellular compartment, intact cell, organ, or organism.

proteomics
The systematic study of the amounts, modifications, interactions, localization, and functions of all or subsets of proteins at the whole-organism, tissue, cellular, and subcellular levels.

proteostasis
The process of maintaining proteins and their activities at appropriate levels, and for making rapid adjustments in these levels in response to changing conditions.

proton-motive force
The energy equivalent of the proton (H^+) concentration gradient and electric potential gradient across a membrane; used to drive ATP synthesis by **ATP synthase**, transport of

molecules against their concentration gradient, and movement of bacterial flagella. (Figure 12-2)

proton pump
A protein that uses an external source of energy (ATP, high-energy electrons) to transport protons across a membrane against an electrochemical gradient. In mitochondria and chloroplasts, a proton pump can generate the proton-motive force used to synthesize ATP.

proton pumping
The process catalyzed by a proton pump in which an external source of energy (ATP, high-energy electrons) is used to transport protons across a membrane against an electrochemical gradient.

proto-oncogene
A normal cellular gene that encodes a protein usually involved in regulation of cell growth or differentiation and that can be mutated into a cancer-promoting **oncogene**, either by changing the protein-coding segment or by altering its expression.

protostomes
A group of bilaterally symmetric animals whose mouth develops close to the blastopore and has a ventral nerve cord. This group includes worms, insects, and mollusks.

pseudogene
DNA sequence that is similar to that of a functional gene but does not encode a functional product; probably arose by sequence drift of duplicated genes.

pulse-chase
A type of experiment in which a radioactive small molecule is added to a cell for a brief period (the pulse) and then is replaced with an excess of the unlabeled form of the same small molecule (the chase). Used to detect changes in the cellular location of a molecule or its metabolic fate over time. (Figure 3-45)

purines
A class of nitrogenous compounds containing two fused heterocyclic rings. Two purines, adenine (A) and guanine (G), are base components of nucleotides found in DNA and RNA. See also **base pair**. (Figure 2-17)

pyrimidines
A class of nitrogenous compounds containing one heterocyclic ring. Two pyrimidines, cytosine (C) and thymine (T), are base components of nucleotides found in DNA; in RNA, uracil (U) replaces thymine. See also **base pair**. (Figure 2-17)

pyruvate
The ionized form of a three-carbon, carboxylic acid that is a key intermediate in energy metabolism.

Q cycle
An evolutionarily conserved mechanism employed by mitochondrial complex III that uses the net oxidation of one $CoQH_2$ molecule to transfer four protons into the mitochondrial intermembrane space and two electrons to two cytochrome *c*. The cytochrome *bf* complex in the chloroplast also uses a Q cycle.

quaternary structure
The number and relative positions of the polypeptide chains in multimeric (multisubunit) proteins. (Figure 3-10b)

Rab proteins
A family of small GTPases that regulate the tethering and targeting of transport vesicles.

Rac protein
A small GTPase that can induce the formation of actin filaments in a lamellipodium by regulating the activity of the Arp2/3 complex through the WAVE complex.

radioisotope
Unstable form of an atom that emits radiation as it decays. Several radioisotopes are commonly used experimentally as labels in biological molecules. (Table 3-1)

Ran protein
A small GTPase that is responsible for most unidirectional transport of large proteins into and out of the nucleus through the nuclear pore.

Ras protein
A monomeric member of the **GTPase superfamily** of switch proteins that is tethered to the plasma membrane by a lipid anchor and functions in intracellular signaling pathways; activated by ligand binding to receptor tyrosine kinases and some other cell-surface receptors. (Figures 16-10 and 16-12)

rate constant
A constant that relates the concentrations of reactants to the rate of a chemical reaction.

rate-limiting step
The slowest step in a multistep process, for example, in a multistep enzymatic reaction or a multistep biosynthetic pathway.

Rb protein
Inhibitor of the E2F transcription factor family and thus a key regulator of cell cycle entry.

reactive oxygen species (ROS)
Highly chemically reactive, oxygen-containing molecules, both radicals and non-radicals, that can damage many biological molecules and can serve as signaling molecules.

reading frame
The sequence of nucleotide triplets (**codons**) that runs from a specific translation start codon in an mRNA to a stop codon. Some mRNAs can be translated into different polypeptides by reading in two different reading frames. (Figure 4-17)

receptor
Any protein that specifically binds another molecule to mediate cell-cell signaling, adhesion, endocytosis, or other cellular process. Commonly denotes a protein located in the plasma membrane, cytosol, or nucleus that binds a specific extracellular molecule (**ligand**), which often induces a conformational change in the receptor, thereby initiating a cellular response. See also **adhesion receptor** and **nuclear receptor**. (Figure 15-3)

receptor protein
A protein that binds a ligand, which could be a small molecule such as adrenalin, a growth factor such as epidermal growth factor, a hormone such as insulin, or another macromolecule, usually with engagement of a signaling cascade as a consequence.

receptor-mediated endocytosis
Uptake of extracellular materials bound to specific cell-surface receptors by invagination of the plasma membrane to form a small membrane-bounded vesicle (early endosome). (Figure 14-29)

recessive
In genetics, referring to that allele of a gene that is not expressed in the **phenotype** when the **dominant** allele is present; also refers to the phenotype of an individual (homozygote) carrying two recessive alleles. Mutations that produce recessive alleles generally result in a loss of the gene's function. (Figure 6-2)

recombinant DNA
Any DNA molecule formed in vitro by joining DNA fragments from different sources.

recombination
Any process in which chromosomes or DNA molecules are cleaved and the fragments are rejoined to give new combinations. Homologous recombination occurs during meiosis,

giving rise to crossing over of homologous chromosomes. Homologous recombination and nonhomologous recombination (i.e., between chromosomes of different morphologic type) also occur during several DNA-repair mechanisms and can be carried out in vitro with purified DNA and enzymes. (Figure 6-27)

redox reaction
An oxidation-reduction reaction in which one or more electrons are transferred from one reactant to another.

reduction
Gain of electrons by an atom or molecule as occurs when a hydrogen atom is added to a molecule or oxygen is removed. The opposite of **oxidation**.

reduction potential (*E*)
The voltage change when an atom or molecule gains an electron; a measure of the tendency of a molecule to gain an electron. For a given reduction reaction, *E* has the same magnitude but opposite sign as the **oxidation potential** for the reverse (oxidation) reaction.

refractory period
The period of time during which a neuron is hyperpolarized and unable to respond to stimuli because the sodium channels are inactivated.

regulated secretion
Proteins are stored inside the cell in specialized vesicles known as *secretory granules* until a signal for exocytosis causes them to be released.

release factor (RF)
One of two types of nonribosomal proteins that recognize stop codons in mRNA and promote release of the completed polypeptide chain, thereby terminating **translation** (protein synthesis). (Figure 5-38)

replication fork
Y-shaped region in double-stranded DNA at which the two strands are separated and replicated during DNA synthesis; also called *growing fork*. (Figure 5-11)

replication origin
Unique DNA segment present in an organism's genome at which DNA replication begins. Eukaryotic chromosomes contain multiple origins, whereas bacterial chromosomes and plasmids usually contain just one.

repolarization

After the action potential passes through a section of a neuron, channel proteins and pumps restore the inside-negative resting potential.

reporter gene
A gene encoding a protein that is easily assayed (e.g., β-galactosidase, luciferase). Reporter genes are used in various types of experiments to indicate activation of the promoter to which the reporter gene is linked.

repression domain
A region of a repressor transcription factor that will inhibit transcription when fused to a DNA-binding domain.

repressor
Specific **transcription factor** that inhibits transcription.

residue
General term for the repeating units in a polymer that remain after covalent linkage of the monomeric precursors.

resolution
The minimum distance between two objects that can be distinguished by an optical apparatus; also called *resolving power*.

respiration
The conversion of energy in nutrients into ATP via a set of reactions involving a series of protein-catalyzed, membrane-associated oxidation and reduction reactions, called an *electron-transport chain*, that are ultimately coupled to the addition of P_i to ADP (oxidative phosphorylation) to form ATP and transfer of electrons to oxygen or other inorganic electron acceptors.

respiratory chain
See **electron-transport chain**.

respiratory control
Dependence of mitochondrial oxidation of NADH and $FADH_2$ on the supply of ADP and P_i for ATP synthesis.

resting K^+ channels
Nongated K^+ ion channels in the plasma membrane that, in conjunction with the high cytosolic K^+ concentration produced by the **Na^+/K^+ ATPase**, are primarily responsible for generating the inside-negative resting **membrane potential** in animal cells.

resting membrane potential
The electrical potential difference that exists across the membrane of a cell that is in a non-excited state. The electrical potential results from the differential distribution of ions across the membrane.

restriction enzyme
Any enzyme that recognizes and cleaves a specific short sequence, the *restriction site*, in double-stranded DNA molecules; used extensively to produce **recombinant DNA** in vitro; also called *restriction endonuclease*. (Figure 6-11)

restriction point
Point in the mammalian cell cycle after which cells are no longer responsive to proliferation regulatory signals.

retrograde
Backward-moving transport vesicle that retrieves proteins from the cis-Golgi cisterna and returns them to the ER.

retromer
A protein complex that coats vesicles budding from late endosomes, which recycle the M6P receptor back to the trans-Golgi network.

retrotransposon
Type of eukaryotic **transposable DNA element** whose movement in the genome is mediated by an RNA intermediate and involves a reverse-transcription step. See also **DNA transposon**. (Figure 7-8)

retrovirus
Type of eukaryotic virus containing an RNA genome that replicates in cells by first making a DNA copy of the RNA. This viral DNA is inserted into cellular chromosomal DNA, forming a provirus, and gives rise to further genomic RNA as well as the mRNAs for viral proteins. (Figure 5-43)

reverse transcriptase
Enzyme found in retroviruses that catalyzes a complex reaction in which a double-stranded DNA is synthesized from a single-stranded RNA template. (Figure 7-14)

RGD motif
The Arg-Gly-Asp sequence motif that serves as a binding site for integrins.

R group

A portion of a molecule, or a chemical group within a larger molecule, that is covalently bonded as an appendage to the main body or core of the molecule. In an amino acid, the R group is the side chain that is attached to the alpha carbon atom and that confers the distinct characteristics of the amino acid.

Rho protein
A protein that can activate myosin light chain kinase and formins in the assembly of contractile structures such as stress fibers. The protein family of which Cdc42, Rac, and Rho are members has also been collectively named *Rho proteins*.

ribonucleic acid (RNA)
See **RNA**.

ribonucleoprotein (RNP) complex
General term for any complex composed of proteins and RNA. Most RNA molecules are present in the cell in the form of RNPs.

ribosomal RNA (rRNA)
Any one of several large RNA molecules that are structural and functional components of ribosomes. Often designated by their sedimentation coefficient: 28S, 18S, 5.8S, and 5S rRNA in higher eukaryotes. (Figure 5-34)

ribosome
A large complex comprising several different **rRNA** molecules and as many as 83 proteins, organized into a large subunit and small subunit; the engine of **translation** (protein synthesis). (Figures 5-34 and 5-35)

ribozyme
An RNA molecule with catalytic activity. Ribozymes function in RNA splicing and protein synthesis.

RNA (ribonucleic acid)
Linear, single-stranded polymer, composed of ribose **nucleotides**. mRNA, rRNA, and tRNA play different roles in protein synthesis; a variety of small RNAs play roles in controlling the stability and translation of mRNAs and in controlling chromatin structure and transcription. (Figure 5-29)

RNA editing
Unusual type of RNA processing in which the sequence of a pre-mRNA is altered.

RNA-induced silencing complex (RISC)

Large multiprotein complex associated with a short single-stranded RNA (**siRNA** or **miRNA**) that mediates degradation or translational repression of a complementary or near-complementary mRNA.

RNA interference (RNAi)
Functional inactivation of a specific gene by a corresponding double-stranded RNA that induces either inhibition of translation or degradation of the complementary single-stranded mRNA encoded by the gene but not that of mRNAs with a different sequence. (Figure 6-41)

RNA polymerase
An enzyme that copies one strand of DNA (the *template* strand) to make the **complementary** RNA strand using as substrates ribonucleoside triphosphates. (Figure 5-24)

RNA polymerase II
An enzyme that synthesizes precursors of mRNAs and most snRNA and miRNAs.

RNA splicing
A process that results in removal of **introns** and joining of **exons** in pre-mRNAs. See also **spliceosome**. (Figures 9-8 and 9-9)

rough endoplasmic reticulum
The region of the endoplasmic reticulum with ribosomes attached to the cytosolic face; site of synthesis of secreted and many membrane proteins.

rRNA
See **ribosomal RNA**.

rubisco (ribulose 1,5-bisphosphate carboxylase)
Enzyme located in chloroplasts that catalyzes the first reaction in the Calvin cycle, the addition of CO_2 to a five-carbon sugar (ribulose 1,5-bisphosphate) to form two molecules of 3-phosphoglycerate. (Figure 12-40)

S (synthesis) phase
See **cell cycle**.

saltatory conduction
The propagation of action potentials from one node of Ranvier to the next along myelinated axons.

sarcoma

A tumor that originates from bone or other types of connective tissue.

sarcomere
Repeating structural unit of striated (skeletal) muscle composed of organized, overlapping thin (**actin**) filaments and thick (**myosin**) filaments and extending from one Z disk to an adjacent one; shortens during contraction. (Figures 17-29 and 17-30)

sarcoplasmic reticulum
Network of membranes in cytoplasm of a muscle cell that sequesters Ca^{2+} ions; release of stored Ca^{2+} induced by muscle stimulation triggers contraction. (Figure 17-32)

saturated
Referring to a compound (e.g., fatty acid) in which all the carbon-carbon bonds are single bonds.

scaffold proteins
Proteins that participate in intracellular signaling by binding several proteins in a signaling pathway and localizing them to a specific region of the cell.

scanning electron microscopy (SEM)
A type of electron microscopy that yields the three-dimensional structure of samples. Samples are coated with a thin film of metal, and then an intense beam of electrons scans the sample to generate secondary electrons that generate the image.

SCF
Ubiquitin-protein ligase that ubiquitinylates inhibitors of S-phase CDKs and many other proteins, marking them for degradation by proteasomes.

Schwann cells
Glial cells in the peripheral nervous system that create myelin sheaths that wrap around and insulate peripheral neuronal axons.

second messenger
A small intracellular molecule (e.g., cAMP, cGMP, Ca^{2+}, DAG, and IP_3) whose concentration increases (or decreases) in response to binding of an extracellular signal and that functions in **signal transduction**. (Figure 15-5)

secondary structure
In proteins, local folding of a polypeptide chain into regular structures including the α helix, β sheet, and β turns.

***secretion (sec)* mutants**

A collection of temperature-sensitive mutants isolated in yeast that block protein secretion at the nonpermissive temperature.

secretory pathway
Cellular pathway for synthesizing and sorting soluble and membrane proteins localized to the endoplasmic reticulum, Golgi, and lysosomes; plasma membrane proteins; and proteins eventually secreted from the cell. (Figure 14-2)

securin
A protein that prevents the onset of chromosome segregation by inhibiting separase, the protease that cleaves cohesin.

segregation
The process that distributes an equal complement of chromosomes to daughter cells during mitosis and meiosis.

selectins
A family of **cell-adhesion molecules** that mediate Ca^{2+}-dependent interactions with specific oligosaccharide moieties in glycoproteins and glycolipids on the surface of adjacent cells or in extracellular glycoproteins. (Figures 20-2 and 20-42)

sensor
Measures an intra- or extracellular property and converts into a signal.

sensory neurons
Neurons that receive external stimuli and transmit them to the brain and spinal cord.

separase
A protease that cleaves cohesin, thereby initiating chromosome segregation.

short interspersed elements (SINEs)
In mammals, repeated sequences related to a 150 (in mice)–300 (in humans) base-pair sequence that has been transposed throughout the genome over evolutionary time by proteins encoded by a long interspersed element (LINE) in the same organism. (Table 7-1)

shuttle vector
Plasmid vector capable of propagation in two different hosts. (Figure 6-15)

signal amplification
The aspect of a signal transduction pathway whereby activation by a small number of signaling molecules leads to an effect on many times that number of target molecules.

signal-anchor sequence (SA)
A hydrophobic N-terminal sequence that serves as an ER-targeting signal sequence and becomes a hydrophobic transmembrane-spanning sequence.

signal-based targeting
A process that involves the targeting of a newly synthesized protein from the cytoplasm to an intracellular organelle.

signaling cascade
Pathway that is activated by an intracellular or extracellular event and then transmits the signal to an **effector**.

signaling molecule
General term for any extracellular or intracellular molecule involved in mediating the response of a cell to its external environment or to other cells.

signal peptide
See **targeting sequence**.

signal recognition particle (SRP)
A cytosolic ribonucleoprotein particle that binds to the ER **signal sequence** in a nascent secretory protein and delivers the nascent chain/ribosome complex to the ER membrane, where synthesis of the protein and translocation into the ER are completed. (Figure 13-5)

signal sequence
See **targeting sequence**.

signal transduction pathway
The set of proteins or small molecules involved in relaying a chemical message from the receptor at the cell surface to the target molecule(s) inside the cell.

silencer sequences
DNA sequences capable of binding transcription regulation factors (repressors). (Figure 8-31)

simple diffusion
Net movement of a molecule across a membrane down its concentration gradient at a rate proportional to the gradient and the permeability of the membrane; also called *passive diffusion*.

simple-sequence (satellite) DNA

Short, tandemly repeated sequences that are found at **centromeres** and **telomeres** as well as at other chromosomal locations and are not transcribed.

single cell RNA-Seq (scRNA-Seq)
A method that allows analysis of the abundance of specific mRNAs from a single cell. Typically, mRNA extracted from a single cell is converted to cDNA with the enzyme reverse transcriptase, and particular gene sequences are amplified by PCR and then sequenced. The relative number of reads from a particular mRNA is used to estimate abundance.

single-pass membrane protein
A membrane protein that contains a single transmembrane-spanning domain.

siRNA (short interfering RNA)
A small, double-stranded RNA, 21–23 nucleotides long with two single-stranded nucleotides at each end. A single strand of the siRNA associates with several proteins to form an **RNA-induced silencing complex (RISC)** that cleaves target RNAs to which the siRNA base pairs perfectly; variously called *short* or *small interfering RNA* and *small inhibitory RNA*. siRNAs can be designed to experimentally inhibit expression of specific genes. See also **miRNA**. (Figure 9-31b)

siRNA knockdown
Technique for experimentally inhibiting translation of a specific mRNA by use of **siRNA**; useful for reducing the activity of a protein, particularly in organisms that are not amenable to classical genetic methods for isolating loss-of-function mutants.

sister chromatid resolution
The process of untangling of the intertwined **sister chromatids** during prophase.

sister chromatids
The two identical DNA molecules created during DNA replication and the associated chromosomal proteins. After DNA replication, each chromosome is composed of two sister chromatids.

SM proteins
Sec1/Munc18-like proteins, which bind syntaxin and are required for SNARE-mediated fusion of synaptic vesicles to a cell membrane.

Smads
Class of transcription factors that are activated by phosphorylation following binding of members of the **transforming growth factor beta (TGFβ)** family of signaling molecules to their cell-surface receptors. (Figure 16-24)

smooth endoplasmic reticulum
The region of the endoplasmic reticulum bereft of ribosomes; site of synthesis of many types of lipids and phospholipids.

SNAREs
Cytosolic and integral membrane proteins that promote fusion of vesicles with target membranes. Interaction of **v-SNAREs** on a vesicle with cognate **t-SNAREs** on a target membrane forms very stable complexes, bringing the vesicle and target membranes into close apposition. (Figure 14-11)

snoRNA (small nucleolar RNA)
A type of small, stable RNA that functions in rRNA processing and base modification in the nucleolus.

snRNA (small nuclear RNA)
One of several small, stable RNAs localized to the nucleus. Five snRNAs are components of the **spliceosome** and function in splicing of pre-mRNA. (Figures 9-9 and 9-10)

snRNPs (small ribonuclear particles) U1, U2, U4, U5, and U6
Five smaller ribonucleoprotein complexes that are assembled on the snRNAs with the same names. The U1, U2, U4, U5, and U6 snRNPs associate transiently with each other and with splice sites in pre-mRNAs to assemble the **spliceosome**.

somatic cell
Any plant or animal cell other than a germ cell.

sorting signal
A relatively short amino acid sequence within a protein that directs the protein to particular transport vesicles as they bud from a donor membrane in the secretory or endocytic pathway. (Table 14-2)

Southern blotting
Technique for detecting specific DNA sequences separated by electrophoresis by hybridization to a labeled nucleic acid probe.

special-pair chlorophylls
Two chlorophyll molecules in a photosynthetic reaction center that can transfer energy to an electron and thus initiate photoelectron transport. These chlorophylls obtain the energy either by directly absorbing light energy or by resonance energy transfer from antenna pigments.

specificity

The ability of immune cells or their products to distinguish between structurally closely related molecules.

sperm
The male gamete — a motile haploid cell that can bind to and fuse with an egg cell, forming a zygote.

S-phase cyclin-CDKs
Cyclin-CDK complexes that promote the initiation of DNA replication.

sphingolipid
Major group of membrane lipids, derived from sphingosine, that contain two long hydrocarbon chains and either a phosphorylated head group (sphingomyelin) or carbohydrate head group (cerebrosides, gangliosides). (Figure 10-8b)

spindle pole bodies
Functionally analogous structure of centrosomes in yeast.

splicing enhancers
The RNA sequences bound by RNA-binding proteins that stimulate RNA splicing.

splicing silencers
The RNA sequences bound by RNA-binding proteins that inhibit RNA splicing.

spliceosome
Large ribonucleoprotein complex that assembles on a pre-mRNA and carries out **RNA splicing**.

SR proteins
A subset of hnRNP proteins that bind exonic splicing enhancers.

SRE-binding proteins (SREBPs)
Cholesterol-dependent transcription factors, localized in the ER membrane, that are activated in response to low cellular cholesterol levels and then stimulate expression of genes encoding proteins involved in cholesterol synthesis and import as well as synthesis of other lipids.

SRP receptor
The receptor for the **signal recognition particle (SRP)**, located in the membrane.

starch

A very long, branched polysaccharide, composed exclusively of glucose units, that is the primary storage carbohydrate in plant cells. (Figure 12-32)

steady state
In cellular metabolic pathways, the condition when the rate of formation and rate of consumption of a substance are equal, so that its concentration remains constant. (Figure 2-23)

stem cell
A self-renewing cell that can divide *symmetrically* to give rise to two daughter cells whose developmental potential is identical to that of the parental stem cell or *asymmetrically* to generate daughter cells with different developmental potentials. (Figures 22-11 and 22-12)

stem-cell niche
A set of cells, extracellular matrices, and hormones that surrounds a stem cell and that maintains its stem-cell properties.

step size
Used when describing the movement of a molecular motor such as myosin, kinesin, or dynein; the distance the motor moves during one cycle of ATP hydrolysis.

stereoisomers
Two compounds with identical molecular formulas whose atoms are linked in the same order but in different spatial arrangements. In *optical isomers*, designated D and L, the atoms bonded to an asymmetric carbon atom are arranged in a mirror-image fashion. *Geometric isomers* include the *cis* and *trans* forms of molecules containing a double bond.

stop-transfer anchor sequence (STA)
A sequence that functions both to stop passage of the polypeptide chain through the translocon and to become a hydrophobic transmembrane segment in the membrane bilayer.

strand invasion
An early step in DNA recombination catalyzed by RecA in bacteria and Rad51 in eukaryotes in which a single-stranded region of DNA with a free 3′ end hybridizes to its complementary strand in a second double-stranded DNA molecule. The complementary strand of this target double-stranded DNA is displaced as a single-stranded loop of DNA over the region of hybridization to the invading strand. (Figures 5-20 and 5-21)

stress fibers

Contractile F-actin- and myosin II-containing structures associated with focal adhesions whose formation is under control of the small GTPase Rho.

stroma
The fluid-filled compartment in a chloroplast between the thylakoid membrane and the inner membrane in which carbon fixation and starch synthesis in stage 4 of photosynthesis take place.

structural motif
A particular combination of two or more secondary structures that form a distinct three-dimensional structure that appears in multiple proteins and that often, but not always, is associated with a specific function.

substrate
Molecule that undergoes a charge in a reaction catalyzed by an enzyme.

substrate-level phosphorylation
Formation of ATP from ADP and P_i catalyzed by cytosolic enzymes in reactions that do not depend on a proton-motive force or molecular oxygen.

suprachiasmatic nucleus (SCN)
A collection of neurons in the region of the brain called the *hypothalamus*. It functions as a master clock or central pacemaker for circadian rhythms in mammals.

supercomplexes
Very large assemblies of protein complexes that participate in pathways. The assembly of the supercomplexes improves the speed and efficiency of these pathways.

superoxide anions
One type of reactive oxygen species in which an unpaired electron is added to molecular oxygen (O_2) to form the anion $O_2^{\bullet -}$.

SWI/SNF chromatin-remodeling complex
A multiprotein chromatin-remodeling co-activator complex with at least one subunit that is homologous to DNA helicases that facilitates nucleosomes sliding along DNA and decondensation of condensed chromatin structures.

symmetric cell division
Separation of a cell into two daughter cells that look and function exactly like the parent cell. In higher eukaryotes, it involves division of the nucleus (mitosis) and of the cytoplasm (cytokinesis); mitosis often is used to refer to both nuclear and cytoplasmic division. (Figure 22-1a).

symport
A type of **cotransport** in which a membrane protein (*symporter*) transports two different molecules or ions across a cell membrane in the *same* direction. See also **antiport**. (Figure 11-3, **[3B]**)

symporter
An integral membrane protein that functions as a cotransporter to transport two molecules in the same direction across a membrane.

synapse
Specialized region between an axon terminal of a neuron and an adjacent neuron or other excitable cell (e.g., muscle cell) across which impulses are transmitted. At a *chemical* synapse, the impulse is conducted by a **neurotransmitter**; at an *electric* synapse, impulse transmission occurs via **gap junctions** connecting the pre- and postsynaptic cells. (Figure 23-3)

synapse elimination
A process during nervous system development during which synaptic connections are removed.

synaptic plasticity
The modification of synaptic connectivity based on neuronal experience, which is a biological manifestation of forming and storing memories.

synaptic vesicles
Small vesicles in axon termini that contain a neurotransmitter and that undergo exocytosis following arrival of an action potential.

synaptonemal complex (SC)
Proteinaceous structure that mediates the association (synapsis) between homologous chromosomes during prophase of meiosis I.

syntelic attachment
Indicates that kinetochores of a **sister chromatid** pair attach to microtubules emanating from the same pole.

synteny
Occurrence of genes in the same order on a chromosome in two or more different species.

talin

An intracellular, mechanosensory protein that functions as a regulated adapter which binds to the cytoplasmic tail of the ß subunit of integrins.

target cell
A cell that senses and responds to a signal, often a protein or small molecule, sent by another cell.

targeting sequence
A relatively short amino acid sequence within a protein that directs the protein to a specific location within the cell; also called *signal peptide, signal sequence,* and *uptake-targeting sequence*. (Table 13-1)

tastants
Water-soluble molecules that activate taste receptors.

taste receptors
A class of receptors that bind to tastants to activate the sense of taste.

Tat
An HIV-encoded protein that prevents Pol II termination when synthesizing the polycistronic HIV mRNA, thereby allowing transcription of the complete proviral DNA genome.

TATA box
A conserved sequence in the **promoter** of many eukaryotic protein-coding genes where the transcription-initiation complex assembles. (Figure 8-11)

TATA box–binding protein (TBP)
A subunit of TFIID (a general transcription factor), TBP is the first protein to bind to a TATA box promoter. All eukaryotic TBPs analyzed to date have very similar C-terminal domains of 180 residues. This domain of TBP folds into a saddle-shaped structure; the two halves of the molecule exhibit an overall dyad symmetry but are not identical. TBP interacts with the minor groove in DNA, bending the helix.

T cell
A lymphocyte that matures in the thymus and expresses antigen-specific receptors that bind antigenic peptides complexed to MHC molecules. There are two major classes: *cytotoxic* T cells (CD8 surface marker, class I MHC restricted, kill virus-infected and tumor cells) and *helper* T cells (CD4 marker, class II MHC restricted, produce cytokines, required for activation of B cells). (Figures 24-36 and 24-40)

T-cell receptor (TCR)
A heterodimeric antigen-binding transmembrane protein containing variable and constant regions and associated with the signal-transducing multimeric CD3 complex. (Figure 24-30)

telomerase
The enzyme which adds repeated telomere sequences that are needed to protect the ends of linear chromosomes.

telomere
Region at each end of a eukaryotic chromosome containing multiple tandem repeats of a short telomeric (TEL) sequence. Telomeres are required for proper chromosome **segregation** and are replicated by a special process that prevents shortening of chromosomes during DNA replication. (Figure 7-40)

telophase
Final mitotic stage, during which the nuclear envelope re-forms around the two sets of separated chromosomes, the chromosomes decondense, and division of the cytoplasm (cytokinesis) is completed. (Figures 18-38 and 18-39)

temperature-sensitive (ts) mutation
A mutation that produces a wild-type phenotype at one temperature (the permissive temperature) but a mutant phenotype at another temperature (the nonpermissive temperature). This type of mutation is especially useful in identification of genes essential for life. (Figure 6-6)

termination
The end of the process of **transcription**.

termination, transcription
Cessation of the synthesis of an RNA chain. (Figure 5-24)

tertiary structure
In proteins, overall three-dimensional form of a polypeptide chain, which is stabilized by multiple noncovalent interactions between side chains. (Figure 3-10a)

thick filaments
Myosin-II-containing filaments of muscle in which the myosin molecules assemble into bipolar structures.

thin filaments
Actin-containing filaments of muscle.

3′ cleavage/polyadenylation
Transcription by RNA polymerase II terminates at any one of multiple termination sites downstream from the poly(A) site, which is located at the 3′ end of the final exon. After the primary transcript is cleaved at the poly(A) site a string of adenosine (A) residues is added. (Figure 9-2)

thylakoid lumen
A single continuous compartment constituted by the spaces within all the thylakoids.

thylakoid membrane
A membrane separating the stroma and the thylakoid spaces. Protons in the thylakoid lumen are pumped across the thylakoid membrane to generate a proton-motive force.

thylakoids
Flattened membranous sacs in a chloroplast that can be arranged in stacks and contain the photosynthetic pigments and **photosystems**. (Figure 12-33)

thrombopoietin
A cytokine that stimulates the development of megakaryocytes, cells that form the platelets involved in blood clotting.

thymosin-β_4
A protein that binds reversibly to ATP-actin in such a way that it cannot polymerize; this interaction provides a reservoir of actin for polymerization when it dissociates from thymosin-β_4.

tight junction
A type of cell-cell junction between the plasma membranes of adjacent epithelial cells that prevents diffusion of macromolecules and many small molecules and ions in the spaces between cells and diffusion of membrane components between the apical and basolateral regions of the plasma membrane. (Figure 20-17)

tolerance
The absence of an immune response to a particular antigen or set of antigens.

Toll-like receptor (TLR)
Member of a class of cell-surface and intracellular receptors that recognize a variety of microbial products. Ligand binding initiates a signaling pathway that induces various responses depending on the cell type. (Figure 24-37)

topogenic sequences

Segments within a protein whose sequence, number, and arrangement direct the insertion and orientation of various classes of transmembrane proteins into the endoplasmic reticulum membrane. (Figure 13-14)

topology
The basic structural arrangement of a membrane protein according to the orientation of transmembrane segment within the membrane. The membrane topology usually remains fixed after it is established during synthesis in the ER.

totipotent
Describes a stem cell that is able to give rise to all of the types of differentiated cells in an organism. Also describes the zygote, as it is able to generate every cell type in the body as well as placental cells.

trans (intercellular) interactions
Binding interactions between molecules on the plasma membranes of different cells.

transcellular pathway
The pathway by which molecules and ions move from one side of an epithelium to the other by crossing the apical and basolateral membranes via transporters, pumps, or channels.

transcellular transport
A two-stage process whereby molecules are imported through the plasma membrane on the **apical** side of the cells and exported through the plasma membrane on the **basolateral** (blood-facing) side (see Figures 11-25d and 11-30).

transcription
Process in which one strand of a DNA molecule is used as a template for synthesis of a **complementary** RNA by RNA polymerase. (Figures 5-23 and 5-24)

transcription-control region
A specific DNA sequence that serves as the binding site for **transcription factors** that regulate **transcription**.

transcription factor
General term for any protein, other than RNA polymerase, required to initiate or regulate transcription in eukaryotic cells. *General* factors, required for transcription of all genes, participate in formation of the transcription-preinitiation complex near the start site. *Specific* factors stimulate (**activators**) or inhibit (**repressors**) transcription of particular genes by binding to their regulatory sequences.

transcription unit
A region in DNA, bounded by an initiation (start) site and termination site, that is transcribed into a single primary transcript.

transcytosis
Mechanism for transporting certain substances across an epithelial sheet that combines **receptor-mediated endocytosis** and **exocytosis**. (Figures 14-25 and 24-11)

transfection
Experimental introduction of foreign DNA into cells in culture, usually followed by expression of genes in the introduced DNA. (Figure 6-34)

transfer RNA (tRNA)
A group of small RNA molecules that function as amino acid donors during protein synthesis. Each tRNA becomes covalently linked to a particular amino acid, forming an aminoacyl-tRNA. (Figures 5-31 and 5-32)

transformation
(1) Permanent, heritable alteration in a cell resulting from the uptake and incorporation of a foreign DNA into the host-cell genome; also called *stable transfection*. (2) Conversion of a "normal" mammalian cell into a cell with cancer-like properties usually induced by treatment with a virus or other cancer-causing agent.

transforming growth factor beta (TGFβ)
A family of secreted signaling proteins that are used in the development of most tissues in most or all animals. Members of the TGFβ family more often inhibit growth than stimulate it. Mutations in TGFβ signal transduction components are implicated in human cancer, including breast cancer. (Figure 16-24)

***trans*-Golgi network (TGN)**
Complex network of membranes and vesicles that serves as a major branch point in the **secretory pathway**. Vesicles budding from this most distal Golgi compartment carry membrane and soluble proteins to the cell surface or to lysosomes. (Figures 14-2 and 14-18)

transition state
State of the reactants during a chemical reaction when the system is at its highest energy level; also called the **transition-state intermediate**.

transition-state intermediate
See **transition state**.

translation

The **ribosome**-mediated assembly of a polypeptide whose amino acid sequence is specified by the nucleotide sequence in an mRNA. (Figure 5-29)

translocon
Multiprotein complex in the membrane of the rough endoplasmic reticulum through which a nascent secretory protein enters the ER lumen as it is being synthesized. (Figure 13-7)

transmission electron microscope (TEM)
A type of electron microscopy whereby electrons pass through the specimen to form an image.

transport protein
See **membrane transport protein**.

transport vesicle
A small, membrane-bounded compartment that carries soluble and membrane "cargo" proteins in the forward or reverse direction in the **secretory pathway**. Vesicles form by budding off from the donor organelle and release their contents by fusion with the target membrane.

transporters
Membrane proteins that undergo conformational changes as they move a wide variety of ions and molecules across cell membranes at a slower rate than channels. See *uniporter*, *symporter*, and *antiporter* in Figure 11-3.

transposable (mobile) DNA element
Any DNA sequence that is not present in the same chromosomal location in all individuals of a species and can move to a new position by **transposition**; also called *mobile DNA element* and *interspersed repeat*. (Table 7-1)

transposition
Movement of a **transposable DNA element** within the genome; occurs by a cut-and-paste mechanism or copy-and-paste mechanism depending on the type of element. (Figure 7-8)

treadmilling
The two ends of a myosin-decorated actin filament grow at different rates. Because the critical concentrations for the two ends are different, at steady state the free ATP–G-actin will be intermediate between C_{c^+} and C_{c^-}, so the (+) end will grow and the (−) end will lose subunits. This phenomenon is known as *treadmilling* because particular subunits appear to move through the filament. (Figure 17-10b)

triacylglycerol (triacylglyceride)
Major form in which fatty acids are stored and transported in animals; consists of three fatty acyl chains esterified to a glycerol molecule.

trophectoderm (TE)
The part of an early mammalian embryo that will form the extra-embryonic tissues, including the placenta but not the embryo proper.

trophic factor
Any of numerous signaling proteins required for the survival of cells in multicellular organisms; in the absence of such signals, cells often undergo "suicide" by **apoptosis**.

tropomodulin
A protein that binds the (−) end of an actin filament to inhibit assembly and disassembly.

tropomyosin
A long rope-like coiled-coil protein that binds along the sides of actin filaments to stabilize them, and also binds the troponin complex involved in Ca^{2+}-mediated skeletal muscle contraction.

t-SNAREs
See **SNAREs**.

tubulin
A family of globular cytoskeletal proteins that polymerize to form the cylindrical wall of **microtubules**. (Figure 18-3)

tumor microenvironment
The complex multicellular structure of a tumor, including multiple subtypes of cancer cells and normal cells that have been recruited to the tumor. T cells from the immune system are an important component of the tumor microenvironment.

tumor suppressor gene
Any gene whose encoded protein directly or indirectly inhibits progression through the cell cycle and in which a loss-of-function mutation is oncogenic. Inheritance of a single mutant allele of many tumor suppressor genes (e.g., *RB, APC,* and *BRCA1*) greatly increases the risk for developing colorectal and other types of cancer. (Figure 25-21)

tumorigenesis
The incremental process by which a tumor develops, starting from an altered somatic cell that acquires the ability to grow in an uncontrolled fashion.

ubiquinone [coenzyme Q (CoQ)]
A hydrophobic small, organic molecule that carries two protons and two electrons in electron-transport pathways, cycling between the oxidized (ubiquinone) and reduced (dihydroquinone) forms.

ubiquitin
A small protein that can be covalently linked to other intracellular proteins, thereby tagging these proteins for degradation by the **proteasome**, sorting to the lysosome, or alteration in the function of the target protein. (Figure 3-32)

ubiquitinylation
The enzymatic, covalent attachment of the small protein ubiquitin to a protein.

uncoupler
Any natural substance (e.g., the protein thermogenin) or chemical agent (e.g., 2,4-dinitrophenol) that dissipates the **proton-motive force** across the inner mitochondrial membrane or thylakoid membrane of chloroplasts, thereby inhibiting ATP synthesis.

uncoupling protein 1 (UCP1)
A transporter protein in the inner mitochondrial membrane, especially in brown-fat adipocytes, that uncouples oxidative phosphorylation and ATP generation by dissipating the proton-motive force via transporting of protons from the innermembrane space into the matrix.

unfolded-protein response
The stress-response pathway that responds to accumulation of unfolded proteins in the ER by activating expression of enzymes for protein folding in the ER.

uniport
A transmembrane protein that transports a single type of molecule down its concentration gradient.

unsaturated
Referring to a compound (e.g., fatty acid) in which one of the carbon-carbon bonds is a double or triple bond.

upstream
(1) For a gene, the direction opposite to that in which RNA polymerase moves during transcription. Nucleotides upstream from the +1 position (the first transcribed nucleotide) are designated −1, −2, and so on. (2) Events that occur earlier in a cascade of steps (e.g., signaling pathway). See also **downstream**.

upstream activating sequence (UAS)
Any protein-binding regulatory sequence in the DNA of yeast and other simple eukaryotes that is necessary for maximal gene expression; equivalent to an enhancer or promoter-proximal element in higher eukaryotes. (Figure 8-18)

vaccine
An innocuous preparation derived from a pathogen and designed to elicit an immune response in order to provide immunity against a future challenge by a virulent form of the same pathogen.

vacuole
A membrane-limited plant organelle that stores water, ions, and small-molecule nutrients and may have a degradative function similar to that of lysosomes in animal cells.

van der Waals interaction
A weak **noncovalent interaction** due to small, transient asymmetric electron distributions around atoms (dipoles). (Figure 2-10)

vector
In cell biology, an autonomously replicating genetic element used to carry a cDNA or fragment of genomic DNA into a host cell for the purpose of gene cloning. Commonly used vectors are bacterial plasmids and modified bacteriophage genomes. See also **shuttle vector**. (Figure 6-13)

viral envelope
A phospholipid bilayer forming the outer covering of some viruses (e.g., influenza and rabies viruses); derived by budding from a host-cell membrane and contains virus-encoded glycoproteins. (Figure 5-41)

virion
An individual viral particle.

virus
A small, intracellular parasite, consisting of nucleic acid (RNA or DNA) enclosed in a protein coat, that can replicate only in a susceptible host cell; widely used in cell biology research. (Figure 5-40)

V_{max}
Parameter that describes the maximal velocity of an enzyme-catalyzed reaction or other process such as protein-mediated transport of molecules across a membrane. (Figures 3-25 and 11-5)

v-SNAREs
A family of small proteins that reside in the membrane of transport vesicles, and through their specific association with similar proteins in the target membrane bring about vesicle membrane fusion. See also **SNAREs**.

voltage-gated channels
Membrane channels that open or close in response to changes in the electrical potential across the cell membrane.

Warburg effect
Describes the observation that most cancer cells predominantly produce energy by glycolysis followed by fermentation of pyruvate to lactic acid. While normal cells use this form of energy production only under the oxygen-limiting (anaerobic) condition, cancer cells metabolize glucose in this manner even in the presence of sufficient oxygen. The process is thus also referred to as *aerobic glycolysis*.

WASp
A nucleation promoting factor involved in the activation of the Arp2/3 complex.

Wee1
protein-tyrosine kinase; phosphorylates CDKs on threonine 14 and tyrosine 15 to inhibit CDK activity.

whole transcriptome shotgun sequencing (RNA-Seq)
A method that allows sequencing of the transcribed portion of the genome. mRNAs extracted from a tissue sample are converted to cDNA with the enzyme reverse transcriptase. The mRNAs present in the sample and their relative abundance can be determined by sequencing of a large number of cDNAs.

wild type
Normal, nonmutant form of a gene, protein, cell, or organism.

Wnt
A family of secreted signaling proteins used in the development of most tissues in most or all animals. Mutations in Wnt signal transduction components are implicated in human cancer, especially colon cancer. Receptors are Frizzled-class proteins with seven transmembrane segments. (Figure 16-26)

xeroderma pigmentosum
A rare inherited disease due to inactivating mutations in any of seven genes encoding proteins involved in DNA nucleotide excision repair. Patients are abnormally sensitive to

the UV in sunlight and have a high incidence of multiple types of skin cancer.

x-ray crystallography
Commonly used technique for determining the three-dimensional structure of macromolecules (particularly proteins and nucleic acids) by passing x-rays through a crystal of the purified molecules and analyzing the diffraction pattern of discrete spots that results. (Figure 3-48)

zinc finger
Several related DNA-binding **structural motifs** composed of secondary structures folded around a zinc ion; present in numerous eukaryotic transcription factors. (Figures 3-7c and 8-25a, b)

References

Chapter 2

Alberty, R. A., and R. J. Silbey. 2005. *Physical Chemistry*, 4th ed. Wiley.

Atkins, P., and J. de Paula. 2005. *The Elements of Physical Chemistry*, 4th ed. W. H. Freeman and Company.

Berg, J. M., J. L. Tymoczko, and L. Stryer. 2007. *Biochemistry*, 6th ed. W. H. Freeman and Company.

Cantor, P. R., and C. R. Schimmel. 1980. *Biophysical Chemistry.* W. H. Freeman and Company.

Davenport, H. W. 1974. *ABC of Acid-Base Chemistry*, 6th ed. University of Chicago Press.

Eisenberg, D., and D. Crothers. 1979. *Physical Chemistry with Applications to the Life Sciences.* Benjamin-Cummings.

Guyton, A. C., and J. E. Hall. 2000. *Textbook of Medical Physiology*, 10th ed. Saunders.

Hill, T. J. 1977. *Free Energy Transduction in Biology.* Academic Press.

Klotz, I. M. 1978. *Energy Changes in Biochemical Reactions.* Academic Press.

Murray, R. K., et al. 1999. *Harper's Biochemistry*, 25th ed. Lange.

Nicholls, D. G., and S. J. Ferguson. 1992. *Bioenergetics 2.* Academic Press.

Oxtoby, D., H. Gillis, and N. Nachtrieb. 2003. *Principles of Modern Chemistry*, 5th ed. Saunders.

Sharon, N. 1980. Carbohydrates. *Sci. Am.* 243(5):90–116.

Tanford, C. 1980. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, 2d ed. Wiley.

Tinoco, I., K. Sauer, and J. Wang. 2001. *Physical Chemistry—Principles and Applications in Biological Sciences*, 4th ed. Prentice Hall.

Van Holde, K., W. Johnson, and P. Ho. 1998. *Principles of Physical Biochemistry.* Prentice Hall.

Voet, D., and J. Voet. 2004. *Biochemistry*, 3d ed. Wiley.

Wood, W. B., et al. 1981. *Biochemistry: A Problems Approach*, 2d ed. Benjamin-Cummings.

Chapter 3

Websites

Entry site into proteins, structures, genomes, and taxonomy: http://www.ncbi.nlm.nih.gov/Entrez/

The protein 3-D structure database: http://www.rcsb.org/

Structural classifications of proteins: http://scop.berkeley.edu/

Sites containing general information about proteins: http://www.expasy.ch/; http://www.proweb.org/; http://scop.berkeley.edu/intro.html

PROSITE database of protein families and domains: http://www.expasy.org/prosite/

Domain organization of proteins and large collection of multiple sequence alignments: http://www.sanger.ac.uk/Software/Pfam/; http://people.cryst.bbk.ac.uk/,ubcg16z/cpn/elmovies.html

MitoCarta: An Inventory of Mammalian Mitochondrial Genes: http://www.broadinstitute.org/pubs/MitoCarta/index.html

Human protein atlas with expression of proteins in different tissues: http://www.proteinatlas.org/

Comparison of Crystallography, NMR and EM

https://www.creative-biostructure.com/comparison-of-crystallography-nmr-and-em_6.htm

Hierarchical Structure of Proteins

Dunker, A. K., et al. 2015. Intrinsically disordered proteins and multicellular organisms. *Semin. Cell Dev. Biol.* **37**:44–55.

Levitt, M. 2009. Nature of the protein universe. *Proc. Nat'l. Acad. Sci. USA* **106**:11079–11084.

Patthy, L. 1999. *Protein Evolution.* Blackwell Science.

Vogel, C., and C. Chothia. 2006. Protein family expansions and biological complexity. *PLoS Comput. Biol.* **2**(5):e48.

Yaffe, M. B. 2006. “Bits” and pieces. *Sci. STKE* **2006**:pe28.

Protein Folding

AlQuraishi, M. 2020. A watershed moment for protein structure prediction. *Nature*, in press.

https://www.nature.com/articles/d41586-019-03951-0

Baker, D. 2019. What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**(4):678–683.

Brandvold, K. R., and R. I. Morimoto. 2015. The chemical biology of molecular chaperones—implications for modulation of proteostasis. *J. Mol. Biol*. **427**:2931–2947.

Coulson, A. F., and J. Moult. 2002. A unifold, mesofold, and superfold model of protein fold use. *Proteins* **46**:61–71.

Daggett, V., and A. R. Fersht. 2003. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* **28**:18–25.

Dobson, C. M. 1999. Protein misfolding, evolution, and disease. *Trends Biochem. Sci*. **24**:329–332.

Jackrel, M. E., et al. 2014. Potentiated Hsp104 variants antagonize diverse proteotoxic misfolding events. *Cell* **156**:170–182.

Knowles, T. P., M. Vendruscolo, and C. M. Dobson. 2014. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol*. **15**:384–396.

Lavery, L. A., et al. 2014. Structural asymmetry in the closed state of mitochondrial Hsp90 (TRAP1) supports a two-step ATP hydrolysis mechanism. *Mol. Cell* **53**:330–343.

Saibil, H. 2013. Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Mol. Cell Biol.* **14**:630–642

Scaiewicz, A., and M. Levitt. 2018. Unique function words characterize genomic proteins. *Proc. Nat’l. Acad. Sci. USA* **115**(26):6703–6708.

Schmidpeter, P. A., and F. X. Schmid. 2015. Prolyl isomerization and its catalysis in protein folding and protein function. *J. Mol. Biol*. **427**:1609–1631.

Taipale, M., D. F. Jarosz, and S. Lindquist. 2010. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol*. **11**:515–528.

Valastyan, J. S., and S. Lindquist. 2014. Mechanisms of protein-folding diseases at a glance. *Dis. Model Mech*. 7:9–14.

Yu, L., D. K. Tanwar, E. D. S. Penha, Y. I. Wolf, E. V. Koonin, and M. K. Basu. 2019. Grammar of protein domain architectures. *Proc. Nat'l. Acad. Sci. USA* **116**(9):3636–3645.

Protein Binding and Enzyme Catalysis

Fersht, A. 1999. *Enzyme Structure and Mechanism*, 3d ed. W. H. Freeman and Company.

Martínez Cuesta, S., et al. 2015. The classification and evolution of enzyme function. *Biophys. J.* **109**:1082–1086.

Radisky, E. S., et al. 2006. Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. *Proc. Nat'l. Acad. Sci. USA* **103**:6835–6840.

Biomolecular Condensates

Alberti, S., A. Gladfelter, and T. Mittag. 2019. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176**(3):419–434.

Bracha, D., M. T. Walls, and C. P. Brangwynne. 2019. Probing and engineering liquid-phase organelles. *Nat. Biotechnol.* **37**(12):1435–1445.

Schaefer, K. N., and M. Peifer. 2019. Wnt/beta-catenin signaling regulation and a role for biomolecular condensates. *Dev. Cell* **48**(4):429–444.

Su, X, et al. 2016. Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**(6285):595–599.

Yoo, H., C. Triandafillou, and D. A. Drummond. 2019. Cellular sensing by phase separation: using the process, not just the products. *J. Biol. Chem.* **294**(18):7151–7159.

Regulating Protein Function

Adams, J. A. 2001. Kinetic and catalytic mechanisms of protein kinases. *Chem. Rev.* **101**(8):2271–2290.

Bellelli, A., et al. 2006. The allosteric properties of hemoglobin: insights from natural and site directed mutants. *Curr. Prot. Pep. Sci.* **7**:17–45.

Campbell, M. G., et al. 2015. 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *eLife*. 10.7554/eLife.06380.

Creixell, P., et al. 2018. Hierarchical organization endows the kinase domain with regulatory plasticity. *Cell Syst.* **7**(4):371–383.e4.

Dahiya, V., and J. Buchner. 2019. Functional principles and regulation of molecular chaperones. *Adv. Protein Chem. Struct. Biol.* **114**:1–60.

Dong, Y., et al. 2019. Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature* **565**(7737):49–55.

Finley, D., and M. A. Prado. 2020. The proteasome and its network: engineering for adaptability. *Cold Spring Harb. Perspect. Biol.* **12**(1). pii: a033985.

Gerlits, O., J. Tian, A. Das, P. Langan, W. T. Heller, and A. Kovalevsky. 2015. Phosphoryl transfer reaction snapshots in crystals: insights into the mechanism of protein kinase a catalytic subunit. *J. Biol. Chem.* **290:**15538–15548.

Glickman, M. H., and A. Ciechanover. 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **82**:373–428.

Goldberg, A. L, S. J. Elledge, and J. W. Harper. 2001. The cellular chamber of doom. *Sci. Am.* **284**:68–73.

Goldberg, A. L. 2003. Protein degradation and protection against misfolded or damaged proteins. *Nature* **426**:895–899.

Haakonsen, D. L., and M. Rape. 2019. Branching out: improved signaling by heterotypic ubiquitin chains. *Trends Cell Biol.* **29**(9):704–716.

Kern, D., and E. R. Zuiderweg. 2003. The role of dynamics in allosteric regulation. *Curr. Opin. Struc. Biol.* **13**:748–757.

Kisselev, A. F., A. Callard, and A. L. Goldberg. 2006. Importance of the different proteolytic sites of the proteasome and the efficacy of inhibitors varies with the protein substrate. *J. Biol. Chem.* **281**:8582–8590.

Lim, W. A. 2002. The modular logic of signaling proteins: building allosteric switches from simple binding domains. *Curr. Opin. Struc. Biol.* **12**:61–68.

Madhusudan, E. A., et al. 1994. cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci.* **3**(2):176–187.

Sahtoe, D. D., and T. K. Sixma. 2015. Layers of DUB regulation. *Trends Biochem. Sci.* **40**(8):456–467.

Schwartz, P. A., and B. W. Murray. 2011. Protein kinase biochemistry and drug discovery. *Bioorg. Chem.* **39**(5–6):192–210.

Sowa, M. E., et al. 2009. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**:389–403.

Ubersax, J. A., and J. E. Ferrell, Jr. 2007. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **8**(7):530–541.

Verba, K.A., and D. A. Agard. 2017. How Hsp90 and Cdc37 lubricate kinase molecular switches. *Trends Biochem. Sci.* **42**(10):799–811.

Purifying, Detecting, and Characterizing Proteins

Arthanari, H., K. Takeuchi, A. Dubey, and G. Wagner. 2019. Emerging solution NMR methods to illuminate the structural and dynamic properties of proteins. *Curr. Opin. Struct. Biol.* **58**:294–304.

Engen, J. R., et al. 2013. Partial cooperative unfolding in proteins as observed by hydrogen exchange mass spectrometry. *Int. Rev. Phys. Chem.* **32**:96–127.

Hames, B. D. *A Practical Approach.* Oxford University Press. A methods series that describes protein purification methods and assays.

Herzik, M. A. Jr., M. Wu, and G. C. Lander. 2019. High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nat. Commun.* **10**(1):1032.

Laskowski, R. A. 2016. Protein structure databases. *Meth. Mol. Biol.* **1415**:31–53.

Liao, M, et al. 2014. Single particle electron cryo-microscopy of a mammalian ion channel. *Curr. Opin. Struc. Biol.* **27**:1–7.

Mitra, A. K. 2019. Visualization of biological macromolecules at near-atomic resolution: cryo-electron microscopy comes of age. *Acta Crystallogr. F. Struct. Biol. Commun.* **75**(Pt 1):3–11.

Nogales, E., and S. H. Scheres. 2015. Cryo-EM: a unique tool for the visualization of macromolecular complexity. *Mol. Cell* **58**:677–689.

Rosenzweig, R., and L. E. Kay. 2014. Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annu. Rev. Biochem*. **83**:291–315.

Zhang, G., et al. 2014. Overview of peptide and protein analysis by mass spectrometry. *Curr. Protoc. Mol. Biol*. **108**:10.21.1–10.21.30.

Proteomics

Azimifar, S. B., et al. 2014. Cell-type-resolved quantitative proteomics of murine liver. *Cell Metab*. **20**:1076–1087.

Calvo, S. E., and V. K. Mootha. 2010. The mitochondrial proteome and human disease. *Annu. Rev. Genomics Hum. Genet.* **11**:25–34.

Chen, C. L., and N. Perrimon. 2017. Proximity-dependent labeling methods for proteomic profiling in living cells. *Wiley Interdiscip. Rev. Dev. Biol.* **6**(4).

Cox, J., and M. Mann. 2011. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu. Rev. Biochem.* **80**:273–299.

Foster, L. J., et al. 2006. A mammalian organelle map by protein correlation profiling. *Cell* **125**:187–199.

Krogan, N. J., et al. 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**:637–643.

Mertins, P., et al. 2018. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat. Protoc.* **13**(7):1632–1661.

Rhee, H. W., et al. 2013. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**(6125):1328–1331.

Rifai, N., M. A. Gillette, and S. A. Carr. 2006. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat. Biotechnol.* **24**:971–983.

Roux, P. P., and P. Thibault. 2013. The coming of age of phosphoproteomics—from large data sets to inference of protein functions. *Mol. Cell. Proteomics* **12**:3453–3464.

Trinkle-Mulcahy, L. 2019. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Res.* 8. pii: F1000 Faculty Rev-135.

Walther, T. C., and M. Mann. 2010. Mass spectrometry-based proteomics in cell biology. *J. Cell Biol*. **190**:491–500.

Chapter 4

Growing and Studying Cells in Culture

Bissell, M. J., A. Rizki, and I. S. Mian. 2003. Tissue architecture: the ultimate regulator of breast epithelial function. *Curr. Opin. Cell Biol.* 15:753–762.

Davis, J. M., ed. 1994. *Basic Cell Culture: A Practical Approach.* IRL Press.

Edwards, B., et al. 2004. Flow cytometry for high-throughput, high-content screening. *Curr. Opin. Chem. Biol.* 8:392–398.

Eggert, U. S., and T. J. Mitchison. 2006. Small molecule screening by imaging. *Curr. Opin. Chem. Biol.* 10:232–237.

Eisenstein, M. 2018. Organoids: the body builders. *Nat. Methods* 15:19–22.

Goding, J. W. 1996. *Monoclonal Antibodies: Principles and Practice. Production and Application of Monoclonal Antibodies in Cell Biology, Biochemistry, and Immunology*, 3d ed. Academic Press.

Griffith, L. G., and M. A. Swartz. 2006. Capturing complex 3D tissue physiology in vitro. *Nat. Rev. Mol. Cell Biol.* 7:211–224.

Herzenberg, L. A., et al. 2002. The future and history of fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clin. Chem*. 48:1819–1827.

Oissi, G., A. Manfrin, and M. P. Lutolf. 2018. Progress and potential in organoid research. *Nat. Rev. Gen.* 19:671–687.

Takebe, T., and J. M. Wells. 2019. Organoids by design. *Science* 364:956–959.

Light Microscopy: Exploring Cell Structure and Visualizing Proteins Within Cells

Egner, A., and S. Hell. 2005. Fluorescence microscopy with super-resolved optical sections. *Trends Cell Biol.* 15:207–215.

Giepmans, B. N. G., et al. 2006. The fluorescent toolbox for assessing protein location and function. *Science* 312:217–224.

Huang, B., H. Babcock, and X. Zhuang. 2010. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* 143:1047–1058.

Inoué, S., and K. Spring. 1997. *Video Microscopy*, 2d ed. Plenum Press.

Komatsu, N., et al. 2011. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* 22:4647–4656.

Lakadamyali, M. 2014. Super-resolution microscopy: going live and going fast. *Chem. Phys. Chem.* 15:630636.

Lippincott-Schwartz, J. 2010. Imaging: visualizing the possibilities. *J. Cell Sci.* 123:3619–3620.

Lippincott-Schwartz, J. 2011. Emerging in vivo analyses of cell function using fluorescence imaging. *Ann. Rev. Biochem.* 80:327–332.

Matsumoto, B., ed. 2002. *Methods in Cell Biology.* Vol. 70: *Cell Biological Applications of Confocal Microscopy.* Academic Press.

Mayer, T. U., et al. 1999. Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 286:971–974.

Mayor, S., and S. Bilgrami. 2007. Fretting about FRET in cell and structural biology. In D. Zuk, ed., *Evaluating Techniques in Biochemical Research.* Cell Press.

Newman, R. H., M. Fosbrink, and J. Zhang. 2011. Genetically encoded fluorescent biosensors for tracking signaling dynamics in living cells. *Chem. Biol.* 115:3614–3666.

Roukos, V., T. Misteli, and C. K. Schmidt. 2010. Descriptive no more: the dawn of high-throughput microscopy. *Trends Cell Biol.* 20:503–506.

Sahl, S. J., S. W. Hell, and S. Kakobs. 2017. Fluorescence nanoscopy in cell biology. *Nat. Rev. Mol. Cell Biol.* 18:685–701.

Schermelleh, L., R. Heintzmann, and H. Leonhardt. 2010. A guide to super-resolution fluorescence microscopy. *J. Cell Biol.* 190:165–175.

Tsien, R. Y. 2009. Indicators based on fluorescence resonance energy transfer (FRET). *Cold Spring Harbor Protoc.*, doi:10.1101/pdb.top57.

Weigert, R., N. Porat-Shliom, and P. Amornphimoltham. 2013. Imaging cell biology in live animals: ready for prime time. *J. Cell Biol.* 201:969–979.

Zipfel, W. R., R. M. Williams, and W. W. Webb. 2003. Non-linear magic: multiphoton microscopy in the biosciences. *Nature Biotechnol.* 21:1369–1377.

Electron Microscopy: High-Resolution Imaging

Beck, M., et al. 2004. Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* 306:1387–1390.

Frey, T. G., G. A. Perkins, and M. H. Ellisman. 2006. Electron tomography of membrane-bound cellular organelles. *Ann. Rev. Biophys. Biomol. Struc.* 35:199–224.

Hyatt, M. A. 2000. *Principles and Techniques of Electron Microscopy*, 4th ed. Cambridge University Press.

Koster, A., and J. Klumperman. 2003. Electron microscopy in cell biology: integrating structure and function. *Nat. Rev. Mol. Cell Biol.* 4:SS6–SS10.

Isolation of Cell Organelles

Bainton, D. 1981. The discovery of lysosomes. *J. Cell Biol.* 91:66s–76s.

de Duve, C. 1975. Exploring cells with a centrifuge. *Science* 189:186–194. The Nobel Prize lecture of a pioneer in the study of cellular organelles.

de Duve, C., and H. Beaufay. 1981. A short history of tissue fractionation. *J. Cell Biol.* 91:293s–299s.

Foster, L. J., et al. 2006. A mammalian organelle map by protein correlation profiling. *Cell* 125:187–199.

Holtzman, E. 1989. *Lysosomes.* Plenum Press.

Palade, G. 1975. Intracellular aspects of the process of protein synthesis. *Science* 189:347–358. The Nobel Prize lecture of a pioneer in the study of cellular organelles.

Chapter 5

Structure of DNA

Arnott, S. 2006. Historical article: DNA polymorphism and the early history of the double helix. *Trends Biochem. Sci.* **31**:349–354.

Berger, J. M., and J. C. Wang. 1996. Recent developments in DNA topoisomerase II structure and mechanism. *Curr. Opin. Struc. Biol.* **6**:84–90.

Dickerson, R. E. 1992. DNA Structure from A to Z. *Meth. Enzymol.* **211**:67–111.

Kornberg, A., and T. A. Baker. 2005. *DNA Replication.* University Science, chap. 1.

Wang, J. C. 1980. Superhelical DNA. *Trends Biochem. Sci.* **5**:219–221.

Wigley, D. B. 1995. Structure and mechanism of DNA topoisomerases. *Ann. Rev. Biophys. Biomol. Struc.* **24**:185–208.

DNA Replication

DePamphilis, M. L., ed. 2006. *DNA Replication and Human Disease*. Cold Spring Harbor Laboratory Press.

Gai, D., Y. P. Chang, and X. S. Chen. 2010. Origin DNA melting and unwinding in DNA replication. *Curr. Opin. Struc. Biol.* **20**:756–762.

Kornberg, A., and T. A. Baker. 2005. *DNA Replication.* University Science.

Kunkel, T. A., and K. Bebenek. 2000. DNA replication fidelity. *Ann. Re.v Biochem.* **69**:497–529.

Langston, L. D., C. Indiani, and M. O'Donnell. 2009. Whither the replisome: emerging perspectives on the dynamic nature of the DNA replication machinery. *Cell Cycle* **8**:2686–2691.

Langston, L. D., and M. O'Donnell. 2006. DNA replication: keep moving and don't mind the gap. *Mol. Cells* **23**:155–160.

Schoeffler, A. J., and J. M. Berger. 2008. DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Quart. Rev. Biophys.* **41**:41–101.

Stillman, B. 2008. DNA polymerases at the replication fork in eukaryotes. *Cell* **30**:259–260.

DNA Repair and Recombination

Barnes, D. E., and T. Lindahl. 2004. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Ann. Rev. Genet.* **38**:445–476.

Bell, C. E. 2005. Structure and mechanism of *Escherichia coli* RecA ATPase. *Mol. Microbiol.* **58**:358–366.

Friedberg, E. C., et al. 2006. DNA repair: from molecular mechanism to human disease. *DNA Repair* **5**:986–996.

Haber, J. E. 2000. Partners and pathways repairing a double-strand break. *Trends Genet.* **16**:259–264.

Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nature Rev. Mol. Cell Biol.* **7**:335–346.

Khuu, P. A., et al. 2006. The stacked-X DNA Holliday junction and protein recognition. *J. Mol. Recog.* **19**:234–242.

Lilley, D. M., and R. M. Clegg. 1993. The structure of the four-way junction in DNA. *Ann. Rev. Biophys. Biomol. Struc.* **22**:299–328.

Mitchell, J. R., J. H. Hoeijmakers, and L. J. Niedernhofer. 2003. Divide and conquer: nucleotide excision repair battles cancer and aging. *Curr. Opin. Cell Biol.***15**:232–240.

Orr-Weaver, T. L., and J. W. Szostak. 1985. Fungal recombination. *Microbiol. Rev.* **49**:33–58.

Shin, D. S., et al. 2004. Structure and function of the double strand break repair machinery. *DNA Repair* **3**:863–873.

Wood, R. D., M. Mitchell, and T. Lindahl. Human DNA repair genes. *Mutat. Res.* **577**:275–283.

Yoshida, K., and Y. Miki. 2004. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci.* **95**:866–871.

Transcription of Protein-Coding Genes and Formation of Functional mRNA

Brenner, S., F. Jacob, and M. Meselson. 1961. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190**:576–581.

Brueckner F., J. Ortiz, and P. Cramer. 2009. A movie of the RNA polymerase nucleotide addition cycle. *Curr. Opin. Struc. Biol.* **19**:294–299.

Murakami, K. S., and S. A. Darst. 2003. Bacterial RNA polymerases: the whole story. *Curr. Opin. Struc. Biol.* **13**:31–39.

Okamoto K., Y. Sugino, and M. Nomura. 1962. Synthesis and turnover of phage messenger RNA in E. *coli* infected with bacteriophage T4 in the presence of chloromycetin. *J. Mol. Biol.* **5**:527–534.

Steitz, T. A. 2006. Visualizing polynucleotide polymerase machines at work. *EMBO J.* **25**:3458–3468.

The Decoding of mRNA by tRNAs

Alexander, R. W., and P. Schimmel. 2001. Domain-domain communication in aminoacyl-tRNA synthetases. *Prog. Nucl. Acid Res. Mol. Biol.* **69**:317–349.

Hoagland, M. B., et al. 1958. A soluble ribonucleic acid intermediate in protein synthesis. *J. Biol. Chem.* **231**:241–257.

Ibba, M., and D. Soll. 2004. Aminoacyl-tRNAs: setting the limits of the genetic code. *Genes & Dev.* **18**:731–738.

Khorana, G. H., et al. 1966. Polynucleotide synthesis and the genetic code. *Cold Spring Harbor Symp. Quant. Biol.* **31**:39–49.

Nakanishi, K., and O. Nureki. 2005. Recent progress of structural biology of tRNA processing and modification. *Mol. Cells* **19**:157–166.

Nirenberg, M., et al. 1966. The RNA code in protein synthesis. *Cold Spring Harbor Symp. Quant. Biol.* **31**:11–24.

Rich, A., and S.-H. Kim. 1978. The three-dimensional structure of transfer RNA. *Sci. Am.* **240**:52–62 (offprint 1377).

Schimmel, P. 2008. Development of tRNA synthetases and connection to genetic code and disease. *Protein Sci.* **17**:1643–1652.

Stepwise Synthesis of Proteins on Ribosomes

Ban N., et al. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**:905–920.

Belousoff, M. J., et al. 2010. Ancient machinery embedded in the contemporary ribosome. *Biochem. Soc. Trans.* **38**:422–427.

Frank, J., and R. L. Gonzalez, Jr. 2010. Structure and dynamics of a processive Brownian motor: the translating ribosome. *Ann. Rev. Biochem.* **79**:381–412.

Hopfield, J. J. 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Nat'l. Acad. Sci. USA* **71**:4135–4139.

Jackson, R. J., et al. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Rev. Mol. Cell Biol.* **11**:113–127.

Korostelev, A., D. N. Ermolenko, and H. F. Noller. 2008. Structural dynamics of the ribosome. *Curr. Opin. Chem. Biol.* **12**:674–683.

Livingstone, M., et al. 2010. Mechanisms governing the control of mRNA translation. *Phys. Biol.* **7**(2):021001.

Sarnow, P., R. C. Cevallos, and E. Jan. 2005. Takeover of host ribosomes by divergent IRES elements. *Biochem. Soc. Trans.* **33**:1479–1482.

Steitz, T. A. 2008. A structural understanding of the dynamic ribosome machine. *Nature Rev. Mol. Cell Biol.* **9**:242–253.

Wimberly, B. T., et al. 2000. Structure of the 30S ribosomal subunit. *Nature* **407**:327–339.

Yusupova, G., and M. Yusupov. 2014. High-resolution structure of the eukaryotic 80S ribosome. *Ann. Rev. Biochem.* **83**:467–486.

Zaher, H. S., and R. Green. 2009. Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**:746–762.

Viruses: Parasites of the Cellular Genetic System

Flint, S. J., et al. 2000. *Principles of Virology: Molecular Biology, Pathogenesis, and Control.* ASM Press.a

Hull, R. 2002. *Mathews' Plant Virology*. Academic Press.

Klug, A. 1999. The tobacco mosaic virus particle: structure and assembly. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **354**:531–535.

Knipe, D. M., and P. M. Howley, eds. 2012. *Fields Virology,* 6th ed. Lippincott Williams & Wilkins.

Chapter 6

Using Genetic Analysis of Mutations to Identify and Study Genes

Costanzo, M., et al. 2016. A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**:1420.

Hartwell, L. H. 1967. Macromolecular synthesis of temperature-sensitive mutants of yeast. *J. Bacteriol.* **93**:1662.

Nüsslein-Volhard, C., and E. Wieschaus. 1980. Mutations affecting segment number and polarity in *Drosophila. Nature* **287**:795–801.

DNA Cloning and Characterization

Maniatis, T., et al. 1978. The isolation of structural genes from libraries of eucaryotic DNA. *Cell* **15**:687–701.

Nasmyth, K. A., and S. I. Reed. 1980. Isolation of genes by complementation in yeast: molecular cloning of a cell-cycle gene. *Proc. Nat'l. Acad. Sci. USA* **77**:2119–2123.

Nathans, D., and H. O. Smith. 1975. Restriction endonucleases in the analysis and restructuring of DNA molecules. *Annu. Rev. Biochem.* **44**:273–293.

Using Sequence Information to Identify Genes and Deduce their Function

Botstein, D. 2015. *Decoding the Language of Genetics.* Cold Spring Harbor Press.

Locating and Identifying Genes That Specify Human Traits

Altschuler D., M. J. Daly, and E. S. Lander. 2008. Genetic mapping in human disease. *Science* **322**:881–888.

Donis-Keller, H., et al. 1987. A genetic linkage map of the human genome. *Cell* **51**:319–337.

Using Cloned DNA Fragments to Study Gene Expression

Saiki, R. K., et al. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487–491.

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**:57–63.

Altering the Function of Specific Genes by Design

Capecchi, M. R. 1989. Altering the genome by homologous recombination. *Science* **244**:1288–1292.

Fire, A., et al. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**:806–811.

Jinek, M., et al. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**:816–821.

Qi, L. S., et al. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**: 1173–83.

Zamore, P. D., et al. 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**:25–33.

Zimmer, A. 1992. Manipulating the genome by homologous recombination in embryonic stem cells. *Annu. Rev. Neurosci.* **15**:115.

Chapter 7

Eukaryotic Gene Structure

Black, D. L. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Ann. Rev. Biochem.* **72**:291–336.

Davuluri, R. V., et al. 2008. The functional consequences of alternative promoter use in mammalian genomes. *Trends Genet.* **24**:167–177.

Wang, E. T., et al. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**:470–476.

Chromosomal Organization of Genes and Noncoding DNA

Celniker, S. E., and G. M. Rubin. 2003. The *Drosophila melanogaster* genome. *Ann. Rev. Genomics Hum. Genet.* **4**:89–117.

Crook, Z. R., and D. Housman. 2011. Huntington's disease: can mice lead the way to treatment? *Neuron* **69**:423–435.

Feuillet, C., et al. 2011. Crop genome sequencing: lessons and rationales. *Trends Plant Sci.* **16**:77–88.

Giardina, E., A. Spinella, and G. Novelli. 2011. Past, present and future of forensic DNA typing. *Nanomedicine* (Lond.) **6**:257–270.

Hannan, A. J. 2010. TRPing up the genome: tandem repeat polymorphisms as dynamic sources of genetic variability in health and disease. *Discov. Med.* **10**:314–321.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature* **431**:931–945.

Jobling, M. A., and P. Gill. 2004. Encoded evidence: DNA in forensic analysis. *Nature Rev. Genet.* **5**:739–751.

Lander, E. S., et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921.

Todd, P. K., and H. L. Paulson. 2010. RNA-mediated neurodegeneration in repeat expansion disorders. *Ann. Neurol.* **67**:291–300.

Venter, J. C., et al. 2001. The sequence of the human genome. *Science* **291**:1304–1351.

Transposable (Mobile) DNA Elements

Curcio, M. J., and K. M. Derbyshire. 2003. The outs and ins of transposition: from mu to kangaroo. *Nature Rev. Mol. Cell Biol.* **4**:865–877.

Goodier, J. L., and H. H. Kazazian, Jr. 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* **135**:23–35.

Jones, R. N. 2005. McClintock's controlling elements: the full story. *Cytogenet. Genome Res.* **109**:90–103.

Lisch, D. 2009. Epigenetic regulation of transposable elements in plants. *Ann. Rev. Plant Biol.* **60**:43–66.

Structural Organization of Eukaryotic Chromosomes

Bannister, A. J., and T. Kouzarides. 2011. Regulation of chromatin by histone modifications. *Cell Res.* **21**:381–395.

Bernstein, B. E., A. Meissner, and E. S. Lander. 2007. The mammalian epigenome. *Cell* **128**:669–681.

Horn, P. J., and C. L. Peterson. 2006. Heterochromatin assembly: a new twist on an old model. *Chromosome Res.* **14**:83–94.

Kurdistani, S. K. 2011. Histone modifications in cancer biology and prognosis. *Prog. Drug Res.* **67**:91–106.

Luger, K. 2006. Dynamic nucleosomes. *Chromosome Res.* **14**:5–16.

Luger, K., and T. J. Richmond. 1998. The histone tails of the nucleosome. *Curr. Opin. Genet. Devel.* **8**:140–146.

Nasmyth, K., and C. H. Haering. 2005. The structure and function of SMC and kleisin complexes. *Ann. Rev. Biochem.* **74**:595–648.

Schalch, T., et al. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**:138–141.

Woodcock, C. L., and R. P. Ghosh. 2010. Chromatin higher-order structure and dynamics. *Cold Spring Harbor Perspect. Biol.* **2**:a000596.

Morphology and Functional Elements of Eukaryotic Chromosomes

Armanios, M., and C. W. Greider. 2005. Telomerase and cancer stem cells. *Cold Spring Harbor Symp. Quant. Biol.***70**:205–208.

Belmont, A. S. 2006. Mitotic chromosome structure and condensation. *Curr. Opin. Cell Biol.* **18**:632–638.

Blackburn, E. H. 2005. Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett.* **579**:859–862.

Cvetic, C., and J. C. Walter. 2005. Eukaryotic origins of DNA replication: could you please be more specific? *Semin. Cell Dev. Biol.* **16**:343–353.

Froenicke, L. 2005. Origins of primate chromosomes as delineated by Zoo-FISH and alignments of human and mouse draft genome sequences. *Cytogenet. Genome Res.* **108**:122–138.

MacAlpine, D. M., and S. P. Bell. 2005. A genomic view of eukaryotic DNA replication. *Chromosome Res.* **13**:309–326.

Ohta, S., et al. 2011. Building mitotic chromosomes. *Curr. Opin. Cell Biol.* **23**:114–121.

Tanaka, T. U. 2010. Kinetochore-microtubule interactions: steps towards bi-orientation. *EMBO J.* **29**:4070–4082.

Chapter 8

Control of Gene Expression in Bacteria

Bush, M., and R. Dixon. 2012. The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. *Microbiol. Mol. Biol. R*. **76**:497–529.

Casino, P., V. Rubio, and A. Marina. 2010. The mechanism of signal transduction by two-component systems. *Curr. Opin. Struc. Biol.* **20**:763–771.

Fürtig, B., et al. 2015. Multiple conformational states of riboswitches fine-tune gene regulation. *Curr. Opin. Struc. Biol.* **30**:112–124.

Muller-Hill, B. 1998. Some repressors of bacterial transcription. *Curr. Opin. Microbiol.* **1**:145–151.

Overview of Eukaryotic Transcription

Djebali, S., et al. 2012. Landscape of transcription in human cells. *Nature* 489:101–108.

Kellis, M., et al. 2014. Defining functional DNA elements in the human genome. *Proc. Nat'l. Acad. Sci. USA* **111**:6131–6138.

RNA Polymerase II Promoters and General Transcription Factors

Sainsbury, S., C. Bernecky, and P. Cramer. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**:129–143.

Regulatory Sequences in Protein-Coding Genes and the Proteins Through Which They Function

de Wit, E., and W. de Laat. 2012. A decade of 3C technologies: insights into nuclear organization. *Gene Dev.* **26**:11–24.

ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**:57–74.

Vaquerizas, J. M., et al. 2009. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **10**:252–263.

Molecular Mechanisms of Transcription Repression and Activation

Berger, S. L. 2007. The complex language of chromatin regulation during transcription. *Nature* **447**:407–412.

Malladi, V. S., et al. 2015. Ontology application and use at the ENCODE DCC. Database (Oxford). doi: 10.1093/database/bav010.

Plaschka, C., et al. 2015. Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature* **518**:376–380.

Rothbart, S. B., and B. D. Strahl. 2014. Interpreting the language of histone and DNA modifications. *Biochim. Biophys. Acta.* **1839**:627–643.

Zaret, K. S., and J. S. Carroll. 2011. Pioneer transcription factors: establishing competence for gene expression. *Gene Dev.* **25**:2227–2241.

Regulation of Transcription-Factor Activity

Kirschke, E., et al. 2014. Glucocorticoid receptor function regulated by coordinated action of the Hsp90 and Hsp70 chaperone cycles. *Cell* **157**:1685–1697.

Epigenetic Regulation of Transcription

Derrien, T., et al. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**:1775–1789.

Gendrel, A. V., and E. Heard. 2014. Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu. Rev. Cell Dev. Bi.* **30**:561–580.

Klose, R. J., and A. P. Bird. 2006. Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* **31**:89–97.

McHugh, C. A., et al. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**:232–236.

Other Eukaryotic Transcription Systems

Moir, R. D., and I. M. Willis. 2015. Regulating maf1 expression and its expanding biological functions. *PLoS Genet.* **11**:e1004896.

Chapter 9

Processing of Eukaryotic Pre-mRNA

Bergkessel, M., G. M. Wilmes, and C. Guthrie. 2009. SnapShot: formation of mRNPs. *Cell* 136:794.

Hocine, S., R. H. Singer, and D. Grünwald. 2010. RNA processing and export. *Cold Spring Harb. Perspect. Biol.* 2(12):a000752.

Houseley, J., and D. Tollervey. 2009. The many pathways of RNA degradation. *Cell* 136:763–776.

Lambowitz, A. M., and S. Zimmerly. 2004. Mobile group II introns. *Annu. Rev. Genet.* 38:1–35.

Moore, M. J., and N. J. Proudfoot. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* 136:688–700.

Sharp, P. A. 2005. The discovery of split genes and RNA splicing. *Trends Biochem. Sci.* 30:279–281.

Shi, Y., and J. L. Manley. 2015. The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev.* 29:889–897.

Valadkhan, S. 2010. Role of the snRNAs in spliceosomal active site. *RNA Biol.* 7:345–353.

Wahl, M. C., C. L. Will, and R. Lührmann. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell* 136:701–718.

Regulation of Pre-mRNA Processing

Licatalosi, D. D., and R. B. Darnell. 2010. RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.* 11:75–87.

Maniatis, T., and B. Tasic. 2002. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* 418:236–243.

Raponi, M., and D. Baralle. 2010. Alternative splicing: good and bad effects of translationally silent substitutions. *FEBS J.* 277:836–840.

Wang, E. T., et al. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476.

Zheng, S., and D. L. Black. 2013. Alternative pre-mRNA splicing in neurons: growing up and extending its reach. *Trends Genet.* 29:442–448.

Zhong, X. Y., et al. 2009. SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Curr. Opin. Genet. Dev.* 19:424–436.

Transport of mRNA Across the Nuclear Envelope

Field, M. C., L. Koreny, and M. P. Rout. 2014. Enriching the pore: splendid complexity from humble origins. *Traffic* 15:141–156.

Folkmann, A. W., et al. 2011. Dbp5, Gle1-IP6 and Nup159: a working model for mRNP export. *Nucleus* 2:540–548.

Grünwald, D., R. H. Singer, and M. Rout. 2011. Nuclear export dynamics of RNA-protein complexes. *Nature* 475:333–341.

Cytoplasmic Mechanisms of Post-Transcriptional Control

Ambros, V. 2004. The functions of animal microRNAs. *Nature* 431:350–355.

Bar-Peled, L., and D. M. Sabatini. 2014. Regulation of mTORC1 by amino acids. *Trends Cell Biol.* 24:400–406.

Buchan, J. R., and R. Parker. 2009. Eukaryotic stress granules: the ins and outs of translation. *Mol. Cell* 36:932–941.

Doma, M. K., and R. Parker. 2007. RNA quality control in eukaryotes. *Cell* 131:660–668.

Ghildiyal, M., and P. D. Zamore. 2009. Small silencing RNAs: an expanding universe. *Nat Rev Genet.* 10(2):94–108. doi: 10.1038/nrg2504.

Processing of rRNA and tRNA

Hopper, A. K., and H. Y. Huang. 2015. Quality control pathways for nucleus-encoded eukaryotic tRNA biosynthesis and subcellular trafficking. *Mol. Cell Biol.* **35**:2052–2058.

Januszyk, K., and C. D. Lima. 2014. The eukaryotic RNA exosome. *Curr. Opin. Struct. Biol.* **24**:132–140.

Kressler, D., E. Hurt, and J. Bassler. 2010. Driving ribosome assembly. *Biochim. Biophys. Acta* **1803**:673–683.

Stahley, M. R., and S. A. Strobel. 2006. RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis. *Curr. Opin. Struct. Biol.* **16**:319–326.

Turowski, T. W., and D. Tollervey. 2015. Cotranscriptional events in eukaryotic ribosome synthesis. *Wiley Interdiscip. Rev. RNA* **6**:129–139

Nuclear Bodies Are Functionally Specialized Nuclear Domains

Handwerger, K. E., and J. G. Gall. 2006. Subnuclear organelles: new insights into form and function. *Trends Cell Biol.* **16**:19–26.

Chapter 10

The Lipid Bilayer: Composition and Structural Organization

McMahon, H., and J. L. Gallop. 2005. Membrane curvature and mechanisms of dynamic cell membrane remodeling. *Nature* **438**:590–596.

Mukherjee, S., and F. R. Maxfield. 2004. Membrane domains. *Annu. Rev. Cell Dev. Biol.* **20**:839–866.

Ploegh, H. 2007. A lipid-based model for the creation of an escape hatch from the endoplasmic reticulum. *Nature* **448**:435–438.

Simons, K., and D. Toomre. 2000. Lipid rafts and signal transduction. *Nat. Rev. Mol. Cell Biol.* **1**:31–41.

Simons, K., and W. L. C. Vaz. 2004. Model systems, lipid rafts, and cell membranes. *Annu. Rev. Biophys. Biomol. Struct.* **33**:269–295.

Tamm, L. K., V. K. Kiessling, and M. L. Wagner. 2001. Membrane dynamics. *Encyclopedia of Life Sciences.* Nature Publishing Group.

Vance, D. E., and J. E. Vance. 2002. *Biochemistry of lipids, lipoproteins, and membranes*, 4th ed. Elsevier.

Van Meer, G. 2006. Cellular lipidomics. *EMBO J.* **24**:3159–3165.

Yeager, P. L. 2001. Lipids. *Encyclopedia of life sciences.* Nature Publishing Group.

Zimmerberg, J., and M. M. Kozlov. 2006. How proteins produce cellular membrane curvature. *Nat. Rev. Mol. Cell Biol.* **7**:9–19.

Membrane Proteins: Structure and Basic Functions

Bowie, J. 2005. Solving the membrane protein folding problem. *Nature* **438**:581–589.

Cullen, P. J., et al. 2001. Modular phosphoinositide-binding domains: their role in signalling and membrane trafficking. *Curr. Biol.* **11**:R882–R893.

Engelman, D. 2005. Membranes are more mosaic than fluid. *Nature* **438**:578–580.

Lanyi, J. K., and H. Luecke. 2001. Bacteriorhodopsin. *Curr. Opin. Struct. Biol.* **11**:415–519.

Lee, A. G. 2005. A greasy grip. *Nature* **438**:569–570.

MacKenzie, K. R., J. H. Prestegard, and D. M. Engelman. 1997. A transmembrane helix dimer: structure and implications. *Science* **276**:131–133.

McIntosh, T. J., and S. A. Simon. 2006. Roles of bilayer material properties in function and distribution of membrane proteins. *Annu. Rev. Biophys. Biomol. Struct*. **35**:177–198.

Schulz, G. E. 2000. β-Barrel membrane proteins. *Curr. Opin. Struct. Biol.* **10**:443–447.

Wucherpfennig, K. W., et al. 2010. *Cold Spring Harb. Perspect. Biol.* **2**:a005140.

Phospholipids, Sphingolipids, and Cholesterol: Synthesis and Intracellular Movement

Bloch, K. 1965. The biological synthesis of cholesterol. *Science* **150**:19–28.

Daleke, D. L., and J. V. Lyles. 2000. Identification and purification of aminophospholipid flippases. *Biochim. Biophys. Acta* **1486**:108–127.

Futerman, A., and H. Riezman. 2005. The ins and outs of sphingolipid synthesis. *Trends Cell Biol.* **15**:312–318.

Hajri, T., and N. A. Abumrad. 2002. Fatty acid transport across membranes: relevance to nutrition and metabolic pathology. *Annu. Rev. Nutr.* **22**:383–415.

Henneberry, A. L., M. M. Wright, and C. R. McMaster. 2002. The major sites of cellular phospholipid synthesis and molecular determinants of fatty acid and lipid head group specificity. *Mol. Biol. Cell* **13**:3148–3161.

Holthuis, J. C. M., and T. P. Levine. 2005. Lipid traffic: floppy drives and a superhighway. *Nat. Rev. Mol. Cell Biol.* **6**:209–220.

Ioannou, Y. A. 2001. Multidrug permeases and subcellular cholesterol transport. *Nature Rev. Mol. Cell Biol.* **2**:657–668.

Kent, C. 1995. Eukaryotic phospholipid biosynthesis. *Annu. Rev. Biochem.* **64**:315–343.

Maxfield, F. R., and I. Tabas. 2005. Role of cholesterol and lipid organization in disease. *Nature* **438**:612–621.

Reese-Wagoner A., et al. 1999. *Biochim. Biophys. Acta*. **23**:1441(2–3):106–116.

Stahl, A., et al. 2001. Fatty acid transport proteins: a current view of a growing family. *Trends Endocrin. Met.* **12**(6):266–273.

van Meer, G., and H. Sprong. 2004. Membrane lipids and vesicular traffic. *Curr. Opin. Cell Biol.* **16**:373–378.

Chapter 11

Facilitated Transport of Glucose and Water

Chen, L. Q., L. S. Cheung, L. Feng, W. Tanner, and W. B. Frommer. 2015. Transport of sugars. *Annu. Rev. Biochem.* **84**:865–894.

Deng, D., C. Xu, P. Sun, J. Wu, C. Yan, M. Hu, and N. Yan. 2014. Crystal structure of the human glucose transporter GLUT1. *Nature* **510**:121–126.

Kreida, S., and S. Tornroth-Horsefield. 2015. Structural insights into aquaporin selectivity and regulation. *Curr. Opin. Struct. Biol.* **33**:126–134.

ATP-Powered Pumps and the Intracellular Ionic Environment

Dyla, M., M. Kjærgaard, H. Poulsen, and P. Nissen. 2019. Structure and mechanism of P-type ATPase ion pumps. *Annu. Rev. Biochem.* doi: 10.1146/annurev-biochem-010611-112801. [Epub ahead of print.]

Harrison, M. A., and S. P. Muench. 2018. The vacuolar ATPase—a nano-scale motor that drives cell biology. *Subcell. Biochem.* **87**:409–459.

Liu, F., Z. Zhang, A. Levit, J. Levring, K. K. Touhara, B. K. Shoichet, and J. Chen. 2019. Structural identification of a hotspot on CFTR for potentiation. *Science* **364**:1184–1188.

Murata, K. et al. 2000. *Nature* 407:599.

Oldham, M. L., A. L. Davidson, and J. Chen. 2008. Structural insights into ABC transporter mechanism. *Curr. Opin. Struct. Biol.* **18**:726–733.

Penmatsa, A., and E. Gouaux. 2014. How LeuT shapes our understanding of the mechanisms of sodium-coupled neurotransmitter transporters. *J. Physiol.* **592**:863–869.

Ramsey, B. W., et al. 2011. A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N. Engl. J. Med.* **365**:1663–1672.

Shinoda, T., H. Ogawa, F. Cornelius, and C. Toyoshima. 2009. Crystal structure of the sodium–potassium pump at 2.4 Å resolution. *Nature* **459**:446–450.

Srikant, S., and R. Gaudet. 2019. Mechanics and pharmacology of substrate selection and transport by eukaryotic ABC exporters. *Nat. Struct. Mol. Biol.* **26**:792–798.

Taylor, N. M. I., I. Manolaridis, S. M. Jackson, J. Kowal, H. Stahlberg, and K. P. Locher. 2017. Structure of the human multidrug transporter ABCG2. *Nature* **546**:504–509.

Zeuthen, T. 2001. *Trends Biochem. Sci.* 26:77.

Nongated Ion Channels and the Resting Membrane Potential

Arakawa T., et al. 2015. Crystal structure of the anion exchanger domain of human erythrocyte band 3. *Science.* **350**:680–684.

Gouaux, E., and R. Mackinnon. 2005. Principles of selective ion transport in channels and pumps. *Science* **310**:1461–1465.

Hibino, H., A. Inanobe, K. Furutani, S. Murakami, I. Findlay, and Y. Kurachi. 2010. Inwardly rectifying potassium channels: their structure, function, and physiological roles. *Physiol. Rev.* **90**:291–366.

Hille, B. 2001. *Ion channels of excitable membranes*, 3d ed. Sinauer Associates.

Cotransport by Symporters and Antiporters

Kuang, Q., P. Purhonen, and H. Hebert. 2015. Structure of potassium channels. *Cell Mol. Life Sci.* **72**:3677–2693.

Wright, E. M., D. D. F. Loo, and B. A. Hirayama. 2011. Biology of human sodium glucose transporters. *Physiol. Rev.* **91**:733–794.

Transcellular Transport

Krishnamurthy, H. et al. 2009. *Nature.* **459**:347–355.

Rao, M. 2004. Oral rehydration therapy: new explanations for an old remedy. *Annu. Rev. Physiol.* **66**:385–417.

Yazdani, S., J. R. Jaldin-Fincati, R. V. S. Pereira, and A. Klip. 2019. Endothelial cell barriers: transport of molecules between blood and tissues. *Traffic.* **20**:390–403.

Zihni, C., C. Mills, K. Matter, and M. S. Balda. 2016. Tight junctions: from simple barriers to multifunctional molecular gates. *Nat. Rev. Mol. Cell. Biol.* **17**:564–580.

New Perspectives

Cheng, Y., 2018. Single-particle cryo-EM—how did it get here and where will it go. *Science.* **361**:876–880.

Nannega, B. L., and T. Gonen. 2019. The cryo-EM method microcrystal electron diffraction (MicroED). *Nat. Methods* **16**:369–379.

Zhang, Q., and V. Cherezov. 2019. Chemical tools for membrane protein structural biology. *Curr. Opin. Struct. Biol.* **58**:278–285.

Chapter 12

First Step of Harvesting Energy from Glucose: Glycolysis

Berg, J., J. Tymoczko, and L. Stryer. 2012. *Biochemistry*, 7th ed. W. H. Freeman and Company. Chaps. 16–20.

Dasgupta, T., et al. 2014. A fundamental trade-off in covalent switching and its circumvention by enzyme bifunctionality in glucose homeostasis. *J. Biol. Chem.* **289**:13010–13025.

Depre, C., M. Rider, and L. Hue. 1998. Mechanisms of control of heart glycolysis. *Eur. J. Biochem.* **258**:277–290.

Fersht, A. 1999. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding.* W. H. Freeman and Company.

Fothergill-Gilmore, L. A., and P. A. Michels. 1993. Evolution of glycolysis. *Prog. Biophys. Mol. Biol.* **59**:105–135.

Nelson, D. L., and M. M. Cox. 2000. *Lehninger principles of biochemistry.* Worth. Chaps. 14–17, 19.

Pilkis, S. J., et al. 1995. 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase: a metabolic signaling enzyme. *Annu. Rev. Biochem.* **64**:799–835.

The Structure of Mitochondria and the Dynamics and Functions of Mitochondria

Ahn, C. S., and C. M. Metallo. 2015. Mitochondria as biosynthetic factories for cancer proliferation. *Cancer Metab.* **J3**:1.

Bonawitz, N. D., D. A. Clayton, and G. S. Shadel. 2006. Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol. Cell* **24**:813–825.

Canfield, D. E. 2005. The early history of atmospheric oxygen: homage to Robert M. Garrels. *Annu. Rev. Earth Pl. Sc.* **33**:1–36.

Friedman J. R., et al. 2011. ER tubules mark sites of mitochondrial division. *Science* **334**:358.

Giorgi, C., et al. 2015. Mitochondria-associated membranes: composition, molecular mechanisms, and physiopathological implications. *Antioxid. Redox. Signal.* **22**:995–1019.

Kalia, R., and A. Frost. 2019. Open and cut: allosteric motion and membrane fission by dynamin superfamily proteins. *Mol. Biol. Cell*. **30**(17):2097–2104.

Kamer, K. J., Y. Sancak, and V. K. Mootha. 2014. The uniporter: from newly identified parts to function. *Biochem. Bioph. Res. Co.* **449**:370–372.

Kaufman, R. J., and J. D. Malhotra. 2014. Calcium trafficking integrates endoplasmic reticulum function with mitochondrial bioenergetics. *Biochim. Biophys. Acta* **1843**:2233–2239.

Lackner, L. L. 2019. The expanding and unexpected functions of mitochondria contact sites. *Trends Cell Biol.* **29**(7):580–590.

Las, G., and O. S. Shirihai. 2014. Miro1: new wheels for transferring mitochondria. *EMBO J.* **33**:939–941.

Mishra, P., and D. C. Chan. 2014. Mitochondrial dynamics and inheritance during cell division, development and disease. *Nat. Rev. Mol. Cell Biol.* **15**:634–646.

Miyawaki, A., et al.1997. Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* **388**:882–887.

Song, M., and G. W. Dorn II. 2015. Mitoconfusion: noncanonical functioning of dynamism factors in static mitochondria of the heart. *Cell Metab.* **21**:195–205.

Spät, A., et al. 2008. High- and low-calcium-dependent mechanisms of mitochondrial calcium signalling. *Cell Calcium* **44**:51–63.

Tan, A. S., et al. 2015. Mitochondrial genome acquisition restores respiratory function and tumorigenic potential of cancer cells without mitochondrial DNA. *Cell Metab.* **21:**81–94.

Valdinocci, D., et al. 2019. Intracellular and intercellular mitochondrial dynamics in Parkinson's disease. *Front. Neurosci.* **13**:930.

Vance, J. E. 2014. MAM (mitochondria-associated membranes) in mammalian cells: lipids and beyond. *Biochim. Biophys. Acta* **1841**:595–609.

van der Merwe, C., et al. 2015. Evidence for a common biological pathway linking three Parkinson's disease-causing genes: parkin, PINK1 and DJ-1. *Eur. J. Neurosci.* **41**:1113–1125.

Wang, X., and H. H. Gerdes. 2015. Transfer of mitochondria via tunneling nanotubes rescues apoptotic PC12 cells. *Cell Death Differ.* **22**:1181–1191.

The Citric Acid Cycle and Fatty Acid Oxidation

Canfield, D. E. 2005. The early history of atmospheric oxygen: homage to Robert M. Garrels. *Annu. Rev. Earth Pl. Sc.* **33**:1–36.

Chan, D. C. 2006. Mitochondria: dynamic organelles in disease, aging, and development. *Cell* **125**(7):1241–1252.

Eaton, S., K. Bartlett, and M. Pourfarzam. 1996. Mammalian mitochondrial beta-oxidation. *Biochem. J.* **320** (Part 2):345–557.

Guest, J. R., and G. C. Russell. 1992. Complexes and complexities of the citric acid cycle in *Escherichia coli. Curr. Top. Cell. Regul.* **33**:231–247.

Krebs, H. A. 1970. The history of the tricarboxylic acid cycle. *Perspect. Biol. Med.* **14**:154–170.

Rasmussen, B., and R. Wolfe. 1999. Regulation of fatty acid oxidation in skeletal muscle. *Annu. Rev. Nutr.* **19**:463–484.

Velot, C., et al. 1997. Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry* **36**:14271–14276.

Wanders, R. J., and H. R. Waterham. 2006. Biochemistry of mammalian peroxisomes revisited. *Annu. Rev. Biochem.* **75**:295–332.

The Electron-Transport Chain and Generation of the Proton-Motive Force

Acin-Pérez, R., et al. 2008. Respiratory active mitochondrial supercomplexes. *Mol. Cell* **32**:529–539.

Babcock, G. 1999. How oxygen is activated and reduced in respiration. *Proc. Nat'l. Acad. Sci. USA* **96**:12971–12973.

Beinert, H., R. Holm, and E. Münck. 1997. Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* **277**:653–659.

Brandt, U. 2006. Energy converting NADH:quinone oxidoreductase (complex I). *Annu. Rev. Biochem.* **75**:165–187.

Brandt, U., and B. Trumpower. 1994. The protonmotive Q cycle in mitochondria and bacteria. *Crit. Rev. Biochem. Mol.* **29**:165–197.

Daiber, A. 2010. Redox signaling (cross-talk) from and to mitochondria involves mitochondrial pores and reactive oxygen species. *Biochim. Biophys. Acta* **6–7**:897–906.

Darrouzet, E., et al. 2001. Large scale domain movement in cytochrome bc1: a new device for electron transfer in proteins. *Trends Biochem. Sci.* **26**:445–451.

Dickinson, B. C., D. Srikun, and C. J. Chang. 2010. Mitochondrial-targeted fluorescent probes for reactive oxygen species. *Curr. Opin. Chem. Biol.* **14**:50–56.

Dudkina, N. V., I. M. Folea, and E. J. Boekema. 2015. Towards structural and functional characterization of photosynthetic and mitochondrial supercomplexes. *Micron* **72**:39–51.

Efremov, R. G., R. Baradaran, and L. A. Sazanov. 2010. The architecture of respiratory complex I. *Nature* **465**:441–445.

Finkel, T. 2011. Signal transduction by reactive oxygen species. *J. Cell Biol.* **194**:7–15.

Grigorieff, N. 1999. Structure of the respiratory NADH:ubiquinone oxidoreductase (complex I). *Curr. Opin. Struct. Biol.* **9**:476–483.

Hosler, J. P., S. Ferguson-Miller, and D. A. Mills. 2006. Energy transduction: proton transfer through the respiratory complexes. *Annu. Rev. Biochem.* **75**:165–187.

Hunte, C., V. Zickermann, and U. Brandt. 2010. Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* **329**:448–451.

Hyde, B. B., G. Twig, and O. S. Shirihai. 2010. Organellar vs cellular control of mitochondrial dynamics. *Semin. Cell Dev. Biol.* **21**:575–581.

Koopman, W. J., et al. 2010. Mammalian mitochondrial complex I: biogenesis, regulation, and reactive oxygen species generation. *Antioxid. Redox Signal.* **12:**1431–1470.

Michel, H., et al. 1998. Cytochrome *c* oxidase. *Annu. Rev. Bioph. Biom. 27*:329–356.

Mitchell, P. 1979. Keilin's respiratory chain concept and its chemiosmotic consequences. *Science* **206**:1148–1159. (Nobel Prize lecture.)

Murphy, M. P. 2009. How mitochondria produce reactive oxygen species. *Biochem. J.* **417**:1–13.

Ramirez, B. E., et al. 1995. The currents of life: the terminal electron-transfer complex of respiration. *P. Natl. Acad. Sci. USA* **92**:11949–11951.

Ruitenberg, M., et al. 2002. Reduction of cytochrome *c* oxidase by a second electron leads to proton translocation. *Nature* **417**:99–102.

Saraste, M. 1999. Oxidative phosphorylation at the fin de siècle. *Science* **283**:1488–1492.

Schafer, E., et al. 2006. Architecture of active mammalian respiratory chain supercomplexes. *J. Biol. Chem.* **281**(22):15370–15375.

Schultz, B., and S. Chan. 2001. Structures and proton-pumping strategies of mitochondrial respiratory enzymes. *Annu. Rev. Bioph. Biom.* **30**:23–65.

Sheeran, F. L., and S. Pepe. 2006. Energy deficiency in the failing heart: linking increased reactive oxygen species and disruption of oxidative phosphorylation rate. *Biochim. Biophys. Acta* **1757**(5–6):543–552.

Sies, H. 2014. Role of metabolic $\mathrm{H_2O_2}$ generation: redox signaling and oxidative stress. *J. Biol. Chem.* **289**:8735–8741.

Tsukihara, T., et al. 1996. The whole structure of the 13-subunit oxidized cytochrome *c* oxidase at 2.8 Å. *Science* **272**:1136–1144.

Walker, J. E. 1995. Determination of the structures of respiratory enzyme complexes from mammalian mitochondria. *Biochim. Biophys. Acta* **1271**:221–227.

Wallace, D. C. 2005. A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annu. Rev. Genet.* **39**:359–407.

Xia, D., et al. 1997. Crystal structure of the cytochrome bc_1 complex from bovine heart mitochondria. *Science* **277**:60–66.

Zaslavsky, D., and R. Gennis. 2000. Proton pumping by cytochrome oxidase: progress and postulates. *Biochim. Biophys. Acta* **1458**:164–179.

Zhang, M., E. Mileykovskaya, and W. Dowhan. 2005. Cardiolipin is essential for organization of complexes III and IV into a supercomplex in intact yeast mitochondria. *J. Biol. Chem.* **280**(33):29403–29408.

Zhang, Z., et al. 1998. Electron transfer by domain movement in cytochrome bc_1. *Nature* **392**:677–684.

Harnessing the Proton-Motive Force to Synthesize ATP

Aksimentiev, A., et al. 2004. Insights into the molecular mechanism of rotation in the F_0 sector of ATP synthase. *Biophys. J.* **86**(3):1332–1344.

Allegretti, M., et al. 2015. Horizontal membrane-intrinsic α-helices in the stator α-subunit of an F-type ATP synthase. *Nature* **521**:237–240.

Bianchet, M. A., et al. 1998. The 2.8 Å structure of rat liver F_1-ATPase: configuration of a critical intermediate in ATP synthesis/hydrolysis. *Proc. Nat'l. Acad. Sci. USA* **95**:11065–11070.

Boyer, P. D. 1997. The ATP synthase—a splendid molecular machine. *Annu. Rev. Biochem. 66*:717–749.

Capaldi, R., and R. Aggeler. 2002. Mechanism of the F_0F_1-type ATP synthase—a biological rotary motor. *Trends Biochem. Sci.* **27**:154–160.

Elston, T., H. Wang, and G. Oster. 1998. Energy transduction in ATP synthase. *Nature* **391**:510–512.

He, J., et al. 2018. Assembly of the membrane domain of ATP synthase in human mitochondria. *P. Natl. Acad. Sci. USA* **20**: 2988–2993.

Hinkle, P. C. 2005. P/O ratios of mitochondrial oxidative phosphorylation. *Biochim. Biophys. Acta* **1706**(1–2):1–11.

Junge, W., and N. Nelson. 2015. ATP synthase. *Annu. Rev. Biochem.* **84**:631–657.

Junge, W., S. Hendrik, and S. Engelbrecht. 2009. Torque generation and elastic power transmission in the rotary F_0F_1-ATPase.*Nature* **459**:364–370.

Kinosita, K., et al. 1998. F_1-ATPase: a rotary motor made of a single molecule. *Cell* **93**:21–24.

Klingenberg, M., and S. Huang. 1999. Structure and function of the uncoupling protein from brown adipose tissue. *Biochim. Biophys. Acta* **1415**:271–296.

Nury, H., et al. 2006. Relations between structure and function of the mitochondrial ADP/ATP carrier. *Annu. Rev. Biochem.* **75**:713–741.

Oliveira, A. S., et al. 2014. Exploring O_2 diffusion in A-type cytochrome *c* oxidases: molecular dynamics simulations uncover two alternative channels towards the binuclear site. *PLoS Comput. Biol.* **10**:e1004010.

Rosen, E. D., and B. M. Spiegelman. 2014. What we talk about when we talk about fat. *Cell* **156**:20–44.

Sharma, V., et al. 2015. Role of subunit III and its lipids in the molecular mechanism of cytochrome *c* oxidase. *Biochim. Biophys. Acta* **1847**:690–697.

Tsunoda, S., et al. 2001. Rotation of the **c** subunit oligomer in fully functional F_0F_1 ATP synthase. *Proc. Nat'l. Acad. Sci. USA* **98**:898–902.

Vercesi, A. E., et al. 2006. Plant uncoupling mitochondrial proteins. *Annu. Rev. Plant Biol.* **57**:383–404.

von Ballmoos, C., A. Wiedenmann, and P. Dimroth. 2009. Essentials for ATP synthesis by F_1F_0 ATP synthases. *Annu. Rev. Biochem.* **78**:649–672.

Wu, J., H. Jun, and J. R. McDermott. 2015. Formation and activation of thermogenic fat. *Trends Genet.* **31**:232–238.

Yasuda, R., et al. 2001. Resolution of distinct rotational substeps by submillisecond kinetic analysis of F_1-ATPase. *Nature* **410**:898–904.

Chloroplasts and Photosynthesis

Bendich, A. J. 2004. Circular chloroplast chromosomes: the grand illusion. *Plant Cell* **16**:1661–1666.

Ben-Shem, A., F. Frolow, and N. Nelson. 2003. Crystal structure of plant photosystem I. *Nature* **426**(6967):630–635.

Blankenship, R. E. 2002. *Molecular mechanisms of photosynthesis.* Blackwell.

Deisenhofer, J., and J. R. Norris, eds. 1993. *The photosynthetic reaction center.* Vols. 1 and 2. Academic Press.

McDermott, G., et al. 1995. Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria. *Nature* **364**:517.

Nelson, N., and C. F. Yocum. 2006. Structure and function of photosystems I and II. *Annu. Rev. Plant Biol.* **57**:521–565.

Prince, R. 1996. Photosynthesis: the Z-scheme revisited. *Trends Biochem. Sci.* **21**:121–122.

Wollman, F. A. 2001. State transitions reveal the dynamics and flexibility of the photosynthetic apparatus. *EMBO J.* **20**:3623–3630.

Use of Light Energy to Generate Molecular Oxygen, NADPH, and ATP in Stages 1–3 of Photosynthesis

Allen, J. F. 2002. Photosynthesis of ATP—electrons, proton pumps, rotors, and poise. *Cell* **110**:273–276.

Amunts, A., et al. 2010. Structure determination and improved model of plant photosystem I. *J. Biol. Chem*. **285**:3478–3486.

Aro, E. M., I. Virgin, and B. Andersson. 1993. Photoinhibition of photosystem II: inactivation, protein damage, and turnover. *Biochim. Biophys. Acta* **1143**:113–134.

Deisenhofer, J., and H. Michel. 1989. The photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis. Science* **245**:1463–1473. (Nobel Prize lecture.)

Deisenhofer, J., and H. Michel. 1991. Structures of bacterial photosynthetic reaction centers. *Annu. Rev. Cell Biol.* **7**:1–23.

Dekker, J. P., and E. J. Boekema. 2005. Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta* **1706**(1–2):12–39.

Finazzi, G. 2005. The central role of the green alga *Chlamydomonas reinhardtii* in revealing the mechanism of state transitions *J. Exp. Bot.* **56**(411):383–388.

Guskov, A., et al. 2010. Recent progress in the crystallographic studies of photosystem II. *ChemPhysChem.* **11**(6):1160–1171.

Haldrup, A., et al. 2001. Balance of power: a view of the mechanism of photosynthetic state transitions. *Trends Plant Sci.* **6**:301–305.

Hankamer, B., J. Barber, and E. Boekema. 1997. Structure and membrane organization of photosystem II from green plants. *Annu. Rev. Plant Phys. 48*:641–672.

Heathcote, P., P. Fyfe, and M. Jones. 2002. Reaction centres: the structure and evolution of biological solar power. *Trends Biochem. Sci.* **27**:79–87.

Horton, P., A. Ruban, and R. Walters. 1996. Regulation of light harvesting in green plants. *Annu. Rev. Plant Phys.* **47**:655–684.

Iwai, M., et al. 2010. Isolation of the elusive supercomplex that drives cyclic electron flow in photosynthesis. *Nature* **464**: 1210–1213.

Joliot, P., and A. Joliot. 2005. Quantification of cyclic and linear flows in plants. *Proc. Nat'l. Acad. Sci. USA* **102**(13):4913–4918.

Jordan, P., et al. 2001. Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* **411**:909–917.

Kühlbrandt, W. 2001. Chlorophylls galore. *Nature* **411**:896–898.

Martin, J. L., and M. H. Vos. 1992. Femtosecond biology. *Annu. Rev. Bioph. Biom.* **21**:199–222.

Nelson, N., and W. Junge. 2015. Structure and energy transfer in photosystems of oxygenic photosynthesis. *Annu. Rev. Biochem.* **84**:659–683.

Penner-Hahn, J. 1998. Structural characterization of the Mn site in the photosynthetic oxygen-evolving complex. *Struct. Bond.* **90**:1–36.

Shikanai, T. 2014. Central role of cyclic electron transport around photosystem I in the regulation of photosynthesis. *Curr. Opin. Biotech.* **26**:25–30.

Suga, M., et al. 2015. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**:99–103.

Tomizioli, M., et al. 2014. Deciphering thylakoid sub-compartments using a mass spectrometry-based approach. *Mol. Cell. Proteomics* **13**:2147–2167.

Tommos, C., and G. Babcock. 1998. Oxygen production in nature: a light-driven metalloradical enzyme process. *Accounts Chem. Res.* **31**:18–25.

ATP and NADPH Drive Carbon Fixation in the Calvin Cycle and Carbohydrate Synthesis in Stage 4 of Photosynthesis

Buchanan, B. B. 1991. Regulation of CO_2 assimilation in oxygenic photosynthesis: the ferredoxin/thioredoxin system. Perspective on its discovery, present status, and future

development. *Arch. Biochem. Biophys.* **288**:1–9.

Gutteridge, S., and J. Pierce. 2006. A unified theory for the basis of the limitations of the primary reaction of photosynthetic CO_2 fixation: was Dr. Pangloss right? *Proc. Nat'l. Acad. Sci. USA* **103**:7203–7204.

Mueller-Cajar, O., M. Stotz, and A. Bracher. 2014. Maintaining photosynthetic CO_2 fixation via protein remodeling: the Rubisco activases. *Photosynth. Res.* **119**:191–201.

Portis, A. 1992. Regulation of ribulose 1,5-bisphosphate carboxylase/oxygenase activity. *Annu. Rev. Plant Phys.* **43**:415–437.

Rawsthorne, S. 1992. Towards an understanding of C_3-C_4 photosynthesis. *Essays Biochem.* **27**:135–146.

Rokka, A., I. Zhang, and E.-M. Aro. 2001. Rubisco activase: an enzyme with a temperature-dependent dual function? *Plant J.* **25**:463–472.

Sage, R., and J. Colemana. 2001. Effects of low atmospheric CO_2 on plants: more than a thing of the past. *Trends Plant Sci.* **6**:18–24.

Schneider, G., Y. Lindqvist, and C. I. Branden. 1992. Rubisco: structure and mechanism. *Annu. Rev. Bioph. Biom.* **21**:119–153.

Tcherkez, G. G., G. D. Farquhar, and T. J. Andrews. 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Nat'l. Acad. Sci. USA* **103**(19):7246–7251.

Wolosiuk, R. A., M. A. Ballicora, and K. Hagelin. 1993. The reductive pentose phosphate cycle for photosynthetic CO_2 assimilation: enzyme modulation. *FASEB J.* **7:**622–637.

Additional references

Aguilera-Alvarado, G. P., and S. Sánchez-Nieto. 2017. Plant hexokinases are multifaceted proteins. *Plant Cell Physiol.* **58**(7):1151–1160.

Anderson, A. J., et al. 2019. Mitochondria-hubs for regulating cellular biochemistry: emerging concepts and networks. *Open Biol.* **9**(8):190126.

Bailey-Serres, J., et al. 2019. Genetic strategies for improving crop yields. *Nature* **575**(7781):109–118.

Bennett, D. I. G., et al. 2019. Models and mechanisms of the rapidly reversible regulation of photosynthetic light harvesting. *Open Biol.* **9**(4):190043. doi: 10.1098/rsob.190043.

Bertholet, A. M., et al. 2019. H(+) transport is an integral function of the mitochondrial ADP/ATP carrier. *Nature* **571**:515–520.

Cho, H. M., and J. R. Ryu. 2019. Drp1-Zip1 interaction regulates mitochondrial quality surveillance system. *Mol Cell.* **73**(2):364–376.e8.

Cohen, S., A. M. Valm, and J. Lippincott-Schwartz. 2018. Interacting organelles. *Curr. Opin. Cell Biol.* **53**:84–91.

Depaoli, M. R., et al. 2019. The enigmatic ATP supply of the endoplasmic reticulum. *Biol. Rev. Camb. Philos. Soc.* **94**(2):610–628.

Gao, J., et al. 2018. Structure and function of the photosystem supercomplexes. *Front. Plant Sci.* **9**:357.

Grossmann, D., and C. Berenguer-Escuder. 2019. Mutations in RHOT1 disrupt endoplasmic reticulum-mitochondria contact sites interfering with calcium homeostasis and mitochondrial dynamics in Parkinson's disease. *Antioxid. Redox Signal.* **31**(16):1213–1234.

Guo, H., et al. 2018. Cryo-EM of ATP synthases. *Curr. Opin. Struct. Biol.* **52**:71–79.

Guo, R., et al. 2017. Architecture of human mitochondrial respiratory megacomplex $\mathrm{I_2III_2IV_2}$. *Cell* **170**, 1247–1257.

Guo, Y., et al. 2018. Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales. *Cell* **175**(5):1430–1442.e17.

Kondo, T., et al. 2019. Microsecond and millisecond dynamics in the photosynthetic protein LHCSR1 observed by single-molecule correlation spectroscopy. *Proc. Nat'l. Acad. Sci. USA*. **116**(23):11247–11252. doi: 10.1073/pnas.1821207116. Epub 2019 May 17.

Lahiri, V., W. D. Hawkins, and D. J. Klionsky. 2019. Watch what you (self-) eat: autophagic mechanisms that modulate metabolism. *Cell Metab.* **29**(4):803–-826.

Newman, L. E., and G. S. Shadel. 2018. Pink1/Parkin link inflammation, mitochondrial stress, and neurodegeneration. *J. Cell Biol.* **217**(10):3327–3329.

Ruprecht, J. J., and E. R. Kunji. 2019. *Curr. Opin. Struct. Biol.* **57**:135–144. Structural changes in the transport cycle of the mitochondrial ADP/ATP carrier.

Scorrano, L., et al. 2019. Coming together to define membrane contact sites. *Nat. Commun.* **10**(1):1287.

Shen, L., et al. 2019. Structure of a C(2)S(2)M(2)N(2)-type PSII-LHCII supercomplex from the green alga Chlamydomonas reinhardtii. *Proc. Nat'l. Acad. Sci. USA.* **116**(42):21246–21255.

Singharoy, A., et al. 2019. Atoms to phenotypes: molecular design principles of cellular energy metabolism. *Cell* **179**(5):1098–1111.

Yang, C., and T. M. Svitkina 2019. Ultrastructure and dynamics of the actin-myosin II cytoskeleton during mitochondrial fission. *Nat. Cell Biol.* **21**(5):603–613.

Yeshaw, W. M., et al. 2019. Human VPS13A is associated with multiple organelles and influences mitochondrial morphology and lipid droplet motility. *eLife*. Feb 11;8. pii: e43561.

Yong, J., et al .2019. Mitochondria supply ATP to the ER through a mechanism antagonized by cytosolic Ca^{2}. *eLife*. Sep 9;8. pii: e49682.

Chapter 13

Targeting Proteins to and Across the ER Membrane

Egea, P. F., R. M. Stroud, and P. Walter. 2005. Targeting proteins to membranes: structure of the signal recognition particle. *Curr. Opin. Struct. Biol.* **15**:213–220.

Osborne, A. R., T. A. Rapoport, and B. van den Berg. 2005. Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* **21**:529–550.

Wickner, W., and R. Schekman. 2005. Protein translocation across biological membranes. *Science* **310**:1452–1456.

Insertion of Membrane Proteins Into the ER

Englund, P. T. 1993. The structure and biosynthesis of glycosylphosphatidylinositol protein anchors. *Annu. Rev. Biochem.* **62**:121–138.

Mothes, W., et al. 1997. Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell* **89**:523–533.

Shao, S., and R. S. Hegde. 2011. Membrane protein insertion at the endoplasmic reticulum. *Annu. Rev. Cell Dev. Biol.* 27:25–56.

Wang, F., et al. 2011. The mechanism of tail-anchored protein insertion into the ER membrane. *Mol. Cell* **43**:738–750.

Protein Modifications, Folding, and Quality Control in the ER

Braakman, I., and N. J. Bulleid. 2011. Protein folding and modification in the mammalian endoplasmic reticulum. *Annu. Rev. Biochem.* **80**:71–99.

Hegde, R. S., and H. L. Ploegh. 2010. Quality and quantity control at the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **22**:437–446.

Helenius, A., and M. Aebi. 2004. Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73**:1019–1049.

Kornfeld, R., and S. Kornfeld. 1985. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **45**:631–664.

Patil, C., and P. Walter. 2001. Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. *Curr. Opin. Cell Biol.* **13**:349–355.

Sevier, C. S., and C. A. Kaiser. 2002. Formation and transfer of disulphide bonds in living cells. *Nat. Rev. Mol. Cell Biol.* **3**:836–847.

Smith, M. H., H. L. Ploegh, and J. S. Weissman. 2011. Road to ruin: targeting proteins for degradation in the endoplasmic reticulum. *Science* **334**:1086–1090.

Targeting of Proteins to Mitochondria and Chloroplasts

Dalbey, R. E., and A. Kuhn. 2000. Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu. Rev. Cell Dev. Biol.* **16**:51–87.

Dolezal, P., et al. 2006. Evolution of the molecular machines for protein import into mitochondria. *Science* **313**:314–318.

Koehler, C. M. 2004. New developments in mitochondrial assembly. *Annu. Rev. Cell Dev. Biol.* **20**:309–335.

Li, H.-M., and C.-C. Chiu. 2010. Protein transport into chloroplasts. *Ann. Rev. Plant Biol.* **61**:157–180.

Matouschek, A., N. Pfanner, and W. Voos. 2000. Protein unfolding by mitochondria: the Hsp70 import motor. *EMBO Rep.* **1**:404–410.

Neupert, W., and M. Brunner. 2002. The protein import motor of mitochondria. *Nat. Rev. Mol. Cell Biol.* **3**:555–565.

Robinson, C., and A. Bolhuis. 2001. Protein targeting by the twin-arginine translocation pathway. *Nat. Rev. Mol. Cell Biol.* **2**:350–356.

Schmidt, O., N. Pfanner, and C. Meisinger. 2010. Mitochondrial protein import: from proteomics to functional mechanisms. *Nat. Rev. Mol. Cell Biol.* **11**:655–667.

Schnell, D. J. 2019. The TOC GTPase receptors: regulators of the fidelity, specificity and substrate profiles of the general protein import machinery of chloroplasts. *Protein J.* **38**:343–350.

Targeting of Peroxisomal Proteins

Dammai, V., and S. Subramani. 2001. The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell* **105**:187–196.

Gould, S. J., and D. Valle. 2000. Peroxisome biogenesis disorders: genetics and cell biology. *Trends Genet.* **16**:340–345.

Hoepfner, D., et al. 2005. Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* **122**:85–95.

Ma, C., G. Agrawal, and S. Subramani. 2011. Peroxisome assembly: matrix and membrane protein biogenesis. *J. Cell Biol.* **193**:7–16.

Purdue, P. E., and P. B. Lazarow. 2001. Peroxisome biogenesis. *Annu. Rev. Cell Dev. Biol.* **17**:701–752.

Smith, J. J., and J. D. Aitchison. 2013. Peroxisomes take shape. *Nat. Rev. Mol. Cell Biol.* **14**:803–817.

Transport into and out of the Nucleus

Chook, Y. M., and G. Blobel. 2001. Karyopherins and nuclear import. *Curr. Opin. Struct. Biol.* **11**:703–715.

Cole, C. N., and J. J. Scarcelli. 2006. Transport of messenger RNA from the nucleus to the cytoplasm. *Curr. Opin. Cell Biol.* **18**:299–306.

Johnson, A. W., E. Lund, and J. Dahlberg. 2002. Nuclear export of ribosomal subunits. *Trends Biochem. Sci.* **27**:580–585.

Kabachinski, G., and T. U. Schwartz. 2015. The nuclear pore complex–structure and function at a glance. *J. Cell Sci*. **128**:423–429.

Ribbeck, K., and D. Gorlich. 2001. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* **20**:1320–1330.

Rout, M. P., and J. D. Aitchison. 2001. The nuclear pore complex as a transport machine. *J. Biol. Chem.* **276**:16593–16596.

Stewart, M. 2010. Nuclear export of mRNA. *Trends Biochem. Sci.* **35**:609–617.

Terry, L. J., and S. R. Wente. 2009. Flexible gates: dynamic topologies and functions for FG nucleoporins in nucleocytoplasmic transport. *Eukaryot. Cell* **8**:1814–1827.

Chapter 14

Techniques for Studying the Secretory Pathway

Beckers, C. J., et al. 1987. Semi-intact cells permeable to macromolecules: use in reconstitution of protein transport from the endoplasmic reticulum to the Golgi complex. *Cell* **50**:523–534.

Kaiser, C. A., and R. Schekman. 1990. Distinct sets of SEC genes govern transport vesicle formation and fusion early in the secretory pathway. *Cell* **61**:723–733.

Lippincott-Schwartz, J., et al. 2001. Studying protein dynamics in living cells. *Nat. Rev. Mol. Cell Biol.* **2**:444–456.

Novick, P., et al. 1981. Order of events in the yeast secretory pathway. *Cell* **25**:461–469.

Orci, L., et al. 1989. Dissection of a single round of vesicular transport: sequential intermediates for intercisternal movement in the Golgi stack. *Cell* **56**:357–368.

Palade, G. E. 1975. Intracellular aspects of the process of protein synthesis. *Science* **189**:347–358.

Molecular Mechanisms of Vesicle Budding and Fusion

Bonifacino, J. S., and B. S. Glick. 2004. The mechanisms of vesicle budding and fusion. *Cell* **116**:153–166.

Hutagalung, A. H., and P. J. Novick. 2011. Role of Rab GTPases in membrane traffic and cell physiology. *Physiol. Rev.* **91**:119–149.

Jahn, R., et al. 2003. Membrane fusion. *Cell* **112**:519–533.

Kirchhausen, T. 2000. Three ways to make a vesicle. *Nat. Rev. Mol. Cell Biol.* **1**:187–198.

McNew, J. A., et al. 2000. Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* **407**:153–159.

Ostermann, J., et al. 1993. Stepwise assembly of functionally active transport vesicles. *Cell* **75**:1015–1025.

Schimmöller, F., I. Simon, and S. Pfeffer. 1998. Rab GTPases, directors of vesicle docking. *J. Biol. Chem.* **273**:22161–22164.

Weber, T., et al. 1998. SNAREpins: minimal machinery for membrane fusion. *Cell* **92**:759–772.

Wickner, W., and A. Haas. 2000. Yeast homotypic vacuole fusion: a window on organelle trafficking mechanisms. *Annu. Rev. Biochem.* **69**:247–275.

Zerial, M., and H. McBride. 2001. Rab proteins as membrane organizers. *Nat. Rev. Mol. Cell Biol.* **2**:107–117.

Early Stages of the Secretory Pathway

Behnia, R., and S. Munro. 2005. Organelle identity and the signposts for membrane traffic. *Nature* **438**:597–604.

Bi, X., et al. 2002. Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature* **419**:271–277.

Glick, B. S., and A. Nakano. 2009. Membrane traffic within the Golgi apparatus. *Annu. Rev. Cell Dev. Biol.* **25**:113–132.

Gurkan, C., et al. 2006. The COPII cage: unifying principles of vesicle coat assembly. *Nat. Rev. Mol. Cell Biol.* **7**:727–738.

Lee, M. C., et al. 2004. Bi-directional protein transport between the ER and Golgi. *Annu. Rev. Cell Dev. Biol.* **20**:87–123.

Letourneur, F., et al. 1994. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell* **79**:1199–1207.

Lord, C., S. Ferro-Novick, and E. A. Miller. 2013. The highly conserved COPII coat complex sorts cargo from the endoplasmic reticulum and targets it to the Golgi. *Cold Spring Harb. Perspect. Biol.* **5**:a013367.

Losev, E., et al. 2006. Golgi maturation visualized in living yeast. *Nature* **441**:1002–1006.

Pelham, H. R. 1995. Sorting and retrieval between the endoplasmic reticulum and Golgi apparatus. *Curr. Opin. Cell Biol.* **7**:530–535.

Later Stages of the Secretory Pathway

Bonifacino, J. S. 2004. The GGA proteins: adaptors on the move. *Nat. Rev. Mol. Cell Biol.* **5**:23–32.

Bonifacino, J. S., and E. C. Dell'Angelica. 1999. Molecular bases for the recognition of tyrosine-based sorting signals. *J. Cell Biol.* **145**:923–926.

Edeling, M. A., C. Smith, and D. Owen. 2006. Life of a clathrin coat: insights from clathrin and AP structures. *Nat. Rev. Mol. Cell Biol.* **7**:32–44.

Fotin, A., et al. 2004. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **432**:573–579.

Ghosh, P., et al. 2003. Mannose 6-phosphate receptors: new twists in the tale. *Nat. Rev. Mol. Cell Biol.* **4**:202–213.

Mellman, I., and W. J. Nelson. 2008. Coordinated protein sorting, targeting and distribution in polarized cells. *Nat. Rev. Mol. Cell Biol.* **9**:833–845.

Schmid, S. 1997. Clathrin-coated vesicle formation and protein sorting: an integrated process. *Annu. Rev. Biochem.* **66**:511–548.

Simons, K., and E. Ikonen. 1997. Functional rafts in cell membranes. *Nature* **387**:569–572.

Steiner, D. F., et al. 1996. The role of prohormone convertases in insulin biosynthesis: evidence for inherited defects in their action in man and experimental animals. *Diabetes Metab.* **22**:94–104.

Tooze, S. A., et al. 2001. Secretory granule biogenesis: rafting to the SNARE. *Trends Cell Biol.* **11**:116–122.

Receptor-Mediated Endocytosis

Brown, M. S., and J. L. Goldstein. 1986. Receptor-mediated pathway for cholesterol homeostasis. *Science* **232**:34–47. (Nobel Prize lecture.)

Kaksonen, M., C. P. Toret, and D. G. Drubin. 2006. Harnessing actin dynamics for clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **7**:404–414.

Rudenko, G., et al. 2002. Structure of the LDL receptor extracellular domain at endosomal pH. *Science* **298**:2353–2358.

Directing Membrane Proteins and Cytosolic Materials to the Lysosome for Degradation

Geng, J., and D. J. Klionsky. 2008. The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. *EMBO Rep.* **9**:859–864.

Henne, W. M., N. J. Buchkovich, and S. D. Emr. 2011. The ESCRT pathway. *Dev. Cell* **21**:77–91.

Katzmann, D. J., et al. 2002. Receptor downregulation and multivesicular-body sorting. *Nat. Rev. Mol. Cell Biol.* **3**:893–905.

Lemmon, S. K., and L. M. Traub. 2000. Sorting in the endosomal system in yeast and animal cells. *Curr. Opin. Cell Biol.* **12**:457–466.

Shintani, T., and D. J. Klionsky. 2004. Autophagy in health and disease: a double-edged sword. *Science* **306**:990–995.

Sundquist, W. I., and H.-G. Kräusslich. 2012. HIV-1 assembly, budding, and maturation. *Cold Spring Harb. Perspect. Med.* **2**:a006924.

Chapter 15

Studying Cell-Surface Receptors and Signal Transduction Proteins

Bagur, R., and G. Hajnoczky. 2017. Intracellular Ca2+ sensing: its role in calcium homeostasis and signaling. *Mol. Cell* **66**:780.

Gross, A., and H. F. Lodish. 2006. Cellular trafficking and degradation of erythropoietin and NESP. *J. Biol. Chem.* **281**:2024–2032.

Miller, C., and B. Turk. 2018. Homing in: mechanisms of substrate targeting by protein kinases. *Trends Biochem. Sci.* **43**:380.

Taylor, S. S., and A. Kornev. 2011. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **36**:65–77.

Structure and Mechanism of G Protein–Coupled Receptors

Benovic, J. 2012. G-protein-coupled receptors signal victory. *Cell* **151**:1–3. (A review of the research leading to the 2012 Nobel Prize.)

Komolov, K., and Benovic, J. 2018. G protein-coupled receptor kinases: past, present and future. *Cell. Signal.* **41**:17.

Ring, A., et al. 2013. Adrenaline-activated structure of β 2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**:575.

Siram, K., and P. Insel. 2018. G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs? *Mol. Pharmacol.* **93**:251.

Smith, J. S., and S. Rajagopal. 2016. The β-arrestins: multifunctional regulators of G protein-coupled receptors. *J. Biol. Chem.* **291**:8969.

Tesmer, J. 2016. Hitchhiking on the heptahelical highway: structure and function of 7TM receptor complexes. *Nat. Rev. Mol. Cell Biol.* **17**: 439.

Weis, W., and B. Kobilka. 2018. The molecular basis of G protein–coupled receptor activation. *Annu. Rev. Biochem.* **87**:897–919.

Regulating Metabolism of Many Cells: G Protein–Coupled Receptors that Activate or Inhibit Adenylyl Cyclase

Agius, L. 2010. Physiological control of liver glycogen metabolism: lessons from novel glycogen phosphorylase inhibitors. *Mini-Rev. Med. Chem.* **10**:1175–1187.

DeWire, S., et al. 2007. β-Arrestins and cell signaling. *Annu. Rev. Physiol.* **69**:483–510.

Lefkowitz, R. J., and S. K. Shenoy. 2005. Transduction of receptor signals by β-arrestins. *Science* **308**:512–517.

Scott, J. D., et al., 2013. Creating order from chaos: cellular regulation by kinase anchoring. *Annu. Rev. Pharmacol.* **53**:187–210.

Taylor, S. S., et al. 2012. Assembly of allosteric macromolecular switches: lessons from PKA. *Nat. Rev. Mol. Cell Biol.* **13**:646–658.

Regulating Protein Secretion and Muscle Contraction: Ca^{2+} Ions as Second Messengers in Multiple Signal Transduction Pathways

DeStefani, D., et al. 2016. Enjoy the trip: calcium in mitochondria back and forth. *Annu. Rev. Biochem.* **85**:161.

Kaufman, R., and J. Amphora. 2014. Calcium trafficking integrates endoplasmic reticulum function with mitochondrial bioenergetics. *Biochim. Biophys. Acta* **1843**:2233–2239.

Nguyen, N., et al. 2018. Store-operated calcium entry mediated by ORAI and STIM. *Compr. Physiol.* **8**:981–1002. doi: 10.1002/cphy.c170031

Soboloff, J., et al. 2012. STIM proteins: dynamic calcium signal transducers. *Nat. Rev. Mol. Cell Biol.* **13**:549–565.

Vision: How the Eye Senses Light

Field, G. D., and A. P. Sampath. 2017. Behavioural and physiological limits to vision in mammals. *Phil. Trans. R. Soc. B* **372**:20160072. doi: 10.1098/rstb.2016.0072

Hofmann, K. P., et al. 2009. A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci.* **34**:540–552.

Pearring, J. N. 2013. Protein sorting, targeting and trafficking in photoreceptor cells. *Prog. Retin. Eye Res.* **36**:24–61.

Smith, S. O. 2010. Structure and activation of the visual pigment rhodopsin. *Annu. Rev. Biophys.* **39**:309–328.

Tsien, A., et al. 2018. Visual cycle proteins: structure, function, and roles in human retinal disease. *J. Biol. Chem.* **293**:13016–13021.

Yue, W., et al. 2019. Elementary response triggered by transducin in retinal rods. *PNAS* **116**: 5144–5153.

Chapter 16

Growth Factors and Their Receptor Tyrosine Kinases

Gutmann, T., et al. 2018. Visualization of ligand-induced transmembrane signaling in the full-length human insulin receptor. *J. Cell. Biol.* **217**(5):1643–1649. doi: 10.1083/jcb.201711047

Haeusler, R., et al. 2017. Biochemical and cellular properties of insulin receptor signalling. *Nat. Rev. Mol Cell Biol.* **19**(1):31–44. doi: 10.1038/nrm.2017.89

Kovacs, E., et al. 2015. A structural perspective on the regulation of the epidermal growth factor receptor. *Annu. Rev. Biochem.* **84**:739.

Lemmon, M., and J. Schlessinger. 2010. Cell signaling by receptor tyrosine kinases. *Cell* **141:**1117–1134.

Tomas, A., et al. 2014. EGF receptor trafficking: consequences for signaling and cancer. *Trends Cell Biol.* **24**:26–34.

The Ras/MAP Kinase Signal Transduction Pathway

Atay, Q., and J. Skotheim. 2017. Spatial and temporal signal processing and decision making by MAPK pathways. *J. Cell. Biol.* **216**(2):317–330. doi: 10.1083/jcb.201609124

Simanshu, D., et al. 2017. RAS proteins and their regulators in human disease. *Cell* **170**:17.

Phosphoinositide Signal Transduction Pathways

Bilanges, B., Y. Posor, and B. Vanhaesebroeck. 2019. PI3K isoforms in cell signaling and vesicle trafficking. *Nat. Rev. Mol. Cell Biol.* **20**:515.

Fruman, D., et al. 2017. The PI3K pathway in human disease. *Cell* **170**:607.

Manning, B., and A. Toker. 2017. AKT/PKB signaling: navigating the network. *Cell* **169**:381.

Cytokines, Cytokine Receptors, and the JAK/STAT Signaling Pathway

Hattangadi, S., et al. 2011. From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs and chromatin modifications. *Blood* **118**:6258–6268.

Villarino, A., et al. 2017. Mechanisms and consequences of Jak–STAT signaling in the immune system. *Nat. Immunol.* **4**:374.

The TGF-β Family of Growth Factors, Their Receptor Serine Kinases, and the Smad Transcription Factors They Activate

Chaikaud, A., and A. Bullock. 2016. Structural basis of intracellular TGF-β signaling: receptors and Smads. *Cold Spring Harb. Perspect. Biol.* **8**:a022111.

Deheuninck, J., and K. Luo. 2009. Ski and SnoN, potent negative regulators of TGF-β signalling. *Cell Res.* **19**:47–57.

Hata, A., and Y.-G. Chen. 2016. TGF-β signaling from receptors to Smads. *Cold Spring Harb. Perspect. Biol.* **8**:a022061.

Hinck, A., et al. 2016. Structural biology and evolution of the TGF-β family. *Cold Spring Harb. Perspect. Biol*. **8**:a022103.

Massagué, J. 2012. TGFβ signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**:617–630.

Signal Transduction Pathways that Utilize Regulated, Site-Specific Protein Cleavage: Notch/Delta and EGF Precursors

Bray, S. 2016. Notch signalling in context. *Nat. Rev. Mol. Cell Biol.* **17**:773.

Kovall, R., et al. 2017. The canonical Notch signaling pathway: structural and biochemical insights into shape, sugar, and force. *Dev. Cell* **41**:228.

Musse, A., et al. 2012. Notch ligand endocytosis: mechanistic basis of signaling activity. *Semin. Cell Dev. Biol.* **23**:429–436.

Yang, G., et al. 2019. Structural basis of Notch recognition by human γ-secretase. *Nature* **565**:192.

Signal Transduction Pathways that Utilize Proteasomal Degradation of Signaling Components: Wnt, Hedgehog, and the Many Hormones that Activate NF-κB

Ferrao, R., et al. 2012. Structural insights into the assembly of large oligomeric signalosomes in the Toll-like receptor–Interleukin-1 receptor superfamily. *Sci. Signal.* **5**:re3.

Gammons, M., and M. Bienz. 2018. Multiprotein complexes governing Wnt signal transduction. *Curr. Opin. Cell Biol.* **51**:42.

He, M., et al. 2017. Microtubule motors drive Hedgehog signaling in primary cilia. *Trends Cell Biol.* **27**:110. doi: 10.1016/j.tcb.2016.09.010

Iwai, K., et al. 2014. Linear ubiquitin chains: NF-κB signalling, cell death and beyond. *Nat. Rev. Mol. Cell Biol.* **15**:503–508.

Kong, J., et al. 2019. Biochemical mechanisms of vertebrate hedgehog signaling. *Development* **146**:dev166892. doi: 10.1242/dev.166892

Kowatsch, C., et al. 2019. Structures of vertebrate Patched and Smoothened reveal intimate links between cholesterol and Hedgehog signaling. *Curr. Opin. Struc. Biol.* **57**:204.

Lee, R., et al. 2016. Hedgehog signalling. *Development* **143**:367. doi: 10.1242/dev.120154

Napetschnig, J., and H. Wu. 2013. Molecular basis of NF-κB signaling. *Annu. Rev. Biophys.* **42**:443.

Nusse, R., and H. Clevers. 2017. Wnt/β-Catenin signaling, disease, and emerging therapeutic modalities. *Cell* **169**:985.

Wiese, K., et al. 2018. Wnt signalling: conquering complexity. Dev. **145**:dev165902. doi: 10.1242/dev.165902

Yau, R., and M. Rape. 2016. The increasing complexity of the ubiquitin code. *Nat. Cell Biol.* **18**:579.

Chapter 17

Microfilaments and Actin Structures

Holmes, K. C., et al. 1990. Atomic model of the actin filament. *Nature* **347**:44–49.

Kabsch, W., et al. 1990. Atomic structure of the actin: DNase I complex. *Nature* **347**:37–44.

Pollard, T. D., and J. A. Cooper. 2009. Actin, a central player in cell shape and movement. *Science* **326**:1208–1212.

Dynamics of Actin Filaments

Paavilainen, V. O., et al. 2004. Regulation of cytoskeletal dynamics by actin-monomer-binding proteins. *Trends Cell Biol.* **14**:386–394.

Theriot, J. A. 1997. Accelerating on a treadmill: ADF/cofilin promotes rapid actin filament turnover in the dynamic cytoskeleton. *J. Cell Biol.* **136**:1165–1168.

Mechanisms of Actin Filament Assembly

Brieher, W. 2013. Mechanisms of actin disassembly. *Mol. Biol. Cell* **24**:2299–2302.

Campellone, K. G., and M. D. Welch. 2010. A nucleator arms race: cellular control of actin assembly. *Nat. Rev. Mol. Cell Biol.* **11**:237–251.

Chen, B., et al. 2014. The WAVE regulatory complex links diverse receptors to the actin cytoskeleton. *Cell* **156**:195–207.

Gouin, E., M. D. Welch, and P. Cossart. 2005. Actin-based motility of intracellular pathogens. *Curr. Opin. Microbiol.* **8**:35–45.

Rotty, J. D., C. Wu, and J. E. Bear. 2013. New insights into the regulation and cellular functions of the Arp2/3 complex. *Nat. Rev. Mol. Cell Biol.* **14**:7–12.

Organization of Actin-Based Cellular Structures

Bennett, V., and D. N. Lorenzo. 2013. Spectrin- and ankyrin-based membrane domains and the evolution of vertebrates. *Curr. Top. Memb.* **72**:1–37.

Fehon, R. G., A. I. McClatchey, and A. Bretscher. 2010. Organizing the cell cortex: the role of ERM proteins. *Nat. Rev. Mol. Cell Biol.* **11**:276–287.

Myosins: Actin-Based Motor Proteins

Berg, J. S., B. C. Powell, and R. E. Cheney. 2001. A millennial myosin census. *Mol. Biol. Cell* **12**:780–794.

Vale, R. D. 2003. The molecular motor toolbox for intracellular transport. *Cell* **112**:467–480.

Vale, R. D., and R. A. Milligan. 2000. The way things move: looking under the hood of molecular motor proteins. *Science* **288**:88–95.

Myosin-Powered Movements

Clark, K. A., et al. 2002. Striated muscle cytoarchitecture: an intricate web of form and function. *Annu. Rev. Cell Dev. Bi.* **18**:637–706.

Hammer, J. A. III, and J. R. Sellers. 2012. Walking to work: roles for class V myosins as cargo transporters. *Nat. Rev. Mol. Cell Biol.* **13**:13–26.

Cell Migration: Mechanism, Signaling, and Chemotaxis

Artemenko, Y., T. J. Lampert, and P. N. Devreotes. 2014. Moving towards a paradigm: common mechanisms of chemotactic signaling in *Dictyostelium* and mammalian leukocytes. *Cell Mol. Life Sci.* **71**:3711–3747.

Etienne-Manneville, S. 2004. Cdc42—the centre of polarity. *J. Cell Sci.* **117**:1291–1300.

Etienne-Manneville, S., and A. Hall. 2002. Rho GTPases in cell biology. *Nature* **420**:629–635.

Pollard, T. D., and G. G. Borisy. 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112**:453–465.

Ridley, A. J., et al. 2003. Cell migration: integrating signals from the front to back. *Science* **302**:1704–1709.

Chapter 18

Microtubule Structure and Organization

Badano, J. L., T. M. Teslovich, and N. Katsanis. 2005. The centrosome in human genetic disease. *Nat. Rev. Mol. Cell Biol.* **6**:194–205.

Brouhard, G. J., and L. M. Rice. 2018. Microtubule dynamics: an interplay of biochemistry and mechanics. *Nat. Rev. Mol. Cell Biol.* **19**:451–463.

Dutcher, S. K. 2001. The tubulin fraternity: alpha to eta. *Curr. Opin. Cell Biol.* **13**:49–54.

Nogales, E., and H.-W. Wang. 2006. Structural intermediates in microtubule assembly and disassembly: how and why? *Curr. Opin. Cell Biol.* **18**:179–184.

Microtubule Dynamics

Brouhard, G. J., and L. M. Rice. 2014. The contribution of αβ-tubulin curvature to microtubule dynamics. *J. Cell Biol.* **207**:323–334.

Desai, A., and T. J. Mitchison. 1997. Microtubule polymerization dynamics. *Annu. Rev. Cell Dev. Bi.* **13**:83–117.

Regulation of Microtubule Structure and Dynamics

Galjart, N. 2010. Plus-end-tracking proteins and their interactions at microtubule ends. *Curr. Biol.* **20**:R528–R537.

Hammond, J. W., D. Cai, and K. J. Verhey. 2008. Tubulin modifications and their cellular functions. *Curr. Opin. Cell Biol.* **20**:71–76.

McNally, F. J., and A. Roll-Mecak. 2018. Microtubule-severing enzymes: from cellular functions to molecular mechanism. *J. Cell Biol.* **217**:4057–4069.

Wloga, D., and J. Gaertig. 2010. Post-translational modifications of microtubules. *J. Cell Sci.* **123**:3447–3455.

Zheng, Y., and P. A. Iglesias. 2013. Nucleating new branches from old. *Cell* **152**:669–670.

Kinesins and Dyneins: Microtubule-Based Motor Proteins

Allan, V. 2014. One, two, three, cytoplasmic dynein is go! *Nature* **345**:271–272.

Bhabha, G., et al. 2014. Allosteric communication in the dynein motor domain. *Cell* **159**:857–868.

Endow, S. A., F. J. Kull, and H. Liu. 2010. Kinesins at a glance. *J. Cell Sci.* **123**:3420–3424.

Fu, M., and E. L. F. Holzbaur. 2014. Integrated regulation of motor-driven organelle transport by scaffolding proteins. *Trends Cell Biol.* **24**:564–574.

Hirokawa, N., et al. 2009. Kinesin superfamily motor proteins and intracellular transport. *Nat. Rev. Mol. Cell Biol.* **10**:682–696.

Reck-Peterson, S., et al. 2018. The cytoplasmic dynein transport machinery and its many cargoes. *Nat. Rev. Mol. Cell Biol.* **19**:382–398.

Trokter, M., and T. Surrey. 2012. Lis1 clamps dynein to the microtubule. *Cell* **150**:877–879.

Vale, R. D. 2003. The molecular motor toolbox for intracellular transport. *Cell* **112**:467–480.

Vale, R. D., and R. A. Milligan. 2000. The way things move: looking under the hood of molecular motor proteins. *Science* **288**:88–95.

Verhey, K. J., and J. W. Hammond. 2009. Traffic control: regulation of kinesin motors. *Nat. Rev. Mol. Cell Biol.* **10**:765–777.

Cilia and Flagella: Microtubule-Based Surface Structures

Gerdes, J. M., E. E. Davis, and N. Katsanis. 2009. The vertebrate primary cilium in development, homeostasis, and disease. *Cell* **137**:32–45.

Jin, H., et al. 2010. The conserved Bardet-Biedl syndrome proteins assemble a coat that traffics membrane proteins to cilia. *Cell* **141**:1208–1218.

Mitosis

Basant, A., and M. Glotzer. 2018. Spatiotemporal regulation of RhoA during cytokinesis. *Curr. Biol*. **28**:R570–R580.

Foley, E. A., and T. M. Kapoor. 2013. Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. *Nat. Rev. Cell Mol. Biol.* **14**:25–37.

Goshima, G., et al. 2009. Augmin: a protein complex required for centrosome-independent microtubule generation within the spindle. *J. Cell Biol.* **181**:421–429.

Hinshaw, S. M., and S. C. Harrison. 2018. Kinetochore function from the bottom up. *Trends Cell Biol.* **28:**22–33.

Kiyomitsu, T. 2015. Mechanisms of daughter cell-size control during cell division. *Trends Cell Biol.* **25**:286–295.

Lampert, F., and S. Westermann. 2011. A blueprint for kinetochores—new insights into the molecular mechanics of cell division. *Nat. Rev. Mol. Cell Biol.* **12**:407–412.

London, N., and S. Biggins. 2014. Signalling dynamics in the spindle checkpoint response. *Nat. Rev. Mol. Cell Biol.* **15**:735–747.

Mitchison, T. J., and E. D. Salmon. 2001. Mitosis: a history of division. *Nat. Cell Biol.* **3**:E17–E21.

Musacchio, A. 2015. The molecular biology of spindle assembly checkpoint signaling dynamics. *Curr. Biol.* **25**:R1002–R1018.

Prosser, S. L., and L. Pelletier. 2017. Mitotic spindle assembly in animal cells: a fine balancing act. *Nat. Rev. Mol. Cell Biol.* **18**:187–201.

Intermediate Filaments

Burke, B., and C. L. Stewart. 2013. The nuclear lamins: flexibility in function. *Nat. Rev. Mol. Cell Biol.* **14**:13–23.

de Leeuw, R., Y. Gruenbaum, and O. Medalia. 2018. Nuclear lamins: thin filaments with major functions. *Trends Cell Biol.* **28**:34–45.

Isermann, P., and J. Lammerding. 2013. Nuclear mechanics and mechanotransduction in health and disease. *Curr. Biol.* **23**:R1113–R1121.

Luxton, G. W. G, and D. A. Starr. 2014. KASHing up with the nucleus: novel functional roles of KASH proteins at the cytoplasmic surface of the nucleus. *Curr. Opin. Cell Biol.* **28**:69–75.

Coordination and Cooperation Between Cytoskeletal Elements

Dogterom, M., and G. H. Koenderink. 2019. Actin-microtubule crosstalk in cell biology. *Nat. Rev. Mol. Cell Biol.* **20**:38–54.

Etienne-Manneville, S., et al. 2005. Cdc42 and Par6-PKCz regulate the spatially localized association of Dlg1 and APC to control cell polarization. *J. Cell Biol.* **170**:895–901.

Kodama, A., T. Lechler, and E. Fuchs. 2004. Coordinating cytoskeletal tracks to polarize cellular movements. *J. Cell Biol.* **167**:203–207.

Schaefer, A. W., et al. 2008. Coordination of actin filament and microtubule dynamics during neurite outgrowth. *Dev. Cell* **15**:146–162.

Chapter 19

References

Overview of the Cell Cycle

Morgan, D. O. 2006. *The Cell Cycle: Principles of Control*. New Science Press.

Blagosklonny, M. V., and A. B. Pardee. 2002. The restriction point of the cell cycle. *Cell Cycle* **1**(2):103–110.

Duronio, R. J., and Y. Xiong. 2013. Signaling pathways that control cell proliferation. *Cold Spring Harb. Perspect. Biol.* **5**:a008904.

Nurse, P. M. 2002. Nobel Lecture. Cyclin dependent kinases and cell cycle control. *Biosci. Rep.* **22**(5-6):487–499.

Reinhardt, H. C., and M. B. Yaffe. 2013. Phospho-Ser/Thr-binding domains: navigating the cell cycle and DNA damage response. *Nat. Rev. Mol. Cell Biol.* **14**:563–580.

Cell Cycle Progression and Control: Feedback Loops and Post-Translational Modification

Besson, A., S. F. Dowdy, and J. M. Roberts. 2008. CDK inhibitors: cell cycle regulators and beyond. *Dev. Cell* **14**(2):159–169.

Bloom, J., and F. R. Cross. 2007. Multiple levels of cyclin specificity in cell-cycle control. *Nat. Rev. Mol. Cell Biol.* **8**(2):149–160.

Ferrell, J. E. Jr., T. Y. Tsai, and Q. Yang. 2011. Modeling the cell cycle: why do certain circuits oscillate? *Cell* **144**(6):874–885.

Jackson, P. K. 2008. The hunt for cyclin. *Cell* 134:199–202.

Lim, S., and P. Kaldis. 2013. CDKs, cyclins and CKIs: roles beyond cell cycle regulation. *Development* **140**(15):3079–3093.

Malumbres, M. 2014. Cyclin-dependent kinases. *Genome Biol.* **15**(6):122.

The Transition from G_1 into S Phase and DNA Replication

Costa, A., I. V. Hood, and J. M. Berger. 2013. Mechanisms for initiating cellular DNA replication. *Ann. Rev. Biochem.* **82**:25–54.

Li, Y., and H. Araki. 2013. Loading and activation of DNA replicative helicases: the key step of initiation of DNA replication. *Genes Cells* **18**(4):266–277.

Losada, A. 2014. Cohesin in cancer: chromosome segregation and beyond. *Nat. Rev. Cancer* **14**(6):389–393.

Sclafani, R. A., and T. M. Holzen. 2007. Cell cycle regulation of DNA replication. *Ann. Rev. Genet*. 41:237–280.

Skibbens, R. V. 2019. Condensins and cohesins – one of these things is not like the other! *J. Cell Sci.* **132**(3):1–11.

Wood, A. J., A. F. Severson, and B. J. Meyer. 2010. Condensin and cohesin complexity: the expanding repertoire of functions. *Nat. Rev. Genet*. **11**(6):391–404.

The G_2/M Transition and the Irreversible Engine of Mitosis

Conduit, P. T., A. Wainman, and J. W. Raff. 2015. Centrosome function and assembly in animal cells. *Nat. Rev. Mol. Cell Biol.* **16**(10):611–624.

DeLuca, J. G., and A. Musacchio. 2012. Structural organization of the kinetochore-microtubule interface. *Curr. Opin. Cell Biol*. **24**(1):48–56.

Ohta, S., et al. 2011. Building mitotic chromosomes. *Curr. Opin. Cell Biol*. **23**(1):114–121.

Smoyer, C. J., and S. L. Jaspersen. 2014. Breaking down the wall: the nuclear envelope during mitosis. *Curr. Opin. Cell Biol.* **26**:1–9.

The Mitotic Spindle, Chromosome Segregation, and Exit from Mitosis

Alfieri, C., S. Zhang,, and D. Barford. 2017. Visualizing the complex functions and mechanisms of the anaphase promoting complex/cyclosome (APC/C). *Open Biol.* **7**:170204.

Basant, A., and M. Glotzer. 2018. Spatiotemporal regulation of RhoA during cytokinesis. *Curr. Biol.* **28**:R570–580.

Conduit, P. T., A. Wainman, and J. W. Raff. 2015. Centrosome function and assembly in animal cells. *Nat. Rev. Mol. Cell. Biol.* **16**(10):611–624.

Craney, A., and M. Rape. 2013. Dynamic regulation of ubiquitin-dependent cell cycle control. *Curr. Opin. Cell Biol.* **25**(6):704–710.

Hetzer, M. W., T. C. Walther, and I. W. Mattaj. 2005. Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Ann. Rev. Cell Dev. Biol.* **21**:347–380.

Sullivan, M., and D. O. Morgan. 2007. Finishing mitosis, one step at a time. *Nat. Rev. Mol. Cell Biol.* **8**(11):894–903.

Surveillance Mechanisms in Cell Cycle Regulation

Aguilera, A., and T. García-Muse. 2013. Causes of genome instability. *Ann. Rev. Genet*. **47**:1–32.

Ciccia, A., and S. J. Elledge. 2010. The DNA damage response: making it safe to play with knives. *Mol. Cell* **40**(2):179–204.

Harper, J. W., and S. J. Elledge. 2007. The DNA damage response: ten years after. *Mol. Cell* **28**(5):739–745.

Musacchio, A. 2015. The molecular biology of spindle assembly checkpoint signaling dynamics. *Curr. Biol* **25**(20):R1002–R1018.

Shiloh, Y., and Y. Ziv. 2013. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat. Rev. Mol. Cell Biol.* **14**(4):197–210.

Meiosis: A Special Type of Cell Division

Lesch, B. J., and D. C. Page. 2012. Genetics of germ cell development. *Nat. Rev. Genet*. **13**(11):781–794.

Lichten, M., and B. de Massy. 2011. The impressionistic landscape of meiotic recombination. *Cell* **147**(2):267–270.

Miller, M. P., A. Amon, and E. Ünal. 2013. Meiosis I: when chromosomes undergo extreme makeover. *Curr. Opin. Cell Biol*. **25**(6):687–696.

Zickler, D., and N. Kleckner. 1999. Meiotic chromosomes: integrating structure and function. *Ann. Rev. Genet.* **33**:603–754.

Chapter 20

Cell-Cell and Cell–Extracellular Matrix Adhesion: An Overview

Hannezo, E., and C. P. Heisenberg. 2019. Mechanochemical feedback loops in development and disease. *Cell* **178**(1):12–25.

Humphrey, J. D., E. R. Dufresne, and M. A. Schwartz. 2014. Mechanotransduction and extracellular matrix homeostasis. *Nat. Rev. Mol. Cell Biol.* **15**:802–812.

Jansen, K. A., et al. 2015. A guide to mechanobiology: where biology and physics meet. *Biochim. Biophys. Acta* **1853**:3043–3052.

The Matrisome Project website (http://matrisomeproject.mit.edu). A compilation of datasets and information about the genes and proteins of the matrisome.

Naba, A., et al. 2016. The extracellular matrix: tools and insights for the "omics" era. *Matrix Biol.* **49**:10–24.

Nieto, M. A. 2013. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science* **342**:1234850.

Padmanabhan, A., et al. 2015. Jack of all trades: functional modularity in the adherens junction. *Curr. Opin. Cell Biol.* **36**:32–40.

Wolfenson, H., B. Yang, and M. P. Sheetz. 2019. Steps in mechanotransduction pathways that control cell morphology. *Annu. Rev. Physiol.* **81**:585–605.

Cell-Cell and Cell–Extracellular Matrix Junctions and Their Adhesion Molecules

Anderson, J. M., and C. M. Van Itallie. 2009. Physiology and function of the tight junction. *Cold Spring Harb. Perspect. Biol.* **1**(2):a002584.

Brasch, J., et al. 2019. Visualization of clustered protocadherin neuronal self-recognition complexes. *Nature* **569**(7755):280–283.

Bush, M., et al. 2019. An ensemble of flexible conformations underlies mechanotransduction by the cadherin-catenin adhesion complex. *Proc. Nat'l. Acad. Sci. USA* **116**(43):21545–21555.

Charras. G., and A. S. Yap. 2018. Tensile forces and mechanotransduction at cell-cell junctions. *Curr. Biol.* **28**(8):R445–R457.

Conrad, M. P., et al. 2016. Molecular basis of claudin-17 anion selectivity. *Cell. Mol. Life Sci.* **73**:185–200.

Gershon, E., V. Plaks, and N. Dekel. 2008. Gap junctions in the ovary: expression, localization and function. *Mol. Cell. Endocrinol.* **282**:18–25.

Glentis, A., V. Gurchenkov, and D. Matic Vignjevic. 2014. Assembly, heterogeneity, and breaching of the basement membranes. *Cell Adh. Migr.* **8**(3):236–245.

Goodenough, D. 1999. *Proc. Nat'l. Acad. Sci. USA* **96**:319.

Lee, J. M., et al. 2006. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *J. Cell Biol.* **172**(7):973–981.

McMillen, P., and S. A. Holley. 2015. Integration of cell-cell and cell-ECM adhesion in vertebrate morphogenesis. *Curr. Opin. Cell Biol.* **36**:48–53.

Mountoufaris, G., et al. 2018. Writing, reading, and translating the clustered protocadherin cell surface recognition code for neural circuit assembly. *Annu. Rev. Cell Dev. Biol.* **34**:471–493.

Nakagawa, S., S. Maeda, and T. Tsukihara. 2010. *Curr. Opin. Struct. Biol.* **20**(4):423–430.

Oda, H., and M. Takeichi. 2011. Structural and functional diversity of cadherin at the adherens junction. *J. Cell Biol.* **193**(7):1137–1146.

Tsukita, S. et al. 2001. *Nat. Rev. Mol. Cell Biol.* **2**:285.

Walko, G., M. J. Castañón, and G. Wiche. 2015. Molecular architecture and function of the hemidesmosome. *Cell Tissue Res.* **360**:529–544.

Wu, Y., P. Kanchanawong, and R. Zaidel-Bar. 2015. Actin-delimited adhesion-independent clustering of E-cadherin forms the nanoscale building blocks of adherens junctions. *Dev. Cell* **32**:139–154.

Yang, C. C., et al. 2015. Differential regulation of the Hippo pathway by adherens junctions and apical-basal cell polarity modules. *Proc. Nat'l. Acad. Sci. USA* **112**:1785–1790.

Zaidel-Bar, R., and B. Geiger. 2010. The switchable integrin adhesome. *J. Cell Sci.* **123**(pt. 9):1385–1388.

Zihni, C., et al. 2016. Tight junctions: from simple barriers to multifunctional molecular gates. *Nat. Rev. Mol. Cell Biol.* **17**(9):564–580.

The Extracellular Matrix I: The Basal Lamina

Bonnans, C., J. Chou, and Z. Werb. 2014. Remodelling the extracellular matrix in development and disease. *Nat. Rev. Mol. Cell Biol.* **15**:786–801.

Brown, K.L., et al. 2017. Building collagen IV smart scaffolds on the outside of cells. *Protein Sci.* **26**(11):2151–2161.

Hohenester, E., and P. D. Yurchenco. 2013. Laminins in basement membrane assembly. *Cell Adh. Migr.* **7**:56–63.

Hynes, R. O. 2014. Stretching the boundaries of extracellular matrix research. *Nat. Rev. Mol. Cell Biol.* **15**:761–763.

Robertson, W. E., et al. 2014. Supramolecular organization of the α121-α565 collagen IV network. *J. Biol. Chem.* **289**:25601–25610.

Sarrazin, S., W. C. Lamanna, and J. D. Esko. 2011. Heparan sulfate proteoglycans. *Cold Spring Harb. Perspect. Biol.* **3**. pii: a004952.

The Extracellular Matrix II: Connective Tissue

Campbell, M. G., et al. 2020. Cryo-EM reveals integrin-mediated TGF-β activation without release from latent TGF-β. *Cell* **180**(3):490–501.

Canty, E. G., and K. E. Kadler. 2005. Procollagen trafficking, processing and fibrillogenesis. *J. Cell Sci.* **118**:1341–1353.

Endo, T. 2019. Mammalian O-mannosyl glycans: biochemistry and glycopathology. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **95**(1):39–51.

Lodyga, M., and B. Hinz. 2020. TGF-β1 — A truly transforming growth factor in fibrosis and immunity. *Semin. Cell Dev. Biol.* **101**:123–139.

Robertson, I. B., et al. 2015. Latent TGF-β-binding proteins. *Matrix Biol.* **47**:44–53.

Shaw, L. M., and B. R. Olsen. 1991. FACIT collagens: diverse molecular bridges in extracellular matrices. *Trends Biochem. Sci.* **16**(5):191–194.

Shoulders, M. D., and R. T. Raines. 2011. Interstrand dipole-dipole interactions can stabilize the collagen triple helix. *J. Biol. Chem.* **286**:22905–22912.

Yoshida-Moriguchi, T., and K. P. Campbell. 2015. Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* **25**:702–713.

Adhesive Interactions in Motile and Nonmotile Cells

Carraher, C. L., and J. E. Schwarzbauer. 2013. Regulation of matrix assembly through rigidity-dependent fibronectin conformational changes. *J. Biol. Chem.* **288**:14805–14814.

Clevers, H. 2016. Modeling development and disease with organoids. *Cell* **165**(7):1586–1597.

Collins, C., and W. J. Nelson. 2015. Running with neighbors: coordinating cell migration and cell-cell adhesion. *Curr. Opin. Cell. Biol.* **36**:62–70.

Früh, S. M., et al. 2015. Molecular architecture of native fibronectin fibrils. *Nat. Commun.* **6**:7275.

Griffith, L. G., and M. A. Swartz. 2006. Capturing complex 3D tissue physiology in vitro. *Nat. Rev. Mol. Cell Biol.* **7**(3):211–224.

Iwamoto, D. V., and D. A. Calderwood. 2015. Regulation of integrin-mediated adhesions. *Curr. Opin. Cell Biol.* **36**:41–47.

Jensen, C., and Y. Teng. 2020. Is it time to start transitioning from 2D to 3D cell culture? *Front. Mol. Biosci.* **7**(Mar 6):33.

Kadry, Y. A., and D. A. Calderwood. 2020 Chapter 22: Structural and signaling functions of integrins. *Biochim. Biophys. Acta Biomembr.* **1862**(5):183206.

Kechagia, J. Z., J. Ivaska, and P. Roca-Cusachs. 2019. Integrins as biomechanical sensors of the microenvironment. *Nat. Rev. Mol. Cell Biol.* **20**(8):457–473.

Sheikh, M. O., et al. 2020. HNK-1 Sulfotransferase modulates α-dystroglycan glycosylation by 3-O-sulfation of glucuronic acid on matriglycan. *Glycobiology* Mar 9. pii: cwaa024. doi:

10.1093/glycob/cwaa024. [Epub ahead of print.]

Nourshargh, S., and R. Alon. 2014. Leukocyte migration into inflamed tissues. *Immunity* **41**:694–707.

Springer, T. A., and M. L. Dustin. 2012. Integrin inside-out signaling and the immunological synapse. *Curr. Opin. Cell Biol.* **24**:107–115.

Wolfenson, H., B. Yang, and M. P. Sheetz. 2019 Steps in mechanotransduction pathways that control cell morphology. *Annu. Rev. Physiol.* **81**:585–605.

Xiong, J. P., et al. 2001. Crystal structure of the extracellular segment of integrin αVβ3. *Science* **294**:339–345.

Plant Tissues

Austefjord, M. W., H. H. Gerdes, and X. Wang. 2014. Tunneling nanotubes: diversity in morphology and structure. *Commun. Integr. Biol.* **7**:e27934.

Chae, K., and E. M. Lord. 2011. Pollen tube growth and guidance: roles of small, secreted proteins. *Ann. Bot.* **108**:627–636.

Darwin, Charles. *Insectivorous Plants*. New York: D. Appleton and Company, 1875.

Darwin, Charles. 1860 letter to Charles Lyell, Nov. 24, 1860 *Darwin Correspondence Project*, https://www.darwinproject.ac.uk/letter/DCP-LETT-2996.xml

Fruleux, A., S. Verger, and A. Boudaoud. 2019. Feeling stressed or strained? A biophysical model for cell wall mechanosensing in plants. *Front. Plant Sci*. **10**:757. doi: 10.3389/fpls.2019.00757. eCollection 2019.

Galletti, R., et al. 2016 Developing a "thick skin": a paradoxical role for mechanical tension in maintaining epidermal integrity? *Development* **143**(18):3249–3258.

Petit, J. D., et al. 2020. Dare to change, the dynamics behind plasmodesmata-mediated cell-to-cell communication. *Curr. Opin. Plant Biol.* **53**:80–89.

Sager, R. E., and J. Y. Lee. 2018. Plasmodesmata at a glance. *J. Cell Sci.* **131**(11). pii: jcs209346. doi: 10.1242/jcs.209346

Sevilem, I., S. R. Yadav, and Y. Helariutta. 2015. Plasmodesmata: channels for intercellular signaling during plant growth and development. *Meth. Mol. Biol.* **1217**:3–24.

Tan, A. S., et al. 2015. Mitochondrial genome acquisition restores respiratory function and tumorigenic potential of cancer cells without mitochondrial DNA. *Cell Metab.* **21**:81–94.

Uluisik, S., et al. 2016. Genetic improvement of tomato by targeted control of fruit softening. *Nat. Biotechnol.* **34**(9):950–952.

Chapter 21

Regulating Blood Glucose Level

Bogan, J. 2012. Regulation of glucose transporter translocation in health and diabetes. *Ann. Rev. Biochem.* **81**:507–532.

Kip, A., T. E. McGraw, and D. James. 2019. Thirty sweet years of GLUT4. *J. Biol. Chem.* **294**:11369–11381.

Petersen, M., D. Vatner, and G. Shulman. 2017. Regulation of hepatic glucose metabolism in health and disease. *Nat. Rev. Endocrinol.* **13**:572–587.

Integrating Cell Growth Signals with Nutrient and Energy Levels

Anandapadamanaban, M., et al. 2019. Architecture of human Rag GTPase heterodimers and their complex with mTORC1. *Science* **366**:203–210.

Henske, E., et al. 2016. Tuberous sclerosis complex. *Nat. Rev. Dis. Primers* **2**:1–18.

Liu, G., and D. Sabatini. 2020. mTOR at the nexus of nutrition, growth, ageing and disease. *Nat. Rev. Mol. Cell. Biol.* **21**:183–203.

Saxton, R., and D. Sabatini. 2017. mTOR signaling in growth, metabolism, and disease. *Cell* **168**:960–976.

Valvezan, A., B. Manning. 2019. Molecular logic of mTORC1 signalling as a metabolic rheostat. *Nat. Metabolism* **1**: 321–333.

Responding to Changes in the Levels of Cholesterol and Unsaturated Fatty Acids

Brown, M., and J. Goldstein. 2009. Cholesterol feedback: from Schoenheimer's bottle to Scap's MELADL. *J. Lipid. Res.* **50**:S15–S27.

Brown, M., A. Radhakrishnan, and J. Goldstein. 2018. Retrospective on cholesterol homeostasis: The central role of Scap. *Annu. Rev. Biochem.* **87**:783–807.

DeBose-Boyd, R., and J. Ye. 2018. **SREBPs in lipid metabolism, insulin signaling, and beyond.**

Trends Biochem. Sci. **43**:358–368.

Responding to Low Oxygen

Ivan, M., and W. Kaelin. 2017. The EGLN-HIF O2-sensing system: multiple inputs and feedbacks. *Mol. Cell* **66**:772–779.

Kaelin, W., and P. Ratcliffe. 2008. **Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol. Cell.*** 30:393–402.

Masson, N., et al. 2019. Conserved N-terminal cysteine dioxygenases transduce responses to hypoxia in animals and plants. *Science* **365**:65–69.

Thompson, C. 2016. Into thin air: how we sense and respond to hypoxia. *Cell* **167**:9–11.

Responding to Elevated Temperatures

Gomez-Pastor, R., E. T. Burchfiel, and D. J. Thiele. 2018. Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **19**(1):4–19.

Krakowiak, J., et al. 2018. Hsf1 and Hsp70 constitute a two-component feedback loop that regulates the yeast heat shock response. *eLife*. e31668. doi: 10.7554/eLife.31668.

Masser, A. E., et al. 2019. Cytoplasmic protein misfolding titrates Hsp70 to activate nuclear Hsf1. *eLife*. 2019 Sep 25;8:e47791. doi: 10.7554/eLife.47791

Sensing Day and Night: Circadian Rhythms

Patke, A., M. W. Young, and S. Axelrod. 2020. Molecular mechanisms and physiological importance of circadian rhythms. *Nat. Rev. Mol. Cell Biol*. **21**(2):67–84.

Sensing and Responding to the Physical Environment

Aragona, M., et al. 2013. A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. *Cell* **154**(5):1047–1059.

Dasgupta, I. and D. McCollum. 2019. Control of cellular responses to mechanical cues through YAP/TAZ regulation. *J. Biol. Chem.* **294**:17693–17706.

Dupont, S., et al. 2011. Role of YAP/TAZ in mechanotransduction. *Nature* **474**:179–183.

Ma, S., et al. 2019. The Hippo pathway: biology and pathophysiology. *Annu. Rev. Biochem.* **88**:577–604.

Sasaki, H. 2017. Roles and regulations of Hippo signaling during preimplantation mouse development. *Dev. Growth Differ*. **59**:12–20.

Zheng, Y., and D. Pan. 2019. The Hippo signaling pathway in development and disease. *Dev. Cell* **50**(3):264–282.

Chapter 22

Early Mammalian Development, Embryonic Stem Cells, and Induced Pluripotent Stem Cells

Ben-David, U., J. Nissenbaum, and N. Benvenisty. 2013. New balance in pluripotency: reprogramming with lineage specifiers. *Cell* **153**:939–940.

Gill, M., and A. Peters. 2018. Toward human egg-like cells *in vitro. Science* **362**:291–292.

Hikabe, O., et al.. 2016. Reconstitution in vitro of the entire cycle of the mouse female germ line. *Nature* **539**:299–303.

Li, M., and J. C. Izpisua Belmonte. 2018. Deconstructing the pluripotency gene regulatory network. *Nat. Cell Biol.* **20**:382–392.

Mallanna, S., and A. Rizzino. 2010. Emerging roles of microRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. *Dev. Biol.* **344**:16–25.

McNeish, J., et al. 2015. From dish to bedside: lessons learned while translating findings from a stem cell model of disease to a clinical trial. *Cell Stem Cell* **17**:8–10.

Mertens, J. et al. 2018. Aging in a dish: iPSC-derived and directly induced neurons for studying brain aging and age-related neurodegenerative diseases. *Ann. Rev. Genet.* **52**:271–293.

Pagliuca, F., et al. 2014. Generation of functional human pancreatic β cells in vitro. *Cell* **159**:428–439.

Robinton, D., and G. Daley. 2014. The promise of induced pluripotent stem cells in research and therapy. *Nature* **481**:295–305.

Shahbazi, M., E. Siggia, and M. Zernicka-Goetz. 2019. Self-organization of stem cells into embryos: a window on early mammalian development. *Science* **364**:948–951.

Soldner, F., and R. Jaenisch. 2018. Stem cells, genome editing, and the path to translational medicine. *Cell* **175**:615–632.

Theunissen, T., and R. Jaenisch. 2017. Mechanisms of gene regulation in human embryos and pluripotent stem cells. *Development* **144**:4496–4509.

Wu, J., A. Ocampo, and J. C. Izpisua Belmonte. 2016. Cellular metabolism and induced pluripotency. *Cell* **166**:1371–1385.

Stem Cells and Niches in Multicellular Organisms

Aichinger, E., et al. 2012. Plant stem cell niches. *Annu. Rev. Plant Biol.* **63**:615–636.

Blanpain, C., and E. Fuchs. 2014. Plasticity of epithelial stem cells in tissue regeneration. *Science* **344**:1243.

Gehart, H., and H. Clevers. 2019. Tales from the crypt: new insights into intestinal stem cells. *Nat. Rev. Gastroenterol. Hepatol.* **16**:19–34.

Goodell, M., H. Nguyen, and N. Shroyer. 2015. Somatic stem cell heterogeneity: diversity in the blood, skin and intestinal stem cell compartments. *Nat. Rev. Mol. Cell Biol.* **16**:299–309.

Heidstra, R., and S. Sabatini. 2014. Plant and animal stem cells: similar yet different. *Nat. Rev. Mol. Cell Biol.* **15**:301–312.

Ikeuchi, M., et al. 2019. Molecular mechanisms of plant regeneration. *Annu. Rev. Plant Biol*. **70**:377–406.

Kahney, E., J. Snedeker, and X. Chen. 2019. Regulation of Drosophila germline stem cells. *Curr. Opin. Cell Biol.* **60**:27–35.

Laurenti, E., and B. Göttgens. 2018. From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**:418–426.

Naik, S., et al. 2018. Two to tango: dialog between immunity and stem cells in health and disease. *Cell* **175**:908–920.

Pierre-Jerome, E., C. Drapek, and P. Benfey. 2018. Regulation of division and differentiation of plant stem cells. *Annu. Rev. Cell Dev. Biol*. **34**:289–310.

Shoshkes-Carmel, M., et al. 2018. Subepithelial telocytes are an important source of Wnts that supports intestinal crypts. *Nature* **557**:242–246.

Tanaka, E., and P. Reddien. 2011. The cellular basis for animal regeneration. *Dev. Cell* **21**:172–185.

Watcham, S., I. Kucinski, and B. Göttgens. 2019. New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood* **133**:1415–1426.

Mechanisms of Cell Polarity and Asymmetric Cell Division

Butler, M. T., and J. B. Wallingford. 2017. Planar cell polarity in development and disease. *Nat. Rev. Mol. Cell Biol.* **18**:375–388.

Cabernard, C., and C. Q. Doe. 2009. Apical/basal spindle orientation is required for neuroblast homeostasis and neuronal differentiation in *Drosophila*. *Dev. Cell* **17**:134–141.

Campanale, J. P., T. Y. Sun, and D. J. Montell. 2017. Development and dynamics of cell polarity at a glance. *J. Cell Sci.* **130**:1201–1207.

Chiou, J., M. K. Balasubramanian, and D. J. Lew. 2017. Cell polarity in yeast. *Ann. Rev. Cell Dev. Biol.* **33**:77–101.

Devenport, D. 2014. The cell biology of planar cell polarity. *J. Cell Biol.* **207**:171–179.

Homem, C. C. F., and J. A. Knoblich. 2012. *Drosophila* neuroblasts: a model for stem cell biology. *Development* **139**:4297–4310.

Knoblich, J. A. 2008. Mechanisms of asymmetric stem cell division. *Cell* **132**:583–597.

Lang, C. F., and E. Munro. 2017. The PAR proteins; from molecular circuits to dynamic self-stabilizing cell polarity. *Development* **144**:3405–3416.

Mellman, I., and W. J. Nelson. 2008. Coordinated protein sorting, targeting and distribution in polarized cells. *Nat. Rev. Mol. Cell Biol.* **9**:833–845.

Rodriguez-Boulan, E., and I. G. Macara. 2014. Organization and execution of the epithelial polarity programme. 2014. *Nat. Rev. Mol. Cell Biol.* **15**:225–242.

Shivas, J. M., et al. 2010. Polarity and endocytosis: reciprocal regulation. *Trends Cell Biol.* **20**:445–452.

Siller, K. H., and C. Q. Doe. 2009. Spindle orientation during asymmetric cell division. *Nat. Cell Biol.* **11**:365–374.

St. Johnston, D., and J. Ahringer. 2010. Cell polarity in eggs and epithelia: parallels and diversity. *Cell* **141**:757–774.

Cell Death and Its Regulation

Adams, J. M., and S. Cory. 2007. Bcl-2-regulated apoptosis: mechanism and therapeutic potential. *Curr. Opin. Immunol.* **19**:488–496.

Ashkenazi, A., and G. Salvesen. 2014. Regulated cell death: signaling and mechanisms. *Annu. Rev. Cell Dev. Biol.* **30**:337–356.

Cory, S. 2018. Phosphatidylserine hide-and-seek. *Proc. Nat'l. Acad. Sci. USA* **115**:12092–12094.

Dondelinger, Y., et al. 2016. An evolutionary perspective on the necroptotic pathway. *Trends Cell Biol*. **26**:721–732.

Galluzzi, L., et al. 2018. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death. *Cell Death Differ.* **25**:486–541.

Gudipaty, S., et al. 2018. Unconventional ways to live and die: cell death and survival in development, homeostasis, and disease. *Annu. Rev. Cell Dev. Biol*. **34**:311–332.

Luna-Vargas, M., and J. Chipuk. 2016. Physiological and pharmacological control of BAK, BAX, and beyond. *Trends Cell Biol.* **26**:906–917.

Newton, K., and G. Manning. 2016. Necroptosis and inflammation. *Annu. Rev. Biochem*. **85**:743–763.

Shan, B., et al. 2018. Necroptosis in development and diseases. *Genes Dev*. **32**:327–340.

Chapter 23

Neurons and Glia: Building Blocks of the Nervous System

Allen, N. J., 2014. Astrocyte regulation of synaptic behavior. *Annu. Rev. Dev. Cell Biol.* **30**:439–463.

Khakh, B., and M. Sofroniew. 2015. Diversity of astrocyte functions and phenotypes in neural circuits. *Nat. Neurosci.* **18**:942–952.

Kriegstein, A., and A. Alvarez-Buylla. 2009. The glial nature of embryonic and adult neural stem cells. *Annu. Rev. Neurosci.* **32**:149–184.

National Institutes of Health. The brain initiative. http://braininitiative.nih.gov/index.htm

Paridaen, J. T. M. L., and W. B. Huttner. 2014. Neurogenesis during development of the vertebrate central nervous system. *EMBO Rep.* **15**:351–364**.**

Voltage-Gated Ion Channels and the Propagation of Action Potentials

Catterall, W. A. 2014. Structure and function of voltage-gated sodium channels at atomic resolution. *Exp. Physiol.* **99**:35–51.

Hille, B. 2001. *Ion channels of excitable membranes*, 3d ed. Sinauer Associates.

Jouhaux, E., and R. Mackinnon. 2005. Principles of selective ion transport in channels and pumps. *Science* **310**:1461–1465.

Long, S. B., X. Tao, E. Campbell, and R. MacKinnon. 2007. Atomic structure of a voltage-dependent K^+ channel in a lipid membrane-like environment. *Nature* **450**:376–382.

Neher, E., and B. Sakmann. 1992. The patch clamp technique. *Sci. Am.* **266**:28–35.

Steinberg, E. E., D. J. Christoffel, K. Deisseroth, and R. C. Malenka. 2015. Illuminating circuitry relevant to psychiatric disorders with optogenetics. *Curr. Opin. Neurobiol.* **30**:9–16.

Communication at Synapses

Burden, S. J. 2011. Snapshot: neuromuscular junction. *Cell* **144**:826–826.e1.

Shen, K., and P. Scheiffele. 2010. Genetics and cell biology of building specific synaptic connectivity. *Annu. Rev. Neurosci.* **33**:473–507.

Siksou, L., A. Triller, and S. Marty. 2011. Ultrastructural organization of presynaptic terminals. *Curr. Opin. Neurosci.* **21**:261–268.

Sudhof, T. C. 2013. Neurotransmitter release: the last millisecond in the life of a synaptic vesicle. *Neuron* **80**:675–680.

Tang, L., et al. 2014. Structural basis for Ca^{2+} selectivity of a voltage-gated calcium channel. *Nature* **505**:56–62.

Unwin, N. 2005. Refined structure of the nicotinic acetylcholine receptor at 4Å resolution. *J. Mol. Biol.* **346**:967–989.

Sensing the Environment: Touch, Pain, Taste, and Smell

Buck, L., and R. Axel. 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**:175–187.

DeMaria, S., and J. Ngai. 2010. The cell biology of smell. *J. Cell Biol.* **191**:443–452.

Liao, M., E. Cao, D. Julius., and Y. Cheng. 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **304**:107–112.

Volkers, L., Y. Mochioukhi, and B. Coste. 2015. Piezo channels: from structure to function. *Eur. J. Physiol.* **467**:95–99.

Yamolinsky, D. A., C. S. Zuker, and N. J. P. Ryba. 2009. Common sense about taste: from mammals to insects. *Cell* **139**:234–244.

Zimmerman, A., L. Bai, and D. D. Ginty. 2014. The gentle touch receptors of mammalian skin. *Science* **346**:950–954.

Forming and Storing Memories

Bailey, C. H., and E. R. Kandel. 2008. Synaptic remodeling, synaptic growth and the storage of long-term memory in *Aplysia*. *Prog. Brain Res.* **169**:179−198.

Ho, V. M., J. A. Lee, and K. C. Martin. 2011. The cell biology of synaptic plasticity. *Science* **334**:623−628.

Huganir, R. L., and R. A. Nicoll. 2013. AMPARs and synaptic plasticity: the last 25 years. *Neuron* **80**:704−717.

Kandel, E. R., Y. Dudai, and M. R. Mayford. 2014. The molecular and systems biology of memory. *Cell* **157**:163−186.

Nabavi, S., et al. 2014. Engineering a memory with LTD and LTP. *Nature* **511**:348−352.

Ramirez, L., et al. 2013. Creating a false memory in the hippocampus. *Science* **341**:387−391.

Shepherd, J. D., and Huganir, R. L. 2007. The cell biology of synaptic plasticity: AMPA receptor trafficking. *Annu. Rev. Cell Dev. Biol.* **23**:613–643.

Chapter 24

Overview of Host Defenses

Akira, S., K. Kiyoshi Takeda, and T. Kaisho. 2001. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat. Immunol.* **2**:675–680.

Heyman, B. 2000. Regulation of antibody responses via antibodies, complement, and Fc receptors. *Annu. Rev. Immunol.* **18**:709–737.

Immunoglobulins: Structure and Function

Amzel, L. M., and R. J. Poljak. 1979. Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* **48**:961–997.

Williams, A. F., and A. N. Barclay. 1988. The immunoglobulin superfamily—domains for cell-surface recognition. *Annu. Rev. Immunol.* **6**:381–405.

Generation of Antibody Diversity and B-Cell Development

Hozumi, N., and S. Tonegawa. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Nat'l. Acad. Sci. USA* **73**:3628–3632.

Jung, D., et al. 2006. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* **24**:541–570.

Schatz, D. G., M. A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene, RAG-1. *Cell* **59**:1035–1048.

The MHC and Antigen Presentation

Bjorkman, P. J., et al. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**:506–512.

Blees, A., et al. 2017. Structure of the human MHC-I peptide-loading complex. *Nature.* **551**:525–528.

Brown, J. H., et al. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**:33–39.

Peters, P. J., et al. 1991. Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature* **349**:669–676.

Rudolph, M. G., R. L. Stanfield, and I. A. Wilson. 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* **24**:419–466.

Zinkernagel, R. M., and P. C. Doherty. 1974. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature* **248**:701–702.

T cells, T-Cell Receptors, and T-Cell Development

Chai, J. 2020. Atomic-resolution view of complete TCR-CD3 revealed. *Protein & Cell* **11**:158–160.

Dembic, Z., et al. 1986. Transfer of specificity by murine alpha and beta T-cell receptor genes. *Nature* **320**:232–238.

Kisielow, P., et al. 1988. Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature $\mathrm{CD4}^{+}8^{+}$ thymocytes. *Nature 333*:742–746.

Miller, J. F. 1961. Immunological function of the thymus. *Lancet* **30**(2):748–749.

Collaboration of Immune-System Cells in the Adaptive Response

Banchereau, J. 2002. The long arm of the immune system. *Sci. Am.* **287**:52–59.

Plotkin, S. A., and W. A. Orenstein. 2003. *Vaccines*, 4th ed. Saunders.

Rathinam, V. A. K., and K. A. Fitzgerald. (2016). Inflammasome complexes: emerging mechanisms and effector functions. *Cell* **165**:792–800.

Shi, J., W. Gao, and F. Shao. (2017). Pyroptosis: gasdermin-mediated programmed necrotic cell death. *Trends Biochem. Sci.* **42**:245–254.

Smith, Jane S. 1990. *Patenting the sun: polio and the Salk vaccine.* Morrow.

20 years of HIV science. 2003. *Nat. Med.* **9**:803–843. A collection of opinion pieces on the prospects for an AIDS vaccine.

Chapter 25

Introduction

Weinberg, R. A. 2006. *The Biology of Cancer.* Garland Science.

How Tumor Cells Differ From Normal Cells

De Bock, K., M. Mazzone, and P. Carmeliet. 2011. Antiangiogenic therapy, hypoxia, and metastasis: risky liaisons, or not? *Nat. Rev. Clin. Oncol*. **8**(7):393–404.

Desgrosellier, J. S., and D. A. Cheresh. 2010. Integrins in cancer: biological implications and therapeutic opportunities. *Nat. Rev. Cancer* **10**(1):9–22.

Giancotti, F. G. 2013. Mechanisms governing metastatic dormancy and reactivation. *Cell* **155**(4):750–764.

Grivennikov, S. I., F. R Greten, and M. Karin. 2010. Immunity, inflammation, and cancer. *Cell* **140**(6):883–899.

Hanahan, D., and R. A. Weinberg. 2011. Hallmarks of cancer: the next generation. *Cell* **144**:646–674.

Joyce, J. A., and J. W. Pollard. 2009. Microenvironmental regulation of metastasis. *Nat. Rev. Cancer* **9**(4):239–252.

Korbel, J. O., and P. J. Campbell. 2013. Criteria for inference of chromothripsis in cancer genomes. *Cell* **152**(6):1226–1236.

Nguyen, D. X., P. D. Bos, and J. Massagué. 2009. Metastasis: from dissemination to organ-specific colonization. *Nat. Rev. Cancer* **9**(4):274–284.

Pfau, S. J., and A. Amon. 2012. Chromosomal instability and aneuploidy in cancer: from yeast to man. *EMBO Rep.* **13**(6):515–527.

Sethi, N., and Y. Kang. 2011. Unravelling the complexity of metastasis—molecular understanding and targeted therapies. *Nat. Rev. Cancer* **11**(10):735–748.

Thiery, J. P., et al. 2009. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**(5):871–890.

Vander Heiden, M. G., L. C. Cantley, and C. B. Thompson. 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**(5930):1029–1033.

Genetic and Genomic Basis of Cancer

Alexandrov. L. B., et al. 2020. The repertoire of mutational signatures in human cancer. *Nature* **258**:94.

Bunting, S. F., and A. Nussenzweig. 2013. End-joining, translocations and cancer. *Nat. Rev. Cancer* **13**(7):443–454.

Curtin, N. J. 2012. DNA repair dysregulation from cancer driver to therapeutic target. *Nat. Rev. Cancer* **12**(12):801–817.

Daley, J. M., and P. Sung. 2014. 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. *Mol. Cell Biol.* **34**(8):1380–1388.

Dawson, M. A., and T. Kouzarides. 2012. Cancer epigenetics: from mechanism to therapy. *Cell* **150**(1):12–27.

Garraway, L. A., and E. S. Lander. 2013. Lessons from the cancer genome. *Cell* **153**(1):17–37.

Grisendi, S., and P. P. Pandolfi. 2005. Two decades of cancer genetics: from specificity to pleiotropic networks. *Cold Spring Harb. Symp. Quant. Biol*. **70**:83–91.

ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. 2020. Pan-cancer analysis of whole genomes. *Nature* **258**:82.

Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol*. **7**:335–346.

Loechler, E. L. 2002. Environmental carcinogens and mutagens. In *Encyclopedia of Life Sciences*. Nature Publishing.

Lujambio, A., and S. W. Lowe. 2012. The microcosmos of cancer. *Nature* **482**(7385):347–355.

Moody, C. A., and L. A. Laimins. 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nat. Rev. Cancer* **10**(8):550–560.

Morrow, P. K., and G. N. Hortobagyi. 2009. Management of breast cancer in the genome era. *Annu. Rev. Med*. **60**:153–165.

Sellers, W. R. 2011. A blueprint for advancing genetics-based cancer therapy. *Cell* **147**(1):26–31.

Vogelstein, B., et al. 2013. Cancer genome landscapes. *Science* **339**:1546–1558. doi: 10.1126/science.1235122.

Wogan, G. N., et al. 2004. Environmental and chemical carcinogenesis. *Semin. Cancer Biol.* **14**:473–486.

Dysregulation of Cell Growth and Developmental Pathways Initiates Tumorigenesis

Dang, C. V. 2012. MYC on the path to cancer. *Cell* **149**(1):22–35.

Holderfield, M., et al. 2014. Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond. *Nat. Rev. Cancer* **14**(7):455–467.

Jiang, J., and C. Hui. 2008. Hedgehog signaling in development and cancer. *Dev. Cell* **15**(6):801–812.

Pickup, M., S. Novitskiy, and H. L. Moses. 2013. The roles of TGFβ in the tumour microenvironment. *Nat. Rev. Cancer* **13**(11):788–799.

Pylayeva-Gupta, Y., E. Grabocka, and D. Bar-Sagi. 2011. RAS oncogenes: weaving a tumorigenic web. *Nat. Rev. Cancer* **11**(11):761–774.

Shaw, A. T., et al. 2013. Tyrosine kinase gene rearrangements in epithelial malignancies. *Nat. Rev. Cancer* **13**(11):772–787.

Bieging, K. T., S. S. Mello, and L. D. Attardi. 2014. Unravelling mechanisms of p53-mediated tumour suppression. *Nat. Rev. Cancer* **14**(5):359–370.

Burkhart, D. L., and J. Sage. 2008. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* **8**(9):671–682.

Malumbres, M., and M. Barbacid. 2001. To cycle or not to cycle: a critical decision in cancer. *Nat. Rev. Cancer* **1**:222–231.

Manning, A. L., and N. J. Dyson. 2012. RB: mitotic implications of a tumour suppressor. *Nat. Rev. Cancer* **12**(3):220–226.

Evasion of Programmed Cell Death and Immune Surveillance Processes

Chen, D., and Mellman, I.. 2013. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**:1–10.

Cotter, T. G. 2009. Apoptosis and cancer: the genesis of a research field. *Nat. Rev. Cancer* **9**(7):501–507.

Fritz, J., and Lenardo, M. 2019. Development of immune checkpoint therapy for cancer. *J. Exp. Med.* **216**:1244–1254.

Hu, Z., P. A. Ott, and C. J. Wu. 2018. Towards personalized, tumour-specific, therapeutic vaccines for cancer. *Nat. Rev. Immunol.* **18**:168–182.

June, C. H., et al. 2018. CAR T cell immunotherapy for human cancer. *Science* **359**:1361–1365.

Schumacher, T. N., W. Scheper, and P. Kvistborg. 2019. Cancer neoantigens *Annu. Rev. Immunol.* **37**:173–200.

Schumacher, T. N., and R. Schreiber. 2018. Neoantigens in cancer immunotherapy. *Science* **348**:69–74.

Shaulian, E., and M. Karin. 2002. AP-1 as a regulator of cell life and death. *Nat. Cell Biol.* **4**:E131–E136.

Srivastava, S., and Riddell, S. 2015. Engineering CAR-T cells: Design concepts. *Trends Immunol.* **36**:494–502.

With its acclaimed author team, cutting-edge content, emphasis on medical relevance, and coverage based on key experiments, *Molecular Cell Biology* has justly earned an impeccable reputation as an exciting and authoritative textbook. Avoiding an encyclopaedic approach, the book grounds its coverage in the experiments that define our understanding of cell biology, engaging students with the exciting breakthroughs that define the field's history and point to its future. The authors, all world-class researchers and teachers, incorporate medically relevant examples where appropriate to help illustrate the connections between cell biology and health and human disease.

Achieve

The ninth edition is now available in Achieve. Achieve supports educators and students throughout the full range of instruction, including assets suitable for pre-class preparation, in-class active learning, and post-class study and assessment. The pairing of a powerful new platform with outstanding biology content provides an unrivalled learning experience.

Features of Achieve include:

- A design guided by learning science research. Co-designed through extensive collaboration and testing by both students and faculty.
- A learning path of powerful content including pre-class, in-class, and post-class activities and assessments.
- Easy integration and gradebook sync with iClicker classroom engagement solutions.
- Powerful content such as the new Smart Biology Animated Videos, which provide cutting-edge 3-D tools for students to help them understand cell biology at a structural level.

Find out more about Achieve at macmillanihe.com

macmillan
international
HIGHER EDUCATION

Cover image © micro_photo / Getty Images